

# THE JOURNAL of the Acoustical Society of America

Vol. 105, No. 6

June 1999

**SOUNDINGS SECTION**

<b>ACOUSTICAL NEWS—USA</b>	2999
USA Meetings Calendar	2999
<b>ACOUSTICAL NEWS—INTERNATIONAL</b>	3001
International Meetings Calendar	3001
<b>BOOK REVIEWS</b>	3005

**GENERAL LINEAR ACOUSTICS [20]**

Sound radiation from a line forced perforated elastic sandwich panel	I. David Abrahams	3009
Surface waves over bead layers	Jean-François Allard, Michel Henry, Julian Tizianel, Luc Kelders, Walter Lauriks	3021
Experimental verification of the opposite effect of fluid loading on the velocity of dilatational waves in thin plates and rods	Waled Hassan, Peter B. Nagy	3026
Active impedance control within a cylindrical waveguide for generation of low-frequency, underwater plane traveling waves	James J. Finneran, Mardi C. Hastings	3035
Wavefront and group velocity in relaxing and bubbly fluids	Johan L. Leander	3044
Parabolic equations for gravity and acousto-gravity waves	Joseph F. Lingeitch, Michael D. Collins, William L. Siegmann	3049
Acoustics of fluid-filled boreholes with pipe: Guided propagation and radiation	V. N. Rama Rao, J. K. Vandiver	3057
The arccos and Lommel formulations—Approximate closed-form diffraction corrections	Charles J. Daly, N. A. H. K Rao	3067
Directivity of a uniform-strength, continuous circular-arc source phased to the spatial position of its diameter	William Thompson, Jr.	3078

**NONLINEAR ACOUSTICS [25]**

A new theoretical approach to the analysis of nonlinear sound beams using the oblate spheroidal coordinate system	Tomoo Kamakura, Tsuneo Ishiwata, Kazuhisa Matsuda	3083
Nonlinear pulse propagation in arbitrarily dispersive media: Tube waves in permeable formations	David Linton Johnson	3087
Static friction threshold effects on hysteresis and frequency spectra in rocks	Abraham Kadish	3097
The secondary field of a parametric source following interaction with sea surface	Lian Sheng Wang, Brian V. Smith, Rodney Coates	3108

(Continued)

## CONTENTS—Continued from preceding page

**AEROACOUSTICS, ATMOSPHERIC SOUND [28]**

- Sound scattering cross-section in a stratified moving atmosphere V. E. Ostashev, G. H. Goedecke, R. Wood, H. Auvermann, S. F. Clifford 3115

**UNDERWATER SOUND [30]**

- Continuous monitoring of fish in a shallow channel using a fixed horizontal sonar Bjarke Pedersen, Mark V. Trevorrow 3126
- Time-evolution modeling of seafloor scatter. I. Concept E. Pouliquen, O. Bergem, N. G. Pace 3136
- Time-evolution modeling of seafloor scatter. II. Numerical and experimental evaluation O. Bergem, E. Pouliquen, G. Canepa, N. G. Pace 3142
- Acoustic hot-spots and breaking wave noise in the surf zone Grant B. Deane 3151
- Target localization for a three-dimensional multibeam sonar imaging system Jules S. Jaffe 3168
- Iterative time reversal in the ocean H. C. Song, W. A. Kuperman, W. S. Hodgkiss, T. Akal, C. Ferla 3176
- A test of basin-scale acoustic thermometry using a large-aperture vertical array at 3250-km range in the eastern North Pacific Ocean Peter F. Worcester, Bruce D. Cornuelle, Matthew A. Dzieciuch, Walter H. Munk, Bruce M. Howe, James A. Mercer, Robert C. Spindel, John A. Colosi, Kurt Metzger, Theodore G. Birdsall, Arthur B. Baggeroer 3185
- Comparisons of measured and predicted acoustic fluctuations for a 3250-km propagation experiment in the eastern North Pacific Ocean John A. Colosi, Edward K. Scheer, Stanley M. Flatté, Bruce D. Cornuelle, Matthew A. Dzieciuch, Walter H. Munk, Peter F. Worcester, Bruce M. Howe, James A. Mercer, Robert C. Spindel, Kurt Metzger, Theodore G. Birdsall, Arthur B. Baggeroer 3202
- Geoacoustic inversion via local, global, and hybrid algorithms Mark R. Fallat, Stan E. Dosso 3219
- Near-caustic behavior in a 270-km acoustical experiment Jeffrey A. Simmen, Stanley M. Flatté, Harry A. DeFerrari, Hien Nguyen, Neil J. Williams 3231
- Mirages in shallow water matched field processing G. L. D'Spain, J. J. Murray, W. S. Hodgkiss, N. O. Booth, P. W. Schey 3245

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

- A new calculation procedure for spatial impulse responses in ultrasound Jørgen Arendt Jensen 3266

**TRANSDUCTION [38]**

- Three-dimensional beam directivity of phase-steered ultrasound Shi-Chang Wooh, Yijun Shi 3275
- Robust control of a sensorless bass-enhanced moving-coil loudspeaker system Mingsian R. Bai, Hsinping Wu 3283
- Propagation characteristics of leaky Lamb waves in a liquid-loaded double-layered substrate consisting of a thin piezoelectric ceramic plate and a thin glass plate Kohji Toda, Kenji Motegi 3290

**STRUCTURAL ACOUSTICS AND VIBRATION [40]**

- Modal characteristics of in-plane vibration of rectangular plates N. H. Farag, J. Pan 3295
- Frequency constants for transverse vibration of annular disks Thomas B. Gabrielson 3311

## CONTENTS—Continued from preceding page

Analysis of acousto-ultrasonic characteristics for an isotropic thin plate	T. Liu, K. M. Liew, S. Kitipornchai, G. Wang	3318
<b>NOISE: ITS EFFECTS AND CONTROL [50]</b>		
Efficiency of a noise barrier with an acoustically soft cylindrical edge for practical use	Tomonao Okubo, Kyoji Fujiwara	3326
Demographic and attitudinal factors that modify annoyance from transportation noise	Henk M. E. Miedema, Henk Vos	3336
<b>ARCHITECTURAL ACOUSTICS [55]</b>		
Speech intelligibility model including room and loudspeaker influences	L. Faiget, R. Ruiz	3345
<b>ACOUSTIC SIGNAL PROCESSING [60]</b>		
On direction finding of an emitting source from time delays	Baruch Berdugo, Miriam A. Doron, Judith Rosenhouse, Haim Azhari	3355
Model-based dispersive wave processing: A recursive Bayesian solution	J. V. Candy, D. H. Chambers	3364
Detection of phase- or frequency-modulated signals in reverberation noise	Valérie Carmillet, Pierre-Olivier Amblard, Geneviève Jourdain	3375
Imposing pattern nulls on broadband array responses	Peter J. Kootsookos, Darren B. Ward, Robert C. Williamson	3390
<b>PHYSIOLOGICAL ACOUSTICS [64]</b>		
Adaptation of $2f_1-2f_2$ distortion product otoacoustic emission in young-adult and old CBA and C57 mice	X.-M. Sun, D. O. Kim	3399
The "inverse problem" solved for a three-dimensional model of the cochlea. III. Brushing-up the solution method	Egbert de Boer, Alfred L. Nuttall	3410
Dolphin hearing: Relative sensitivity as a function of point of application of a contact sound source in the jaw and head region	B. Møhl, W. W. L. Au, J. Pawloski, P. E. Nachtigall	3421
<b>PSYCHOLOGICAL ACOUSTICS [66]</b>		
Intensity discrimination of Gaussian-windowed tones: Indications for the shape of the auditory frequency-time window	Nicolle H. van Schijndel, Tammo Houtgast, Joost M. Festen	3425
Speech intelligibility and localization in a multi-source environment	Monica L. Hawley, Ruth Y. Litovsky, H. Steven Colburn	3436
Acoustic pursuit of invisible moving targets by cats	Ralph E. Beitel	3449
Sound localization in noise in hearing-impaired listeners	Christian Lorenzi, Stuart Gatehouse, Catherine Lever	3454
Temporal integration of loudness in listeners with hearing losses of primarily cochlear origin	Søren Buus, Mary Florentine, Torben Poulsen	3464
System identification of feedback in hearing aids	Johan Hellgren, Thomas Lunner, Stig Arlinger	3481
<b>SPEECH PERCEPTION [71]</b>		
Missing-data model of vowel identification	Alain de Cheveigné, Hideki Kawahara	3497
<b>SPEECH PROCESSING AND COMMUNICATION SYSTEMS [72]</b>		
Identification of resynthesized /hVd/ utterances: Effects of formant contour	James M. Hillenbrand, Terrance M. Nearey	3509
<b>MUSIC AND MUSICAL INSTRUMENTS [75]</b>		
Three-dimensional mechanical admittance: Theory and new measurement method applied to the violin bridge	Xavier Boutillon, Gabriel Weinreich	3524

## CONTENTS—Continued from preceding page

<b>Behavior of piano-action in a grand piano. I. Analysis of the motion of the hammer prior to string contact</b>	Eiji Hayashi, Masami Yamane, Hajime Mori	3534
<b>Time-domain modeling and numerical simulation of a kettledrum</b>	Leïla Rhaouti, Antoine Chaigne, Patrick Joly	3545
<b>Preferred self-to-other ratios in choir singing</b>	Sten Ternström	3563
<b>BIOACOUSTICS [80]</b>		
<b>The underwater audiogram of the West Indian manatee (<i>Trichechus manatus</i>)</b>	Edmund R. Gerstein, Laura Gerstein, Steven E. Forsythe, Joseph E. Blue	3575
<b>LETTERS TO THE EDITOR</b>		
<b>Characteristics of thermoacoustic resonance in a ducted burner [28]</b>	Sangyeon Cho, Soogab Lee	3584
<b>INDEX TO VOLUME 105</b>		3588
<b>SUBJECT INDEX TO VOLUME 105</b>		3593
<b>AUTHOR INDEX TO VOLUME 105</b>		3617

**NOTES CONCERNING ARTICLE ABSTRACTS**

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

**Document Delivery:** Copies of journal articles can be ordered from the new *Articles in Physics* online document delivery service (URL: <http://www.aip.org/articles.html>).

# ACOUSTICAL NEWS—USA

**Elaine Moran**

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

**Editor's Note:** Readers of this Journal are asked 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.

## James Candy elected IEEE Fellow



ASA Fellow James. V. Candy was elected a Fellow of The Institute of Electrical and Electronics Engineers, Inc. (IEEE) "for contributions to model-based ocean acoustic signal processing." The IEEE Fellowship is bestowed upon a very limited number of Senior Members to recognize their worldwide achievements in electro and information technology.

Dr. Candy is Director of the Center for Advanced Signal & Image Sciences at Lawrence Livermore National Laboratory. He earned a B.S.E.E. at the University of Cincinnati and M.S.E. and Ph.D. degrees in Electrical

Engineering at the University of Florida in Gainesville.

James Candy is the current Chair of the ASA Interdisciplinary Technical Group on Signal Processing in Acoustics.

## ASME International honors ASA members

ASA Fellow Andrew F. Seybert, professor in the mechanical engineering department of the University of Kentucky, and ASA Member Jason T. Weissenburger, founder and president of Engineering Dynamics International, St. Louis, have been named Fellows of ASME International (The American Society of Mechanical Engineers). The Fellow grade is conferred upon an ASME International member with at least ten years active engineering practice who has made significant contributions to the field.

Andrew Seybert earned his doctorate from Purdue University, West Lafayette, IN. Jason Weissenburger earned his doctorate from Washington University in St. Louis, MO.

## ASA members receive NHCA Awards

The National Hearing Conservation Association (NHCA) announced recipients of awards made at their recent meeting in Atlanta.

Dennis Driscoll of Associates in Acoustics in Evergreen, CO, was awarded the Outstanding Lecture Award for his presentation entitled "Noise Control Survey Procedures."

William W. Clark of the Central Institute for the Deaf in St. Louis, MO, shared the Outstanding Poster Award with his coauthor, Nancy Nadler, for their poster presentation entitled "From the Lab to the Living Room: Are Noise Toys Hazardous to Hearing?"

Daniel L. Johnson of Bruel, Bertrand and Johnson in Provo, UT, received the NHCA Outstanding Hearing Conservationist award. Dr. Johnson currently serves as the ASA Standards Director.

## Award for Excellence in Noise Control Engineering

The Institute of Noise Control Engineering has established an Award for Excellence in Noise Control Engineering. An engraved plaque and an honorarium of \$3000 will be given to a person or group of persons from one organization who has demonstrated an outstanding application of noise control engineering to a product or process that contributes to a quieter environment.

It is essential for the nomination that the accomplishment be current and its excellence clearly demonstrated. The accomplishment must be new and must advance knowledge of noise control engineering technology. The application should have a wide practical significance to noise control engineering for a broad spectrum of the public.

Nominations for this award are presently being accepted. A nomination package is available from the Managing Director of the Institute of Noise Control Engineering—USA, P.O. Box 3206 Arlington Branch, Poughkeepsie, New York 12603; Telephone 1-914-463-0201; E-mail: [hq@inco.org](mailto:hq@inco.org). Nominations must be received by the INCE Managing Director by July 31, 1999.

## 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.

### 1999

- 6–7 June 1999 SEM Spring Conference, Cincinnati, OH [Katherine M. Ramsay, Conference Manager, Society for Experimental Mechanics, Inc., 7 School St., Bethel, CT 06801; Tel.: 203-790-6373; Fax: 203-790-4472; E-mail: [meetings@sem1.com](mailto:meetings@sem1.com)].
- 25–26 June 1999 Biomedical Imaging Symposium: Visualizing the Future of Biology and Medicine, Bethesda, MD [BECON Symposium Website (<http://www.nih.gov/grants/becon/meeting99/index.htm>)].
- 27–30 June ASME Mechanics and Materials Conference, Blacksburg, VA [Mrs. Norma Guynn, Dept. of Engineering Science and Mechanics, Virginia Tech, Blacksburg, VA 24061-0219; Fax: 540-231-4574; E-mail: [nguyenn@vt.edu](mailto:nguyenn@vt.edu); WWW: <http://www.esm.vt.edu/mmconf/>].
- 6–11 July 1999 Clarinetfest, Ostend, Belgium [International Clarinet Association, Keith Koons, Music Dept., Univ. of Central Florida, P.O. Box 161354, Orlando, FL 32816-1354].
- 30 Sept.–2 Oct. Seventh Annual Conference on the Management of the Tinnitus Patient, Iowa City, IA [Rich Tyler, Dept. of Otolaryngology, Head and Neck Surgery, The University of Iowa, 200 Hawkins Dr., #E230 GH, Iowa City, IA 52242-1078, Tel.: 319-356-2471; Fax: 319-353-6739, E-mail: [rich-tyler@uiowa.edu](mailto:rich-tyler@uiowa.edu), WWW: <http://www.medicine.uiowa.edu/otolaryngology/news/news.html>].
- 7–10 Oct. Symposium on Occupational Hearing Loss, Philadelphia, PA [American Institute for Voice and Ear Research, Attn: Barbara-Ruth Roberts, 1721 Pine St., Philadelphia, PA 19103, Tel.: 215-545-2068; Fax: 215-735-2725. Deadline for submission of abstracts: 1 May].
- 1–5 Nov. 138th meeting of the Acoustical Society of America, Columbus, OH [Acoustical Society of America, 500 Sunnyside Blvd., Woodbury, NY 11797, Tel.: 516-576-2360; Fax: 516-576-2377, E-mail: [asa@aip.org](mailto:asa@aip.org); WWW: [asa.aip.org](http://asa.aip.org)].
- 2–4 Dec. ACTIVE 99, Fort Lauderdale, FL [Institute of Noise Control Engineering, P.O. Box 3206 Arlington Branch, Poughkeepsie, NY 12603, Tel.: 914-462-4006; Fax: 914-463-020; E-mail: [INCEUSA@aol.com](mailto:INCEUSA@aol.com); [users.aol.com/inceusa/inco.html](http://users.aol.com/inceusa/inco.html)].
- 6–8 Dec. INTER-NOISE 99, Fort Lauderdale, FL [Institute of Noise Control Engineering, P.O. Box 3206 Arlington Branch, Poughkeepsie, NY 12603, Tel.: 914-462-4006; Fax: 914-463-020; E-mail: [INCEUSA@aol.com](mailto:INCEUSA@aol.com); [users.aol.com/inceusa/inco.html](http://users.aol.com/inceusa/inco.html)].

# BOOK REVIEWS

**James F. Bartram**

94 Kane Avenue, Middletown, Rhode Island 02842

*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).]*

## The New Stereo Soundbook, Second Edition

**Ron Streicher and F. Alton Everest**

*Audio Engineering Associates, Pasadena, CA.  
xvi+266 pp. Price: \$54.50.*

The transmission of sounds using the conversion of acoustical energy into electrical energy and then back again has been with us for over 130 years. The magnetic recording of these electrical manifestations of sonic energy is about 100 years old. The start of the practical electroacoustic recording of sounds for playback dates from the late 1920's. The growth of radio and the birth of the "talkies" spurred tremendous technological advancement in recording technology and also in microphone technology. The technique of recording became important as the tools for recording became more diverse.

Stereophonic recording, transmission and playback was "invented" by two different people at two different organizations. Alan Blumlein of EMI invented a two-channel stereophonic system and coincident microphone arrays that are the antecedents for today's two-channel stereo music recordings. Harvey Fletcher and his team at Bell Labs invented the three-channel stereophonic system that has evolved into the three front channels used for most motion picture soundtracks as well as for many home theater motion picture releases.

Both of these systems strived to create the illusion of "being there." Both systems depended, initially, on microphone selection, deployment and recording (or in the case of the Bell Labs work, careful transmission) technique. Both systems depend on the psychoacoustics of the human hearing system to achieve their goal.

Although the first edition of *The New Stereo Handbook* was published less than a decade ago, new release formats, as well as the growing sophistication of the listening public, provided the authors with sufficient new material to warrant, and almost require, a second edition. In its 14 chapters, the second edition of *The New Stereo Handbook* provides a concise yet fairly comprehensive review of the history of stereo recording and reproduction, microphones and microphone techniques, binaural recording, and the limitations, and the practical uses thereof, of the human auditory system.

The format of the book is logical and has a linear form, although some subjects are covered only briefly. After a brief review of early stereophonic efforts, the authors review the psychophysics of hearing and how these aspects of hearing are used by stereophonic techniques to create the illusion of sonic space and depth. Production techniques, including both the practical implications as well as what can be termed philosophical implications, are covered. Various stereo microphone techniques are also described. Other concepts such as auditory spaciousness, coloration, as well as the listening environment and the optimization of same are also examined.

Depending on the sophistication of the user, the progression one makes through and of the book can be either tutorial, review, or introductory in nature. The authors have provided a compact compendium of most of the major and minor facets of stereophonic recording and human hearing and perception. The first chapter of the book, "Early Stereophonics," introduces monaural, binaural, monophonic, stereophonic, and biphonic systems. A very short and brief history of stereophonics, the advances of the 1930's, and both Blumlein's and Fletcher's work are presented.

In the next chapter, "How Stereo Information is Conveyed," basic concepts of the physics and physical nature of sound are presented. Monophonic and stereophonic (two-channel) systems are then introduced in some detail, with attention to time arrival and level differences. The concepts of interaural coherence and spaciousness are introduced.

The human auditory system and how the brain perceives the various acoustical cues received at the ears is dealt with in Chapter 3, "Stereo and the Auditory System." The lingua franca of hearing is presented. The auditory system structure is very briefly covered, as is cochlear function and mechanics. Masking and auditory filters, binaural masking, perceptual space, binaural unmasking versus frequency, speech in noise, auditory canal transfer function(s), reflected sound and its audibility on music and speech, first wavefront arrival effects, audibility of echoes and localization in the medial plane as well as the vertical plane are succinctly reviewed.

Practical parameters in the recording of sound are then tackled in the chapter titled "Philosophical and Pragmatic Approaches to Stereo." Questions such as identifying the listener, what medium is to be used for the final delivery of the material, when and where is the listener listening (while shopping, at home, etc.), and why is the listener listening are discussed. Other decisions that should be addressed include are we recreating or creating an event and what is the perspective that the listener is to be presented with. Then there are the technical considerations—amplitude limits and limitations of both the recording and the playback systems, the dynamic range of the original material, if applicable, and the dynamic range in the listening environment, phase, bandwidth (spectrum) and distortion limits. Recording formats, sonic versus visual images and the importance of remembering that any human endeavor is a compromise of physical (technical) and philosophical factors—in this case the authors state "illusion or reality... you can't have both."

The next three chapters (5, 6, and 7) deal with, respectively, "Two-microphone Stereo Techniques," "Binaural Recording and Reproduction," and "Coincident-microphone Stereo Techniques." In Chapter 5, phantom center, phantom imaging, the "hole in the middle," reverberation, critical distance, and reverberation in mono versus stereo as well as basic stereophonic perspectives versus multiple microphone systems are discussed. Chapter 6 deals with equipment, recording reproduction via headphones and loudspeakers, the problem of crosstalk, and dummy head developments. In Chapter 7, intensity stereo, the "Blumlein" technique and stereo, polar patterns, XY stereo, microphone matching, the mid/side (MS) technique, advantages and disadvantages of each technique, as well as a critique of other major coincident microphone techniques as well as subjective evaluation of these techniques are clearly explained. A handy chart that presents microphone polar patterns versus azimuth angle, pickup arc ( $-3$  dB and  $-6$  dB), relative output at 90 and 180 degrees, angle(s) for zero output (if applicable), random energy efficiency (rejection), and distance factor (reach) allows for quick appraisal of the major microphone polar patterns.

The authors then spend a chapter on "Audibility of Reflections." A review of the audibility and filtering effects of reflections, in terms of different signals, angles and spectrums of several sounds is followed by a discussion of the audible effects of these—what and where in time and frequency are timbre, echo, spaciousness, multiple images, speech disturbance and loudness summation (temporal integration). These are graphically presented in three concise diagrams (from Floyd Toole). Comb filtering and critical bands as well as their influence and effect on coloration and spaciousness, the comb filtering inherent in stereo microphone pickups, common comb filter producers (such as floor bounce) and the inevitable comb filter produced by the stereo listening setup are presented for the reader to ponder, i.e., how to minimize and how to use these.

Chapter 9 then considers spaced microphone arrays and techniques. Common arrays—ORTF, NOS, OSS, and SASS—are reviewed as well as what these initials mean. The authors then develop the actual arrangement of the microphones and the effect of these on lateral imaging and phase (problems). Interference (comb filtering) are again reviewed. The "Decca Tree" array is discussed as well as the use of auxiliary, or flanking, microphones.

The authors then turn their attention to “Multimicrophone Techniques”—that is, those situations where panpots, nonreal time recording, and mixdown are used—what is colloquially known as studio multi-track recording. The authors then visit the basic “rules” of and for microphones—directional patterns, reverberation rejection (random energy efficiency), and the various  $X$ -to-1 ratios for the placement of microphones. A short review of audio design using a film analogy is presented to demonstrate how to build up a musical soundscape. Spot microphone technique as well as the techniques that can be used in a closed recording session, as opposed to a live session, and still create the illusion of a live performance, are also offered.

Chapter 11 discusses several of the pseudostereo processes and effects using monophonic source material. Among the processes reviewed are those that use frequency response differences (Janovsky, 1948), reverberation chamber with and without delay (Schroeder, 1958; Lockner and Keet, 1960), the use of comb filters (Lauridsen, 1954), and, phase shifting (Schroeder, 1961). Surprisingly, the authors do not mention some of the HRTF processes (such as SRS) in this discussion.

“Auditory Spaciousness” is discussed in Chapter 12 in the context of how a recording is perceived. Definition versus auditory spaciousness, delayed sound and spatial impression, reverberation (time and level) and spaciousness, spaciousness in mono, the effect of lateral reflections, as well as some of the psychoacoustic effects of and on spaciousness, are reviewed. The authors expand a bit on the importance of the level of the reflections, the spectrum of the reflections as well as the temporal characteristics of reflections on the spaciousness of a recording (or room).

A little bit more of the history of multi-dimensional and surround sound systems is covered in the next chapter. In 1939, “Fantasound” was the first commercial realization of a truly multi-channel sound system, by Disney Studios, for the now classic film “Fantasia.” This and other three-channel film formats, rear channel systems (such as Hafler, circa 1970), and the now ubiquitous four-channel film systems (based upon the old Sansui Quadraphonic patents) are described. The full four-channel “Quad” system (LF, RF, LR, and RR) favored by audiophiles using four-channel magnetic tape, as well as the matrixed versions (4:2:4) used for the ill-fated “Quad” systems of the late 1960’s and early 1970’s (which, as alluded to above, became very successful for commercial motion picture soundtrack use) are also examined. The authors discuss how this led to the development of consumer-based products for surround sound and ambience enhancement. The authors also ask and then discuss the questions how many channels are enough as well as how many channels do we need. From this a discussion of

Ambisonics (Gerzon circa 1970s) is launched with a fairly concise but complete discussion of this technique, including the tetrahedral microphone array used and the encoding and decoding of the signals captured by the four microphone capsules in the array. Note that the Ambisonic technique also captures information in the vertical plane, and this can be used for playback. Binaural systems, production techniques for surround sound, ambience recording, enhanced ambience recording (i.e., enhancing what is there), localization in recordings, monitoring, as well as surround sound and the illusion of reality are briefly touched upon to close this chapter.

The final chapter is devoted to, and is titled, “Optimizing the Listening Environment.” The authors state that good listening is comprised of three interlinked rings: equipment (both recording and reproduction), human perception, and, room acoustics. Expanding upon the brief discussion in Chapter 8, the authors introduce the reader to the fact that sound acts differently in an enclosure than outdoors, that the response of the room is influenced by its physical dimensions (which lead to room modes—axial, tangential and oblique), the need for low-frequency control, and that in the mid-high frequency range the effects of reflections (delay) can be substantial—both in and on the room as well as on the human auditory system. The point of this chapter is that don’t expect excellent listening results from excellent recordings and equipment if the room is badly asymmetrical and/or poorly treated with respect to reflections and low frequency control.

The back of the book features an appendix which contains, an introduction to as well as, Blumlein’s British patent #394,325—“Improvements in and relating to Sound-transmission, Sound-recording and Sound-reproducing Systems,” a reading of which is in itself illuminating and informative almost 70 years after its filing. A glossary and index follow.

To close, the new edition is larger than the original which makes for easier reading. The second edition also has a list of references following each chapter, that which, while brief, contain many useful books and articles that deal with the subject matter in greater depth.

*The New Stereo Handbook*, second edition, is a useful reference book in itself to not only practitioners of the recording arts, but also to the end-users, that is the listeners, of recordings.

NEIL A. SHAW  
*Menlo Scientific Acoustics, Inc.*  
*Topanga, California*

# Sound radiation from a line forced perforated elastic sandwich panel

I. David Abrahams

*Department of Mathematics, University of Manchester, Oxford Road, Manchester M13 9PL,  
United Kingdom*

(Received 16 April 1998; revised 31 October 1998; accepted 11 March 1999)

Composite barriers, consisting of thin plates separated by light matrix structures, are widely used for fuselage construction in the aircraft industry, and in partitions in the building trade. The acoustical properties of such materials can vary considerably by altering the interior geometry, and perforations can be added to one or both sides. With perforations the interior cavities can act as Helmholtz resonators, causing a substantial modification to the overall transmission and reflection properties of such barriers. Leppington [Proc. R. Soc. London, Ser. A **427**, 385–399 (1990)] devised an effective boundary condition for a perforated sandwich plate structure, valid in the limit of low frequency (acoustical waves long compared with the typical dimensions of the hole/cavity construction), and obtained transmission and reflection coefficients for infinite planar structures. This article investigates the radiation properties of perforated sandwich plates by examining a simple infinite one-dimensional model (employing Leppington's effective boundary condition) which is loaded by a line force or moment. The radiated far field, and unattenuated subsonic plate wave coefficients, are found explicitly, and are plotted over a range of frequencies for two physical configurations, namely an aluminum plate in water and in air. It is revealed that, unlike the usual thin plate equation, the model discussed herein has two bi-directional unattenuated plate waves, and for the structure in air the two waves are of similar magnitude over most frequencies. Surprisingly, these amplitudes are shown to become very large at a frequency below that of the structure's Helmholtz resonance frequency. Further, the field radiating into air is also significantly modified by the cavity/perforations well away from the Helmholtz frequency. © 1999 Acoustical Society of America. [S0001-4966(99)04006-0]

PACS numbers: 43.20.Fn, 43.30.Jx, 43.40.Dx, 43.40.Fz [CBB]

## INTRODUCTION

Sound transmission through panels has long been a subject of importance in a variety of applications, including room acoustics<sup>1</sup> and aeroacoustics. In the latter application, for instance in the fabrication of aircraft fuselages, the bodies are often of a sandwich construction composed of two thin-plates separated by stiffeners or ribs in between.<sup>2</sup> The transmission and reflection of sound waves through the sandwich structure is determined by the properties of the plates and the enveloping fluid. In aircraft, and in particular helicopter designs, it is usually the case that sound transmission is required to be minimized over a specific range of frequencies. Sound transmission can be ameliorated to some extent by adding perforations to one or both sides of the flexible structure.<sup>3,4</sup> Leppington<sup>5</sup> examined the simple but physically realistic model of a perforated sandwich panel; namely one constructed of a honeycomb cellular structure (usually made of thin foil or similar material) bounded by two planar thin elastic plates. Perforations were included on one side of the structure. Leppington's model can accommodate variations in size of the cavities, number of perforations as a fraction of the number of cavities, perforation hole size, etc. The only restriction on the model is that the wavelength of the acoustic waves is long compared with the length scales of the composite structure, i.e., cavity width and height, thickness of thin plates and perforation hole size. In this limit the sound field will experience the "averaged" effect of the

sandwich structure rather than scattering/radiation from the individual cells. Leppington exploited this fact to solve the transmission and reflection problem, of long waves incident on the sandwich structure, by the method of matched asymptotic expansions. He showed that the cavities with small holes behave as individual Helmholtz resonators which dramatically alter the pressure field on the perforated (upper) side of the plate near the resonance frequency. The upshot is that away from a resonance the barrier behaves as a single elastic plate (in the long wavelength limit), with fluid on both sides, but near the Helmholtz frequency the sound transmission is significantly reduced with almost all energy reflected. In the latter case it is as if there is no fluid above the barrier.

In his article, Leppington<sup>5</sup> concluded with the derivation of a new plate equation which included the effects of the cavities and perforations. This reproduced his results found by matching, and is a very useful approximate boundary condition that can be employed in more complicated model problems. In particular, it is important to understand how the sandwich panel scatters acoustic waves as well as its transmission and reflection properties. In the former case only can energy be fed into the flexural wave mode of the structure and this could have significant consequences on the sound field above and below the composite plate. Jones<sup>6</sup> discussed the canonical scattering problem of a semi-infinite sandwich panel attached to a rigid half-screen and irradiated by plane waves from below. Unfortunately, the semi-infinite geometry



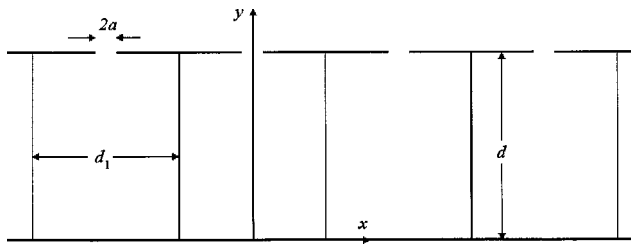


FIG. 1. The perforated plate construction.

adds significantly to the mathematical complexity, and the boundary value problem is reduced to a matrix Wiener-Hopf problem. Such equations cannot, in general, be solved exactly (see, for example, Ref. 7), although an approximate procedure has recently been presented by the author<sup>8</sup> which can be applied<sup>9</sup> to the class of matrices to which the one examined by Jones<sup>6</sup> belongs. However, Jones merely restricted attention to the case when the effect of the perforations almost vanished, and hence obtained a perturbation solution for which the standard scalar Wiener-Hopf technique is applicable. Jones did not obtain any numerical results.

On reexamining Jones' model problem using the method described in Ref. 9 by the author, and performing a numerical investigation over all values of the perforation parameter  $\tau$ , various results regarding the scattered field and traveling plate waves were deduced. These were sufficiently perplexing<sup>10</sup> to justify appraisal of a simpler radiation problem, which avoids the complexity of a Wiener-Hopf analysis. In particular, we demonstrate that Leppington's plate model<sup>5</sup> has two unattenuated wave modes for all nonzero values of the fluid/plate parameters. Apparently, this result has hitherto not been noticed by others. The problem to be discussed here consists of an infinite planar perforated elastic plate immersed above and below in a compressible stationary fluid. A time-harmonic line force provides the excitation and we seek the cylindrically radiating sound field plus the coupled surface wave terms. The paper is constructed as follows: in the following section the boundary value problem is posed, and in Sec. II the dispersion relation for unattenuated waves on the perforated plate is derived and then shown to have two real roots. In Sec. III the boundary value problem is solved by recourse to Fourier transforms, from which the radiated far-field and plate wave coefficients are deduced. In the final section we examine the behavior of these quantities for variations in the numerical values of parameters (consistent with physical models) and concluding remarks are offered.

## I. THE BOUNDARY VALUE PROBLEM

The boundary value problem is summarized in Fig. 1. It consists of a compressible, stationary fluid occupying the whole of space, outside of the plate, for which time-harmonic small disturbances are governed by

$$\frac{\partial^2 \phi}{\partial x^2} + \frac{\partial^2 \phi}{\partial y^2} + \frac{\partial^2 \phi}{\partial z^2} + \frac{\omega^2}{c_0^2} \phi = 0. \quad (1)$$

Here  $(x, y, z)$  are dimensional Cartesian coordinates,  $\phi(x, y, z, t)$  is the velocity potential of the acoustic perturba-

tion,  $\omega$  is the oscillation radian frequency, and  $c_0$  is the sound speed in the fluid at rest. The fluid pressure fluctuations,  $p$ , and the density fluctuations,  $\rho$ , are related to the potential via

$$p(x, y, z, t) = c_0^2 \rho(x, y, z, t) = \Re\{\rho_0 i \omega \phi(x, y, z) e^{-i \omega t}\}, \quad (2)$$

where  $\rho_0$  is the ambient fluid density, and similarly the velocity fluctuations are

$$\mathbf{u}(x, y, z, t) = \Re\{\nabla \phi(x, y, z) e^{-i \omega t}\}. \quad (3)$$

For brevity we omit the exponential time-factor and real part braces until the end.

Omitting all details regarding the derivation of an effective boundary condition for a perforated elastic plate (instead referring the reader to the article by Leppington<sup>5</sup>), we state the equations for a structure consisting of two thin elastic plates, each of bending stiffness  $B$ , density  $\rho_p$  and thickness  $h$ , separated by a light but acoustically rigid honeycomb structure of width  $d$ . Round perforations of hole size  $2a$  are included on the top plate, and for the sake of mathematical simplicity, the honeycomb arrangement is taken so that each cavity is a cuboid with height  $d$ , and widths in the  $x$  and  $z$  directions of  $d_1$ , as shown in Fig. 1. Then, for  $kd \ll 1$ ,  $kd_1 \ll 1$ ,  $kh \ll 1$ ,  $ka \ll 1$  where  $k = \omega/c_0$ , we find that the composite plate structure satisfies (see, e.g., Ref. 11)

$$\left\{ 2B \left( \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial z^2} \right)^2 - 2\omega^2 \rho_p h \right\} \eta(x, z) - F(x, z) = i\omega \rho_0 \phi(x, 0^-, z) - i\omega \rho_0 \phi(x, 0^+, z) \quad (4)$$

on  $y=0$ , in which  $\eta(x, z)$  is the deflection of the plate in the  $y$  (vertical direction) and the right hand side is the pressure jump across the plate [i.e.,  $0^+$  ( $0^-$ ) indicates that  $y$  goes to zero from above (below)]. The function  $F(x, z)$  is the imposed plate forcing to be specified later. The vertical velocity of the plate must equal the fluid velocity on the plate's underside, but the holes on the top plate means that this is not so in general for the fluid velocity above. Hence,

$$\eta(x, z) = \frac{i}{\omega} \frac{\partial \phi}{\partial y}(x, 0^-, z), \quad (5)$$

and the pressure above the plate is related to the jump in velocity across the structure<sup>5</sup> via

$$\frac{\partial \phi}{\partial y}(x, 0^+, z) - \frac{\partial \phi}{\partial y}(x, 0^-, z) + k\tau \phi(x, 0^+, z) = 0 \quad (6)$$

in which the perforation parameter  $\tau$  is

$$\tau = 2fkda/[2a - k^2V], \quad (7)$$

$f$  is the fraction of cells with apertures,  $V$  is the volume of each cell and  $d$ ,  $a$  are the plate spacing and perforation hole radius, respectively. For the cuboid cavities mentioned above, with a single hole in each one, we find

$$\tau = 2kda/[2a - k^2dd_1^2], \quad (8)$$

and note that the cavities resonate (Helmholtz resonance) when  $\tau \rightarrow \infty$ , or when

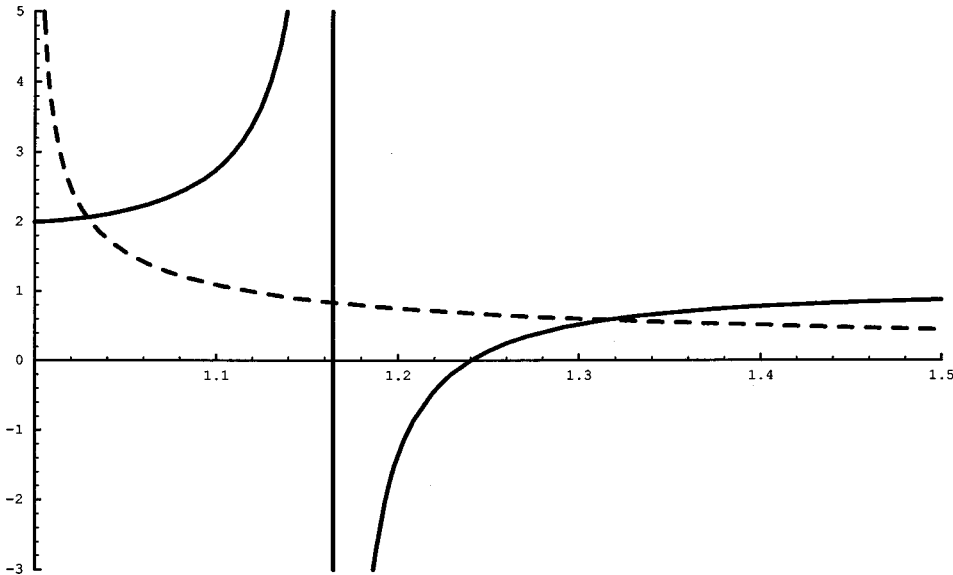


FIG. 2. Roots of the dispersion relation (23) given by the intersection of the function on the right (dashed line) and left (solid line) hand sides of (24).

$$k^2 = \frac{2a}{dd_1^2} \Rightarrow \omega = \sqrt{\frac{2a}{d} \frac{c_0}{d_1}}. \quad (9)$$

It is worth observing from (9) that the constraint

$$kd_1 \ll 1 \quad (10)$$

remains valid at or near a Helmholtz resonance if and only if

$$a \ll d. \quad (11)$$

Note that in the Leppington model used above the long-wavelength assumption allows the sandwich plate to be approximated by one of infinitesimal thickness [(4) and (6)]; by this means, all the details of the perforations and cell geometry are conveniently wrapped up in the single parameter  $\tau$ . As  $\tau \rightarrow 0$  then the perforation's influence vanishes and we recover the usual flexural thin-plate equation, across which the fluid velocity is continuous. As already stated, large  $\tau$  indicates that a Helmholtz resonance is approached, so that the perforations have a significant effect on the fluctuating field. Here we take  $\tau$  real, indicating that the flow in the cavity is lossless. However, if viscous motions through the apertures are assumed, then  $\tau$  will be complex, which will limit the maximum size of  $|\tau|$ . Other effects could also be included without difficulty, such as two or more different families of cavities, each yielding a different resonance value, i.e.,  $|\tau|$  will have multiple peaks when plotted over wave number.

As a final step in posing the problem, it is convenient to nondimensionalize using the acoustic wavelength and sound speed

$$(\bar{x}, \bar{y}, \bar{z}) = (kx, ky, kz), \quad \bar{\phi}_j = \frac{k}{c_0} \phi_j, \quad j = 1, 2, \quad (12)$$

where we write  $\phi_0$  ( $\phi_1$ ) as the acoustic potential below (above) the composite plate. Also, for the purposes of this article, we take a simple line force and bending moment acting upon the plate on  $\bar{x} = 0, \nabla \bar{z}$ . Thus, the forcing and plate geometry are independent of  $\bar{z}$  and so the boundary value problem may be considered as two-dimensional, the  $(\bar{x}, \bar{y})$ -plane. Summarizing, and dropping the overbars on the dimensionless variables for convenience, we have

$$\left( \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + 1 \right) \phi_0(x, y) = 0, \quad y \leq 0, \quad -\infty < x < \infty, \quad (13)$$

$$\left( \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + 1 \right) \phi_1(x, y) = 0, \quad y \geq 0, \quad -\infty < x < \infty, \quad (14)$$

and on the plate

$$\left( \frac{\partial^4}{\partial x^4} - \mu^4 \right) \frac{\partial \phi_0}{\partial y} + \alpha (\phi_1 - \phi_0) = A \delta(x) + B \delta'(x), \quad y = 0, \quad -\infty < x < \infty, \quad (15)$$

with the fluid velocity jump determined by

$$\frac{\partial \phi_1}{\partial y} - \frac{\partial \phi_0}{\partial y} + \tau \phi_1 = 0, \quad y = 0, \quad -\infty < x < \infty. \quad (16)$$

Here, the  $\mu$  and  $\alpha$  are

$$\mu^4 = \frac{\rho_p h c_0^4}{\omega^2 B}, \quad \alpha = \frac{\rho_0 c_0^5}{2 \omega^3 B}, \quad (17)$$

which, with  $\tau$  defined in (8), give three nondimensional fluid/structure coupling parameters for the problem. Note that  $\mu$  is the (dimensionless) wave number of free waves on the plate *in vacuo*. The first term on the right hand side of (15) is a line force of magnitude  $A$  and the second is a bending moment of strength  $B$  [ $\delta(x)$  is the usual generalized function and the dash denotes differentiation with respect to its argument]. These coefficients,  $A$  and  $B$ , can be chosen as desired. Equations (13)–(16), together with the condition of outgoing waves at infinity constitute the boundary value problem to be solved.

## II. UNATTENUATED PLATE WAVES

To examine free plate modes for the homogeneous coupled/plate system it is convenient to try the following form for a surface wave:

$$\phi_s = e^{isx - \gamma(s)|y|} \times \begin{cases} a, & y > 0, \\ b, & y < 0, \end{cases} \quad s \text{ real.} \quad (18)$$

Clearly (by direct substitution)  $\phi_s$  satisfies (13) and (14) if

$$\gamma^2(s) = s^2 - 1, \quad (19)$$

and so the wave will decay exponentially away from the plate if and only if

$$|s| > 1, \quad (20)$$

or  $\gamma(s) = (s^2 - 1)^{1/2} > 0$ . For  $|s| < 1$  then  $\gamma(s)$  is pure imaginary, and so (18) represents a plane wave which takes energy away from the plate. Hence, for study of surface waves in which the energy is constrained close to the plate the condition (20) is required. Substituting (18) into the homogeneous form of (15) yields

$$(s^4 - \mu^4)\gamma(s)b + \alpha(a - b) = 0 \quad (21)$$

and similarly (16) gives

$$-\gamma(s)(a + b) + \tau a = 0. \quad (22)$$

The two equations yield a nontrivial solution if and only if the determinant vanishes:

$$K(s) = [(s^4 - \mu^4)\gamma(s) - \alpha](\tau - \gamma(s)) + \alpha\gamma(s) = 0, \quad (23)$$

the left hand side of which is referred to as the dispersion relation for  $s$ . Rearranging gives an alternative form of constraint on  $s$ , namely

$$\frac{(s^4 - \mu^4)\gamma(s) - 2\alpha}{(s^4 - \mu^4)\gamma(s) - \alpha} = \frac{\tau}{\gamma(s)}. \quad (24)$$

The last expression is useful as it is simple to examine the limiting cases  $\tau \rightarrow 0$ ,  $\tau \rightarrow \infty$ . As  $\tau \rightarrow 0$  the equation reduces to

$$(s^4 - \mu^4)\gamma(s) - 2\alpha = 0, \quad (25)$$

whereas for  $\tau \rightarrow \infty$ ,

$$(s^4 - \mu^4)\gamma(s) - \alpha = 0, \quad (26)$$

which correspond to the usual thin-plate dispersion relations when the fluid is on both and one side, respectively (see, e.g., the article by Cannell<sup>12</sup>). This is as it should be; when  $\tau \rightarrow 0$  the perforations are closed yielding the usual plate equation, but as  $\tau \rightarrow \infty$  the perforation flow is large and so the effect of the fluid above the plate becomes negligible.

For all values of  $\mu$ ,  $\alpha$ , it is known that (25) and (26) each have one and only one positive real root (and one real negative root due to even behavior in  $s$ ) at  $s_2^*$ ,  $s_1^*$  say, respectively, where  $s_2^* > s_1^*$ . To see this, plot  $s^4 - \mu^4$  and  $c/\gamma(s)$ ,  $c > 0$ , against  $s$  for  $s \geq 1$ . As the first function is monotonic increasing with a finite value at  $s = 1$ , and the second is monotonic decreasing with value  $+\infty$  at  $s = 1$ , then there is clearly one unique point of intersection for arbitrary real values of  $\mu$  and  $c$  ( $> 0$ ). The full dispersion curve can now be examined by plotting the left hand side (unbroken curve) and right hand side (dashed curve) of (24) for  $s > 1$  in Fig. 2. Again, it is easy to prove that the form of the function on the left hand side is as shown for all values of  $\alpha$ ,  $\mu$  and the dashed curve is for a finite nonzero value of  $\tau$ . That is, there are two, and two only, points of intersection corre-

sponding to two roots  $s_1$ ,  $s_2$ , say, of the original dispersion equation (23). In fact, a little elementary calculus shows that the unbroken curve may, in  $|s| > 1$ , have a turning point at the location  $\sqrt{(2 + \sqrt{4 + 5\mu^4})/\sqrt{5}}$ , which is always to the left of the point of divergence in Fig. 2,  $s_1^*$ , but this does not alter the number of roots of (24). As  $\tau \rightarrow 0$  the dashed curve moves down, hence moving the intersection points (roots) to

$$\tau \rightarrow 0, \quad s_1 \rightarrow 1, \quad s_2 \rightarrow s_2^*. \quad (27)$$

As will be shown in the next section, when  $\tau = 0$  the root  $s_1 = 1$  in fact gives no contribution, hence recovering the single plate unattenuated mode for the case of no perforations. As  $\tau \rightarrow \infty$ , the dashed curve is pushed upwards, yielding roots at

$$\tau \rightarrow \infty, \quad s_1 \rightarrow s_1^*, \quad s_2 \sim \tau. \quad (28)$$

Again, it will be shown that the  $s_2$  plate wave term has vanishing coefficient as  $\tau \rightarrow \infty$ , hence recovering just the single unattenuated plate wave term,  $s_1^*$ .

It is to be stressed that for all values of  $\alpha$ ,  $\mu$ ,  $\tau$  the dispersion relation (24) yields two, and two only, positive real roots. One can be assigned to be a modified form of the unperforated plate wave number, but the other is entirely new and until now undiscovered. The purpose of the following sections is, for a typical boundary value problem, to see if the new plate wave term has a contribution to the sound field of the same order as the usual plate wave term (or perhaps larger) for typical values of fluid/plate parameters. If it has, then the effect of the perforations can be seen as significant on the scattered/radiated field in such models.

### III. SOLUTION OF THE BOUNDARY VALUE PROBLEM

To solve the system of equations (13)–(16) it is convenient to define the Fourier transform pair

$$F(s) = \int_{-\infty}^{\infty} f(x) e^{isx} dx, \quad (29)$$

$$f(x) = \frac{1}{2\pi} \int_{\mathcal{C}} F(s) e^{-isx} ds, \quad (30)$$

where  $\mathcal{C}$  is the contour running along the real line, from  $-\infty$  to  $+\infty$ , indented above any singularities on the negative half-line and below those occurring on the positive half-line (see Fig. 3). The reason for passing above (below) singularities on the left (right) hand part of the real line is that, for the given time dependence, wave terms will be recovered that are outward propagating and *not* incoming. The Fourier transform of the reduced wave equations (13) and (14) is

$$\frac{d^2 \Phi_j}{dy^2} (s, y) + (1 - s^2) \Phi_j(s, y) = 0, \quad j = 0, 1, \quad (31)$$

which has the bounded/outgoing solution

$$\Phi_0(s, y) = C(s) e^{\gamma(s)y}, \quad y \leq 0, \quad (32)$$

$$\Phi_1(s, y) = D(s) e^{-\gamma(s)y}, \quad y \geq 0, \quad (33)$$

with  $\gamma(s)$  as given by (19). Here  $\gamma(s)$  is made single valued by cutting the plane from  $s = 1$  to  $+\infty$  in the first quadrant

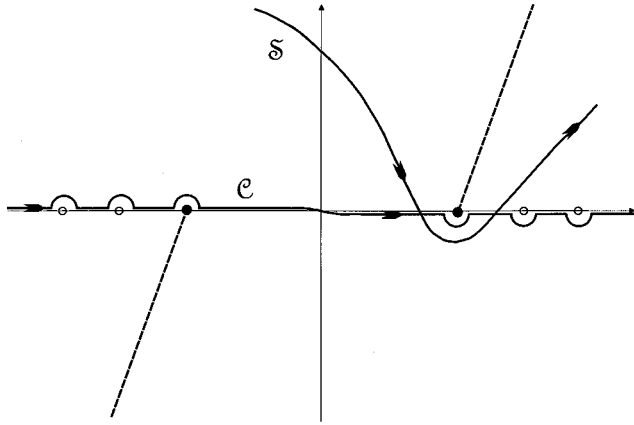


FIG. 3. The inverse transform path  $C$  of (30) passing above the branch point at  $-1$  and the simple poles at  $-s_1, -s_2$ , and below the branch point at  $+1$  and the poles at  $s_1, s_2$ . The branch cuts are denoted by dashed lines, and the steepest descent path for  $\theta > \pi/2$  is shown as  $S$ .

and  $s = -1$  to  $-\infty$  in the third quadrant, as shown by the dashed lines in Fig. 3, and choosing  $\gamma(0) = -i$ . Then  $\gamma(s)$  is negative imaginary for  $s$  real and  $|s| < 1$ , which implies that (32) and (33) are outgoing as  $y \rightarrow -\infty, +\infty$ , respectively, in view of the chosen time dependence (2). Also, for  $|s| > 1$  then  $\gamma(s) > 0$  and so  $\Phi_0$  and  $\Phi_1$  decay exponentially as  $y \rightarrow -\infty, +\infty$ , respectively.

The Fourier transform of the perforated plate equations (15) and (16) gives

$$(s^4 - \mu^4) \frac{d\Phi_0}{dy} + \alpha(\Phi_1 - \Phi_0) = A - isB, \quad (34)$$

$$\frac{d\Phi_1}{dy} - \frac{d\Phi_0}{dy} + \tau\Phi_1 = 0 \quad (35)$$

on  $y=0$ , which allows  $C(s)$  and  $D(s)$  to be determined by direct substitution of (32) and (33). This yields

$$\gamma(s)(s^4 - \mu^4)C(s) + \alpha(D(s) - C(s)) = A - isB, \quad (36)$$

$$-\gamma(s)(D(s) + C(s)) + \tau D(s) = 0, \quad (37)$$

or inverting gives

$$\begin{pmatrix} C(s) \\ D(s) \end{pmatrix} = \frac{1}{K(s)} \begin{pmatrix} \tau - \gamma(s) & -\alpha \\ \gamma(s) & \gamma(s)(s^4 - \mu^4) - \alpha \end{pmatrix} \times \begin{pmatrix} A - isB \\ 0 \end{pmatrix}, \quad (38)$$

where  $K(s)$  is the dispersion function written in (23). The solution to the problem is therefore given by the inverse transforms:

$$\begin{aligned} \phi_0(x, y) &= \frac{1}{2\pi} \int_C \frac{(\tau - \gamma(s))(A - isB)}{K(s)} e^{-isx + \gamma(s)y} ds, \\ &y \leq 0, \end{aligned} \quad (39)$$

$$\begin{aligned} \phi_1(x, y) &= \frac{1}{2\pi} \int_C \frac{\gamma(s)(A - isB)}{K(s)} e^{-isx - \gamma(s)y} ds, \\ &y \geq 0, \end{aligned} \quad (40)$$

in which  $C$  runs from  $-\infty$  to  $+\infty$  on the real line but passing above the branch point at  $-1$  and the simple poles at  $-s_1, -s_2$  [zeros of  $K(s)$ ], and below the branch point at  $+1$  and the poles at  $s_1, s_2$  (see Fig. 3).

The most useful information to extract from the inverse integral representations is the far-field cylindrical wave pattern and the amplitudes of the unattenuated surface waves. The former is obtained by determining the contribution from the integral at the saddle point (when  $r = x^2 + y^2 \rightarrow \infty$ ). To achieve this end we write

$$x = r \cos(\theta), \quad y = r \sin(\theta), \quad -\pi < \theta < \pi, \quad (41)$$

and introduce the transformation

$$s = -\cos(|\theta| + iu), \quad (42)$$

which maps  $s=0$  to the point  $u = i|\theta| - i\pi/2$ . Thus

$$\gamma(s) = -\sin(|\theta| + iu) \quad (43)$$

has the correct value ( $-i$ ) at the image point of  $s=0$ , and the exponents of the integrands in (39) and (40) become

$$-isx - \gamma(s)|y| = ircosh(u). \quad (44)$$

The saddle point occurs at

$$u = 0 \Leftrightarrow s = -\cos|\theta|, \quad (45)$$

which lies between the image of the branch points, and the path of steepest descent from this point is the curve given by

$$\Re\{\cosh(u)\} = 1. \quad (46)$$

The contour of integration  $C$  is deformed into the steepest descent path  $S$ , shown in Fig. 3 for  $|\theta| > \pi/2$ , which, in the  $u$ -plane, is the path starting at  $-i\pi/2 - \infty$  passing through the origin at an angle of  $\pi/4$  radians to the real line, and ending at  $+i\pi/2 + \infty$ . Then, by the usual arguments (see, e.g., p. 34 of the monograph by Noble<sup>13</sup>), as  $r \rightarrow \infty$ ,

$$\phi_0(r, \theta) \sim D(\theta) \frac{e^{ir + i\pi/4}}{\sqrt{2\pi r}}, \quad -\pi \leq \theta \leq 0, \quad (47)$$

$$\phi_1(r, \theta) \sim D(\theta) \frac{e^{ir + i\pi/4}}{\sqrt{2\pi r}}, \quad \pi \geq \theta \geq 0, \quad (48)$$

with

$$\begin{aligned} D(\theta) &= \frac{\sin(\theta)(A + iB \cos(\theta))}{K(-\cos(\theta))} \\ &\times \begin{cases} (-i\tau - \sin(\theta)), & -\pi \leq \theta \leq 0, \\ \sin(\theta), & \pi \geq \theta \geq 0. \end{cases} \end{aligned} \quad (49)$$

Note, when  $\tau=0$ , then the potential field is odd in  $y$ , i.e.,  $\phi_0(r, -\theta) = -\phi_1(r, \theta)$  as expected, and as  $\tau \rightarrow \infty$ ,  $\phi_1(r, \theta) = \mathcal{O}(1/\tau)$  and

$$\begin{aligned} \phi_0(r, \theta) &\sim \frac{-i \sin(\theta)(A + iB \cos(\theta))}{[i(\cos^4(\theta) - \mu^4) \sin(\theta) - \alpha]} \frac{e^{ir + i\pi/4}}{\sqrt{2\pi r}}, \\ &\pi \leq \theta \leq 0. \end{aligned} \quad (50)$$

In the deformation of (39) and (40) from  $C$  to the steepest descent path  $S$  the poles corresponding to the unattenuated surface waves are crossed for values of  $|\theta|$  sufficiently close

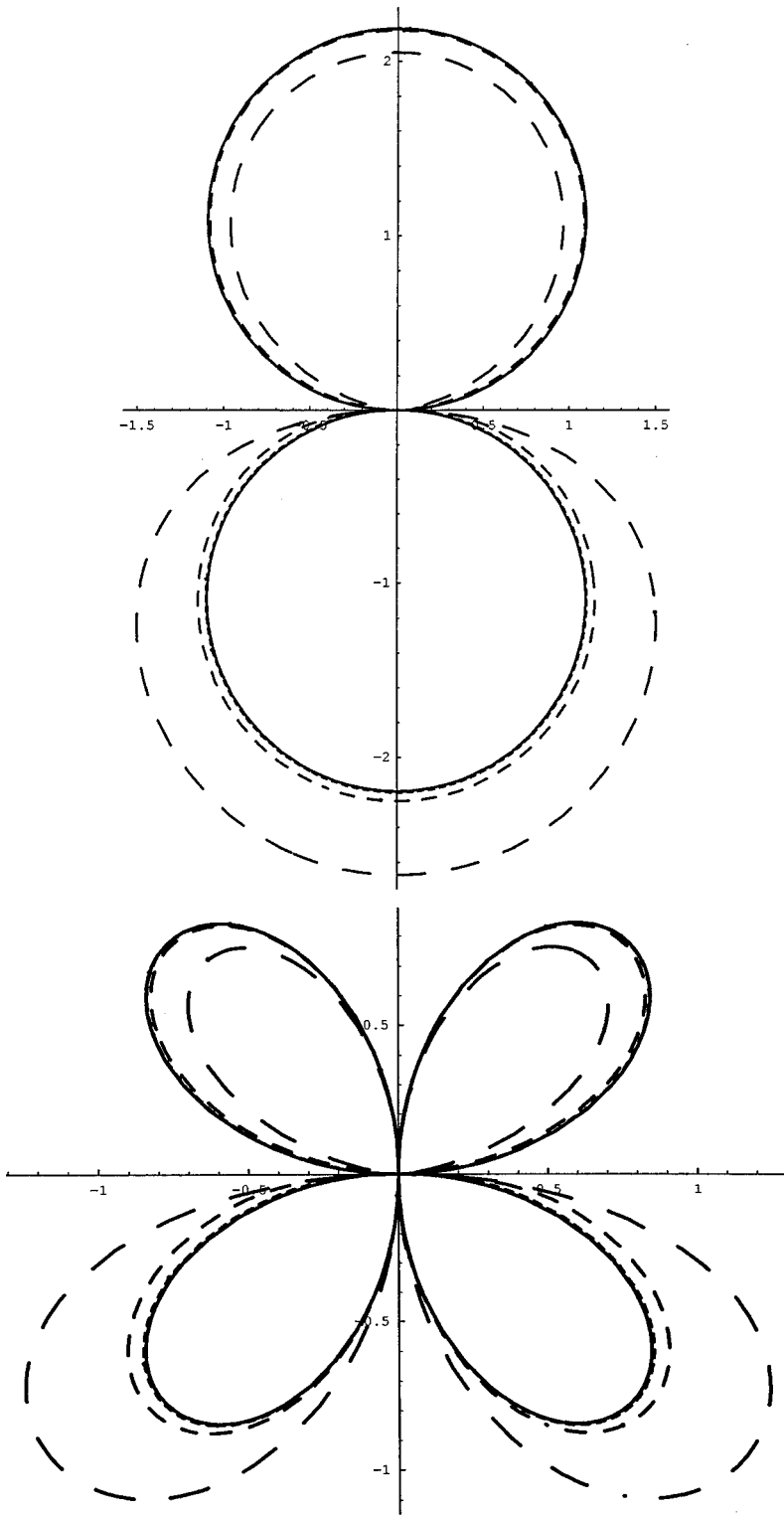


FIG. 4. Polar plots of the scaled amplitude of the far-field diffraction coefficient (49),  $|D(\theta)| \times 10^{15}/f^3$ , for an aluminum plate in water [parameter values as in (53)–(57)]. The curves are for frequencies 500 Hz (solid line), 1 KHz, 2 KHz, and 3 KHz (largest dash size), and forcing coefficients are (a)  $A=1, B=0$ , and (b)  $A=0, B=1$ .

to 0 or  $\pi$ . Picking up the residue contribution at these poles,  $s_1, s_2$ , gives

$$\phi_{s_j}(x, y) = \frac{i(A + is_j \operatorname{sgn}(x)B)}{K'(s_j)} e^{is_j x - \gamma(s_j)|y|} \times \begin{cases} \gamma(s_j), & y > 0, \\ \tau - \gamma(s_j), & y < 0, \end{cases} \quad (51)$$

for  $j=1,2$ , where  $s_j$  are the two positive real roots of  $K(s)$

$=0$ , (23),  $K'(s_j)$  is the derivative of  $K(s)$  evaluated at  $s_j$ , and, as expected, the applied force,  $A$ , and couple,  $B$ , yield even and odd behavior, respectively, in  $x$ . It will prove useful later to define the plate wave coefficient  $P_j$  as

$$P_j = \frac{i\gamma(s_j)(A + is_j \operatorname{sgn}(x)B)}{K'(s_j)}. \quad (52)$$

We can confirm directly from expression (51) that, as

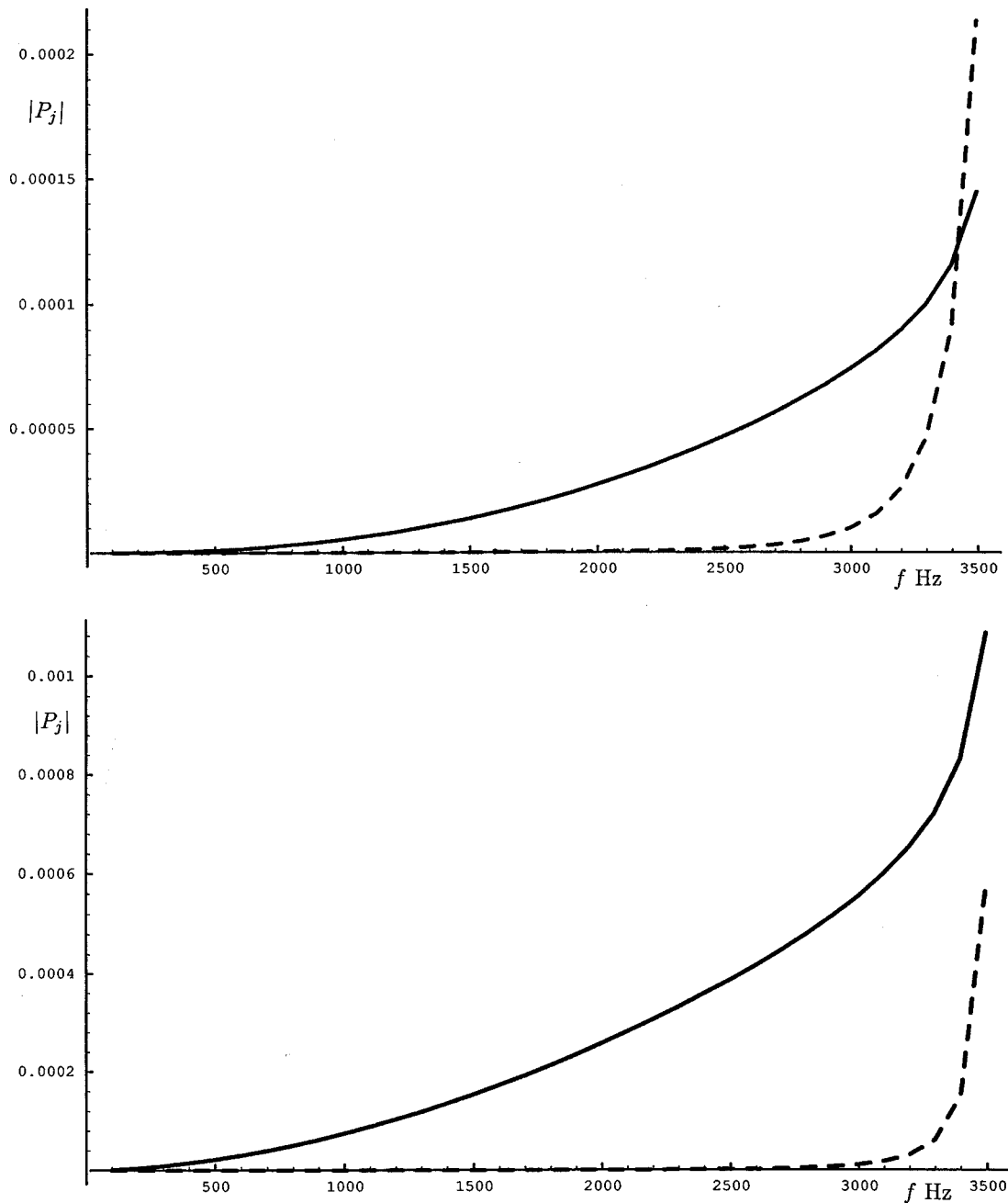


FIG. 5. Plate wave amplitudes (52)  $|P_1|$  (dashed line corresponding to root  $s_1$ ),  $|P_2|$  (solid curve for root  $s_2$ ) versus frequency,  $f$ , for an aluminum plate in water. Forcing coefficients are (a)  $A=1$ ,  $B=0$ , and (b)  $A=0$ ,  $B=1$ .

$\tau \rightarrow 0$ , the lower wave number root  $s_1$  (which tends to unity in this limit) has amplitude which goes to zero. Similarly, as  $\tau \rightarrow \infty$ , the upper root  $s_2 \sim \tau$ ,  $K'(s_2) \sim -\tau^5$ , and so  $\phi_{s_2} \rightarrow 0$ . Thus, in both limits only a single plate wave is recovered. This completes the solution of the boundary value problem, and in the following section we will numerically evaluate the far-field (47), (48) and plate wave terms (51) for a physically realistic range of values.

#### IV. NUMERICAL RESULTS AND CONCLUDING REMARKS

For brevity we will confine attention to a couple of sets of physically reasonable values for the nondimensional parameters. The composite plate will be taken to be made of

aluminum, with a thin acoustically rigid honeycomb spacer (of cuboid shape) in between; the values of the plate/cavity constants are

$$\rho_p = 2.8 \times 10^3 \text{ Kg/m}, \quad h = 2.5 \times 10^{-3} \text{ m}, \quad (53)$$

$$E = 7.4 \times 10^{10} \text{ N/m}^2,$$

$$\nu = 0.33, \quad d = 2 \times 10^{-2} \text{ m}, \quad d_1 = 2 \times 10^{-2} \text{ m} \quad (54)$$

(see Fig. 1), which are typical for normal engineering applications. Note that  $E$  and  $\nu$  are Young's modulus and Poisson's ratio, respectively, for the plate, from which we can deduce the bending stiffness:

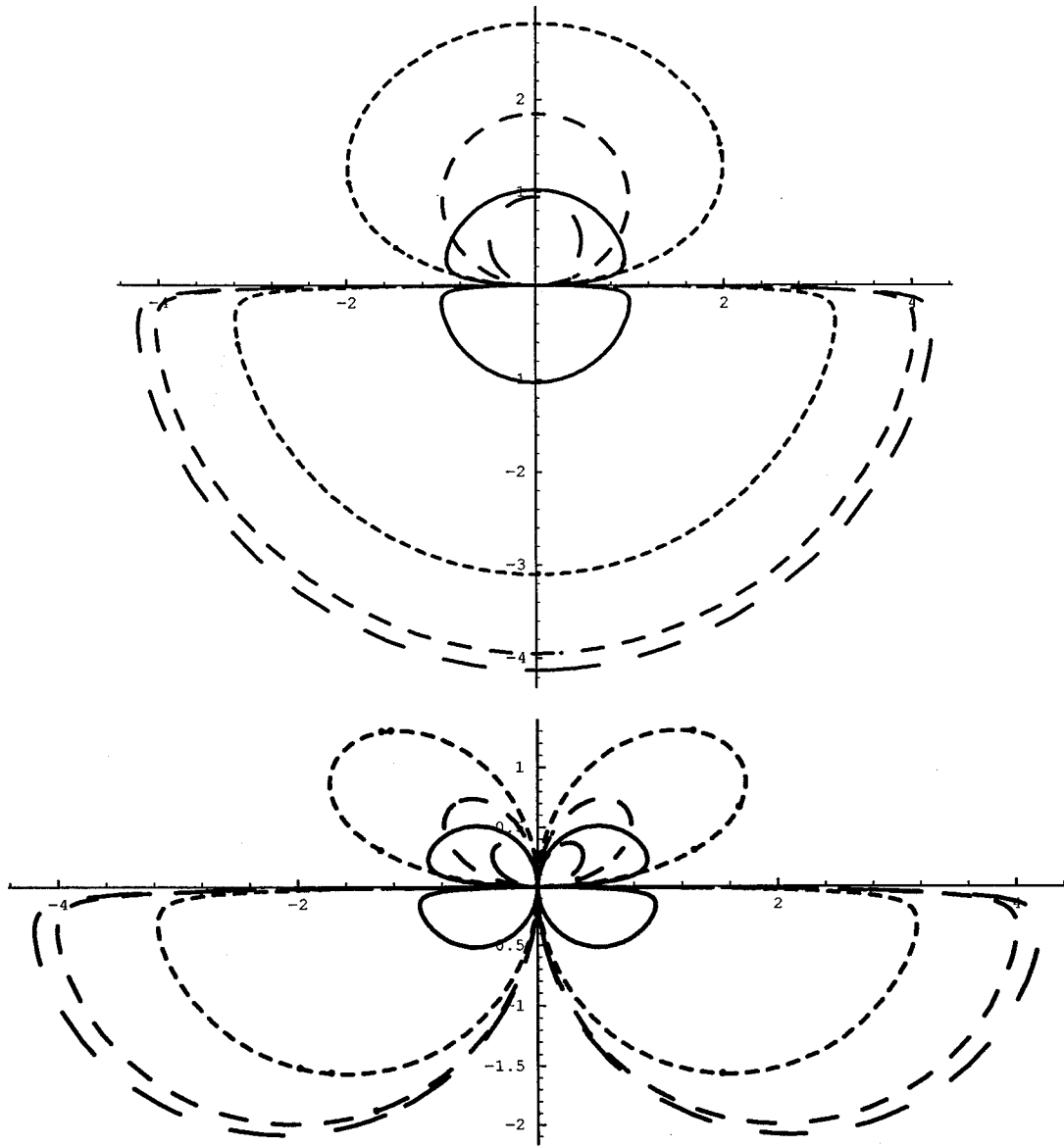


FIG. 6. Polar plots of the scaled amplitude of the far-field diffraction coefficient (49),  $|D(\theta)| \times 10^5 / f$ , for an aluminum plate in air [parameter values as in (53)–(56) and (59)]. The curves are for frequencies 200 (solid line), 600, 764, and 800 (largest dash size), and forcing coefficients are (a)  $A=1$ ,  $B=0$ , and (b)  $A=0$ ,  $B=1$ .

$$B = \frac{Eh^3}{12(1-\nu^2)} = 108.13 \text{ N m.} \quad (55)$$

Also, each cavity is assumed to have a single round perforation of radius

$$a = 1 \times 10^{-3} \text{ m.} \quad (56)$$

The embedding fluid is taken in the first instance as water, with density and sound speed

$$\rho_0 = 10^3 \text{ Kg/m}^3, \quad c_0 = 1439 \text{ m/s,} \quad (57)$$

and so

$$\mu \sim 289/f^{1/2}, \quad \alpha \sim 32717\mu^4/f, \quad \tau \sim \frac{3621f}{\sqrt{10(3621^2 - f^2)}} \quad (58)$$

in which  $f$  is the frequency of the forcing term. Note that we take  $f$  to lie below the Helmholtz frequency, 3621 Hz, in view of the assumptions used in Sec. I.

Figure 4(a) and (b) offer plots of the modulus of the far-field diffraction pattern,  $|D(\theta)|$ , over all  $\theta$  ( $-\pi < \theta < \pi$ ) for the aluminum plate in water and taking the loading, respectively, as a unit force,  $B=0$ ,  $A=1$ , and unit moment  $A=0$ ,  $B=1$ . Each polar plot gives four different values of forcing frequency,  $f=0.5, 1, 2, 3$  KHz (the corresponding  $\tau=0.0445, 0.0945, 0.251, 0.8354$ , respectively) and are marked in increasing dash size. The different curves are scaled using the factor  $10^{15}/f^3$  purely to allow them to be drawn on the same figure. As can be deduced from  $|D(\theta)|$ , the first field is dipole in form, whereas Fig. 4(b) exhibits quadrupole structure. The lobes in  $y > 0$  are roughly the same size as those in  $y < 0$ , i.e., the energy radiated above and below is more-or-less equal, except when the frequency ap-

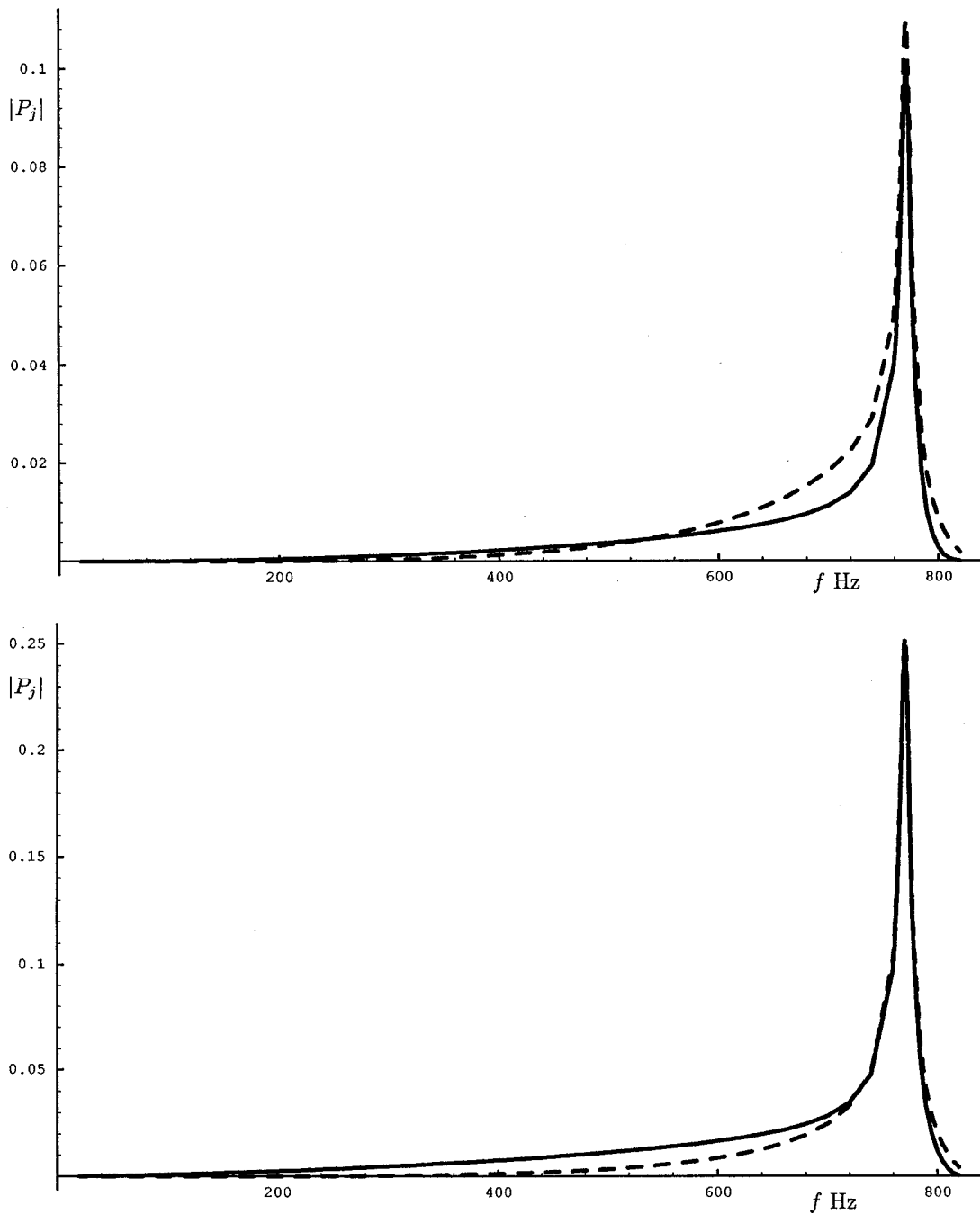


FIG. 7. Plate wave amplitudes (52)  $|P_1|$  (dashed line corresponding to root  $s_1$ ),  $|P_2|$  (solid curve for root  $s_2$ ) versus frequency,  $f$ , for an aluminum plate in air. Forcing coefficients are (a)  $A=1, B=0$ , and (b)  $A=0, B=1$ .

proaches the Helmholtz resonance value. Clearly the only significantly asymmetric curve is the one for  $f=3$  KHz. Related to Fig. 4 are the corresponding plate wave amplitude moduli  $|P_j|$  (52), where  $A=1, B=0$  in Fig. 5(a),  $A=0, B=1$  in Fig. 5(b), and  $j=1,2$  relates to the two plate wavenumbers  $s_1, s_2$  ( $s_1 < s_2$ ) from (23). These figures strongly indicate the fact that the perforations are unimportant over most of the frequency range, that is,  $s_1$  has negligible amplitude in comparison with  $s_2$ , recognizable as the usual thin-plate wave component. So, the effect of the perforations on the radiated and plate wave fields only becomes important at frequencies very close to the Helmholtz resonance value.

The above figures could perhaps indicate the lack of

necessity for including a complicated perforated plate model except very close to the Helmholtz value. However, it is worth investigating the problem in a different physical situation, namely for an aluminum plate in air. Keeping plate/cavity values as those in (53)–(56), the fluid density and sound speed are now taken as

$$\rho_0 = 1.2 \text{ Kg/m}^3, \quad c_0 = 330 \text{ m/s}, \quad (59)$$

respectively. This gives the three dimensionless constants as

$$\mu \sim 66/f^{1/2}, \quad \alpha \sim 9\mu^4/f, \quad \tau \sim \frac{830f}{\sqrt{10(830^2 - f^2)}} \quad (60)$$



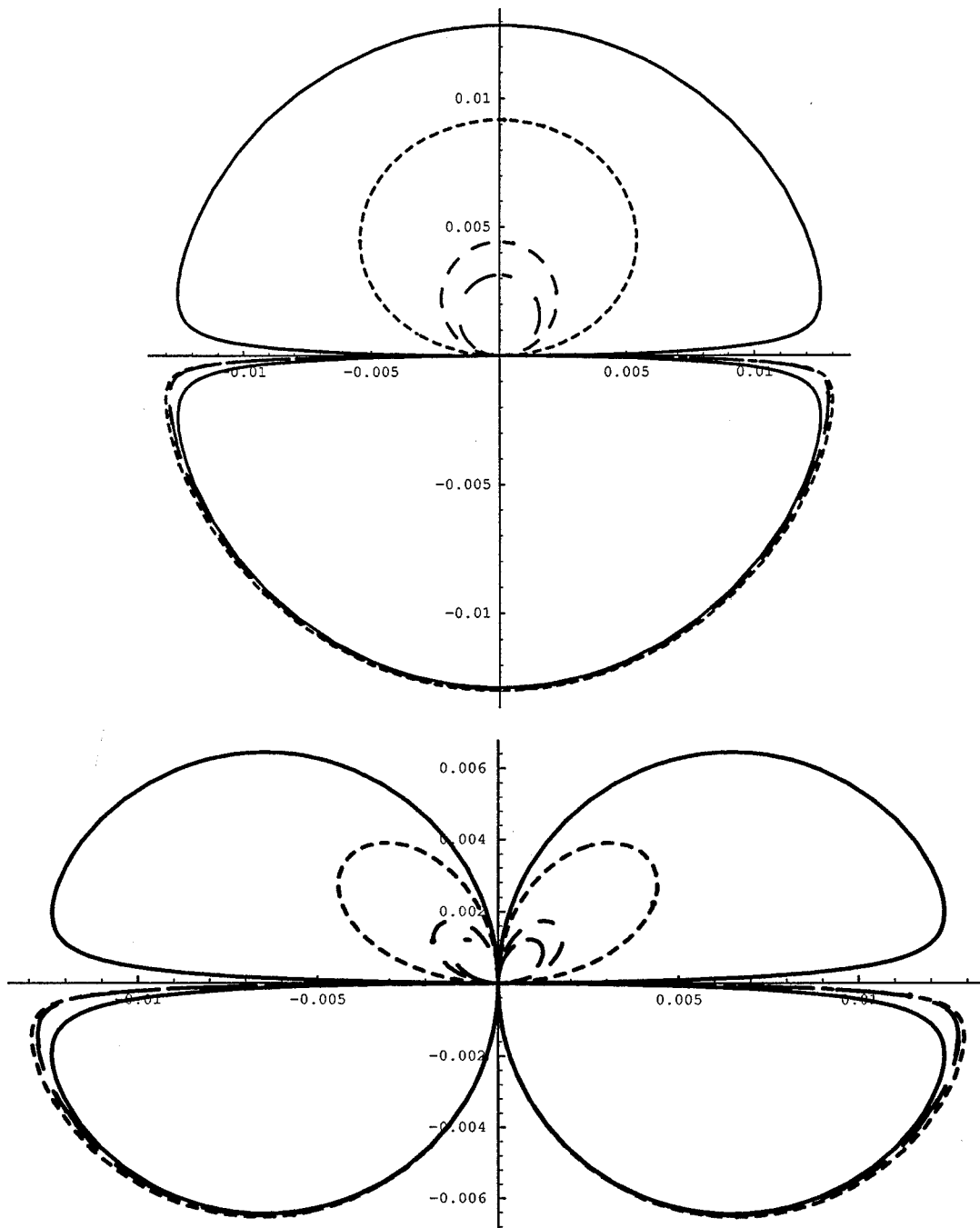


FIG. 8. Polar plots of  $|D(\theta)|$  (49) for an aluminum plate in air, with parameter values as in (53)–(56) and (59) and fixed frequency  $f=500$  Hz but with varying perforation parameter  $\tau$ . The curves are for values of  $\tau=0$  (solid line), 1, 2.75, and 4 (largest dash size), and forcing coefficients are (a)  $A=1$ ,  $B=0$ , and (b)  $A=0$ ,  $B=1$ .

in terms of the frequency  $f$ , and the structure has a Helmholtz resonance value of 830 Hz. The far-field amplitude plots,  $|D(\theta)|$ , for line force ( $B=0$ ) and moment ( $A=0$ ) are given in Fig. 6(a) and (b), for frequency  $f=200$  (continuous curve), 600, 764, and 800 Hz (largest dash size) (corresponding  $\tau$  values are 0.08, 0.48, 1.9, and 4.3, respectively). The amplitudes have again been scaled, this time by the factor  $10^5/f$ , in order to show the curves with approximately similar size in one figure. Notice that, despite the scaling factor decreasing with  $f$ , the curves increase in size in the region  $y < 0$ . However, there is a marked diminution in value of the ratio of energy flow above the plate as compared with radi-

ated energy below as the frequency increases. Not only can the cavity perforations be seen to be important in terms of the energy partitioning for frequencies above say 200 Hz, but in the line force case the remarkably uniform field (in  $\theta$ ) in  $y > 0$  is significantly modified as  $f$ , and hence  $\tau$  increases. This is to be contrasted with the field in  $y < 0$ , and Fig. 4(a) and (b) which do not exhibit any noticeable “beaming” effect. Most marked are the two noticeable lobes in  $y > 0$ , in Fig. 6(b), whose angle of radiation increases with frequency.

Turning now to the plate wave coefficients (52), these are computed for the parameter values (53)–(56) and (59) over the range of frequencies below resonance. As before, in

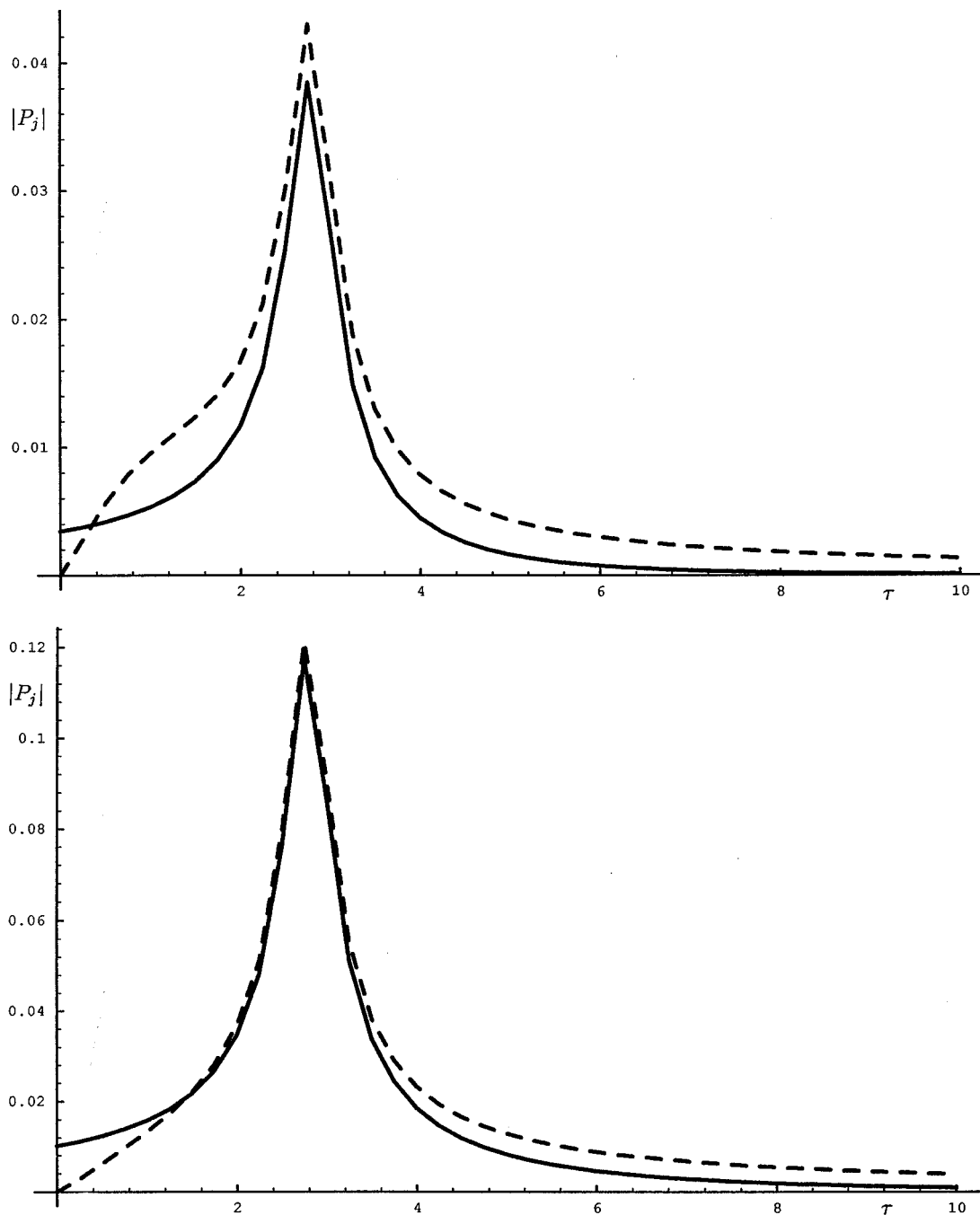


FIG. 9. Plate wave amplitudes (52)  $|P_1|$  (dashed line),  $|P_2|$  (solid curve) versus perforation parameter  $\tau$ , for an aluminum plate in air. Forcing coefficients are (a)  $A=1, B=0$ , and (b)  $A=0, B=1$ .

Fig. 7(a) and (b) the curve corresponding to the small wave number  $s_1$  is dashed, and the higher root,  $s_2$ , which becomes the usual propagating wave number as  $\tau \rightarrow 0$ , is the solid line. Immediately noticeable is the fact that, unlike the aluminum plate/water case [Fig. 5(a) and (b)], the coefficients take comparable values over most of the range of the ordinate. Therefore, energy partitioning has been affected by the modified plate model well away from the Helmholtz resonance value. Indeed, Fig. 7(a) shows that the new wave ( $s_1$ ) coefficient actually exceeds that of  $s_2$  at about 540 Hz (when  $\tau$  is just 0.357). However, the most significant, and unexpected, feature of graphs 7(a) and (b) is the presence of a very large peak at about 764 Hz [hence the reason for taking

this as one of the values in Fig. 6(a) and (b)]. This apparent “resonance” phenomenon does not seem to be easily attributable to any physical confluence of natural frequencies, etc. Nevertheless, at or very close to this value both wave modes attain very large (but finite) amplitudes, greater than 100 times larger than amplitudes at 200 Hz (corresponding to  $10^4$  times the energy in these surface waves)! Note that this is true for both traveling waves  $s_1, s_2$ , and for arbitrary forcing (any values of  $A, B$ ).

To illustrate the “pseudo-resonance” phenomenon a little more clearly, it is interesting to plot the radiated far-field and plate wave coefficients, at a fixed frequency, over a range of  $\tau$  values. This is tantamount to varying the cavities’

resonance parameters, such as by altering the perforation radius or cavity interior lengths. In Fig. 8(a) and (b),  $|D(\theta)|$  is plotted unscaled at a frequency of 500 Hz for four  $\tau$  values: 0 (solid line), 1, 2.75, 4 (largest dashed), and for the two cases of line force and moment loadings, respectively. Clearly, the radiated field into the region  $y < 0$  is remarkably insensitive to  $\tau$  values, whereas in  $y > 0$  there is a **monotonic decrease** in amplitude of the field with increasing  $\tau$ . Figure 9(a) and (b) show typical plate wave coefficients for the same frequency and forcing as in Fig. 8(a) and (b), respectively and with  $\tau$  varying from 0 to 10. As expected from the results of Sec. II [(27) and (28)], the coefficient of the  $s_1$  wave vanishes and  $s_2 \rightarrow s_2^*$  as  $\tau \rightarrow 0$ , and the coefficient of the  $s_2$  wave vanishes and  $s_1 \rightarrow s_1^*$  as  $\tau \rightarrow \infty$ . Clearly indicated is the fact that the amplitude of the  $s_2$  wave is only significantly greater than the  $s_1$  term for very small values of  $\tau$ . It can thus be concluded that for aluminum plates in air, perforated composite plate structures will have a significant effect on the radiation of sound into the enveloping fluid, even for modest  $\tau$  values. The very large amplitude of both plate waves at the “pseudo-resonance,” occurring in Fig. 9(a) and (b) at  $\tau = 2.775$ , is a totally unexpected feature, which may have important physical consequences. For example, there are significant structural dangers in employing finite panels with sharp resonance characteristics. Note, however, that nothing regarding a peak in the surface wave amplitudes at  $\tau \approx 2.75$  is apparently discernible from the cylindrical wave fields in Fig. 8(a) and (b).

In conclusion, this article has investigated the simple problem of sound radiation from an infinite perforated composite plate. It was shown that Leppington’s plate model<sup>5</sup> yields two distinct unattenuated surface waves for all non-zero values of the plate/fluid parameters. In the case of a sandwich construction for typical engineering applications, say composed of thin aluminum plates, the effect of perforations and an interior cavity was noticeable for problems in air but not so marked in underwater applications. In the former case, it is clear that **significant** modification to the energy radiation into the fluid above the plate (perforated side), and into the surface wave terms, occurs as  $\tau$  increases

from zero. An unexpected result is the discovery of a “resonant” frequency or  $\tau$  value at which both plate wave amplitudes increase dramatically. This is likely to be of significant engineering interest, if it proves not to be purely an artifact of the mathematical model, and is to be investigated further in a forthcoming article on the topic by the author.<sup>10</sup>

## ACKNOWLEDGMENTS

The author wishes to acknowledge useful discussions with Professor A. N. Norris, Rutgers University, and Professor F. G. Leppington, Imperial College, University of London, regarding the work contained in this article.

<sup>1</sup>F. Fahy, *Sound and Sources of Sound* (Academic, London, 1987).

<sup>2</sup>F. P. Grooteman and A. de Boer, “A symmetrical element formulation for sound transmission through fuselage walls; theory and application,” paper communicated to Euromech 369 Colloquium on *Fluid-Structure Interaction in Acoustics*, The Netherlands (1997).

<sup>3</sup>J. E. Ffowcs Williams, “The acoustics of turbulence near sound absorbent liners,” *J. Fluid Mech.* **51**, 737–749 (1972).

<sup>4</sup>F. G. Leppington and H. Levine, “Reflexion and transmission at a plane screen with periodically arranged circular or elliptical apertures,” *J. Fluid Mech.* **61**, 109–127 (1973).

<sup>5</sup>F. G. Leppington, “The effective boundary conditions for a perforated elastic sandwich panel in a compressible fluid,” *Proc. R. Soc. London, Ser. A* **427**, 385–399 (1990).

<sup>6</sup>C. M. A. Jones, “Scattering by a semi-infinite sandwich panel perforated on one side,” *Proc. R. Soc. London, Ser. A* **431**, 465–479 (1990).

<sup>7</sup>I. D. Abrahams, “Radiation and scattering of waves on an elastic half-space; a noncommutative matrix Wiener-Hopf problem,” *J. Mech. Phys. Solids* **44**, 2125–2154 (1996).

<sup>8</sup>I. D. Abrahams, “On the solution of Wiener-Hopf problems involving non-commutative matrix kernel decompositions,” *SIAM (Soc. Ind. Appl. Math.) J. Appl. Math.* **57**, 541–567 (1997).

<sup>9</sup>I. D. Abrahams, “On the non-commutative factorization of Wiener-Hopf kernels of Khrapkov type,” *Proc. R. Soc. London, Ser. A* **454**, 1719–1743 (1998).

<sup>10</sup>I. D. Abrahams, “Diffraction of sound by a semi-infinite sandwich panel,” *Proc. R. Soc. London* (submitted).

<sup>11</sup>M. C. Junger and D. Feit, *Sound, Structures and Their Interaction* (MIT, Cambridge, MA, 1986).

<sup>12</sup>P. A. Cannell, “Acoustic edge scattering by a heavily loaded elastic half-plane,” *Proc. R. Soc. London, Ser. A* **350**, 71–89 (1976).

<sup>13</sup>B. Noble, *The Wiener-Hopf Technique* (Chelsea, New York, 1988), 2nd ed.

# Surface waves over bead layers

Jean-François Allard, Michel Henry, and Julian Tizianel

Laboratoire d'Acoustique de l'Université du Maine, UMR CNRS 6613, Avenue Olivier Messiaen 72085  
Le Mans Cedex 9, France

Luc Kelders and Walter Lauriks

Laboratorium voor Akoestiek en Thermische Fysica, Department Natuurkund, Katholieke Universiteit  
Leuven, Celestijnenlaan 200 D, B 3001 Heverlee, Belgium

(Received 23 September 1998; accepted for publication 25 February 1999)

For thin bead layers saturated by air, a pole of the reflection coefficient related to a trapped mode inside the layer and a surface wave in air is predicted, and detected with the Tamura method [J. Acoust. Soc. Am. **88**, 2259 (1990)]. For semi-infinite layers, the Brewster angle is observed with the same method at oblique incidence at 15 kHz. A description of the related Zenneck wave is performed, at high and low frequencies. © 1999 Acoustical Society of America.  
[S0001-4966(99)01506-4]

PACS numbers: 43.20.Fn, 43.60.-c [DEC]

## INTRODUCTION

In a previous paper,<sup>1</sup> sound propagation in air saturating random packings of beads has been described. Due to the small density of air compared to that of the beads, the assumption of a motionless porous frame is valid, leading to a simple model for sound propagation. Air is replaced inside the material by an equivalent fluid having an effective density  $\rho_1$ , and a compressibility  $\beta$  normalized to the adiabatic compressibility of air, that can be predicted with a good approximation by the following general expressions (the time dependence is  $\exp(-i\omega t)$ , the model by Johnson *et al.*<sup>2</sup> is used for  $\rho_1$ , and the model by Lafarge<sup>3</sup> for  $\beta$ )

$$\rho_1 = \alpha_\infty \rho \left[ 1 + \frac{i\eta\phi}{\rho\alpha_\infty k_0 \omega} \left( 1 - \frac{4\alpha_\infty^2 k_0^2 i\omega\rho}{\Lambda^2 \phi^2 \eta} \right)^{(1/2)} \right], \quad (1)$$

$$\beta = \gamma - (\gamma - 1) \left[ 1 + \frac{i\eta\phi}{\rho\alpha_\infty k_0' \text{Pr}\omega} \left( 1 - \frac{4k_0'^2 i\omega\rho \text{Pr}}{\Lambda'^2 \phi^2 \eta} \right)^{(1/2)} \right]^{-1}, \quad (2)$$

where  $\alpha_\infty$  is the tortuosity,  $\rho$  the density,  $\eta$  the viscosity, and Pr the Prandtl number of air,  $\gamma$  the ratio of the specific heats,  $k_0$  and  $k_0'$  are the viscous and the thermal permeability,  $\Lambda$  and  $\Lambda'$  are the viscous and the thermal characteristic lengths,  $\phi$  is the porosity, and  $\omega$  the radian frequency. The wave number  $k_1$  and the characteristic impedance  $Z_1$  in air inside the material are

$$k_1 = \omega \left( \rho_1 \frac{\beta}{\gamma P} \right)^{(1/2)}, \quad (3)$$

$$Z_1 = \left( \rho_1 \frac{\gamma P}{\beta} \right)^{(1/2)}, \quad (4)$$

where  $P$  is the atmospheric static pressure. The beads used in Ref. 1 and in the present study have a diameter  $d$  sharply distributed around  $d = 1.5$  mm. The measured values for the parameters are:<sup>1</sup>  $\phi = 0.4$ ,  $k_0 = 1.5 \times 10^{-9}$  m<sup>2</sup>,  $\alpha_\infty = 1.37$ ,  $\Lambda = 0.9 \times 10^{-4}$  m,  $k_0' = 5 \times 10^{-9}$  m<sup>2</sup>, and  $\Lambda' = 0.32 \times 10^{-4}$  m.

In Ref. 1, the surface impedance of a layer of beads was measured and compared with predictions, to validate the

model and the set of measured parameters. The present work is concerned about poles and zeros of the reflection coefficient of a layer of the same beads saturated by air and for the related surface waves. Many papers have been written concerning long-range sound propagation over ground considered as a porous medium, including the description of poles of the reflection coefficient and the related surface waves. Details of the current state of the art can be found in Ref. 4. The models for sound propagation in these works are generally phenomenological, like the Delany and Bazley model,<sup>5</sup> and not well suited for granular materials of porosity noticeably smaller than 1. A more precise prediction for the location of poles and zeros can be obtained with the model of Ref. 1, which is used in the present work. Experimental localization of a predicted pole for a thin fibrous layer has been performed recently<sup>6</sup> with the near-field acoustical holography method (NAH) developed by Tamura.<sup>7,8</sup> In what follows, similar work is performed for thin layers of beads. Theoretical<sup>9,10</sup> and experimental<sup>11</sup> studies of the surface waves over semi-infinite porous layers saturated by heavy fluids have been performed previously. The full Biot model<sup>12</sup> is used to describe the interactions between the porous frame and the heavy fluid. The simple case of air-saturated beads, where the beads are motionless, brings complementary information. For semi-infinite or very thick layers, zeros and poles, related to the Brewster angle and to a guided wave similar to the electromagnetic Zenneck wave, are predicted. Evidence of zero at high frequencies is obtained with the NAH method.

## I. REFLECTION COEFFICIENT FOR POROUS LAYER-ELECTROMAGNETIC ANALOGIES

In this section, previous results concerning poles and zeros of the reflection coefficient of porous layers are recalled. Analogies between the reflection of acoustic waves by porous layers and of electromagnetic transverse magnetic (TM) waves by dielectrics<sup>13</sup> are shown. Electromagnetic TM waves incident in the  $(x,z)$  plane (see Fig. 1) present a magnetic field perpendicular to the plane and an electric field

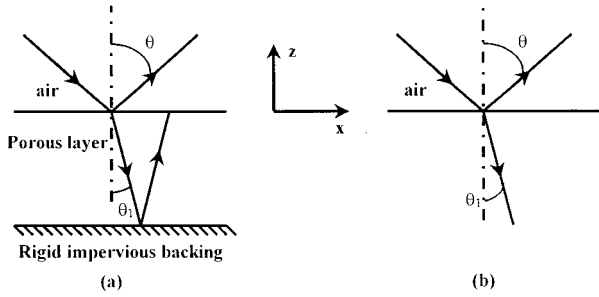


FIG. 1. Representation of reflection and refraction of a plane wave by a porous layer of finite thickness (a) and a semi-infinite layer (b).

parallel to the plane. The electromagnetic analogies provide well-established concepts and methods for the description of acoustic surface waves. A plane wave impinging on a porous layer set on a rigid impervious backing with an angle of incidence  $\theta$  and an angle of refraction  $\theta_1$  is represented in Fig. 1(a). The layer and the backing are replaced by a semi-infinite layer in Fig. 1(b). As in the work by Brekhovskikh and Godin,<sup>14</sup> the  $z$ -axis direction is opposite to the layer. The components of the wave number vector for the reflected wave are given by

$$k_x = k \sin \theta, \quad (5)$$

$$k_z = k \cos \theta, \quad (6)$$

where  $k$  is the wave number in free air, and for the refracted wave the components are given by

$$k_{1x} = k_1 \sin \theta_1, \quad (7)$$

$$k_{1z} = -k_1 \cos \theta_1. \quad (8)$$

The reflection coefficients for the layer of finite thickness  $l$  and for the semi-infinite layer are given by, respectively,

$$R = \frac{-\cos \theta_1 + (i/\phi)(m/n)(\cot k_1 l \cos \theta_1) \cos \theta}{\cos \theta_1 + (i/\phi)(m/n)(\cot k_1 l \cos \theta_1) \cos \theta}, \quad (9)$$

$$R = \frac{-\cos \theta_1 + (m/\phi n) \cos \theta}{\cos \theta_1 + (m/\phi n) \cos \theta}, \quad (10)$$

where  $m = \rho_1/\rho$ ,  $n = k_1/k$ , and the refraction angle  $\theta_1$  is defined by

$$n \sin \theta_1 = \sin \theta. \quad (11)$$

It has been shown that strong analogies exist between the reflection of electromagnetic TM waves (such that the magnetic field has only one component different from zero,  $H_y$ , and the electric field  $\mathbf{E}$  two components  $E_x, E_z$ ) by a dielectric and of acoustic waves by a fluid layer.<sup>15</sup> The equations describing the reflection of TM waves and acoustic waves are very similar if the following correspondence is used:  $H_y \rightarrow$  pressure  $p$ ,  $(E_x, E_z) \rightarrow$  velocity components  $(v_z, -v_x)$ , electric permittivity of vacuum  $\epsilon_0$  and of dielectric  $\epsilon \rightarrow$  density of air  $\rho$ , and of the equivalent fluid  $\rho_1$ , respectively. The magnetic permeability  $\mu_0$  is the same for air and the dielectric, and corresponds to the compressibility, which must be the same in air and in fluid for a strict equivalence, leading to  $m = n^2$ . Setting  $\phi = 1$  to replace the porous layer

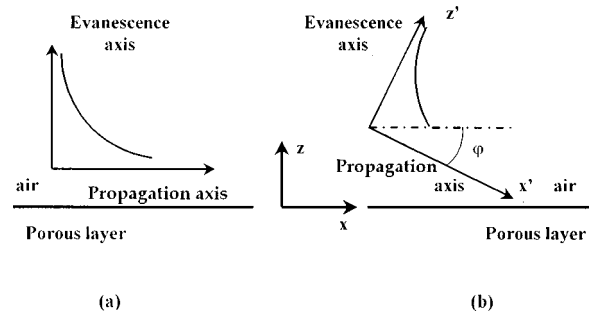


FIG. 2. Representation of surface waves in a lossless medium over a lossless plane structure (a), over a structure with losses (b).

by a fluid layer and replacing  $m$  by  $n^2$  in Eqs. (9) and (10) gives

$$R = \frac{-\cos \theta_1 + i n (\cot k_1 l \cos \theta_1) \cos \theta}{\cos \theta_1 + i n (\cot k_1 l \cos \theta_1) \cos \theta}, \quad (12)$$

$$R = \frac{-\cos \theta_1 + n \cos \theta}{\cos \theta_1 + n \cos \theta}, \quad (13)$$

which are the reflection coefficients of the TM wave with  $n = (\epsilon/\epsilon_0)^{1/2}$  for a grounded dielectric and a semi-infinite dielectric. [See Eq. (50), p. 475, and Eq. (4), p. 455 of Ref. 13.] Poles of the reflection coefficient and related surface waves over bead layers are studied in the next two sections. Precise definitions of the different kinds of guided waves for open radiating structures, like a plane interface between two media, have been given by Tamir and Oliner<sup>16</sup> and references therein in the electromagnetic domain. For Collin,<sup>13</sup> a surface wave is a discrete eigenvalue solution to the wave equation, then related to a pole of the reflection coefficient, intimately bound to the surface of the structure. The field in air is characterized by an exponential decay away from the surface. The ideal plane surface wave for a lossless structure has a wave number vector  $\mathbf{k}$  defined by

$$k_x = (k^2 + \alpha^2)^{1/2}, \quad (14)$$

$$k_z = i\alpha, \quad (15)$$

where  $\alpha$  is a real number. The  $z$  axis is the evanescence axis, and the wave propagates in the  $x$  direction with a velocity  $c_x$

$$c_x = \omega/k_x, \quad (16)$$

smaller than the sound speed in free air  $c$ . The ideal surface wave is represented in Fig. 2(a). When the condition for zero loss is removed, the propagation axis makes an angle  $\phi$  with  $x$ , as shown in Fig. 2(b), the two wave number components are complex, and the structure absorbs the energy transported by the surface wave. It may be noted that if the surface wave propagates in a lossless medium, like air with a good approximation at audible frequencies, the evanescence and the propagation axis are always perpendicular, as shown in Fig. 2(a) and (b).

## II. POLES AND ZEROS OF THE REFLECTION COEFFICIENT FOR A THIN LAYER OF BEADS

An infinite number of poles exist<sup>15</sup> for the reflection coefficient  $R$  given by Eq. (9). The reflection coefficient is an

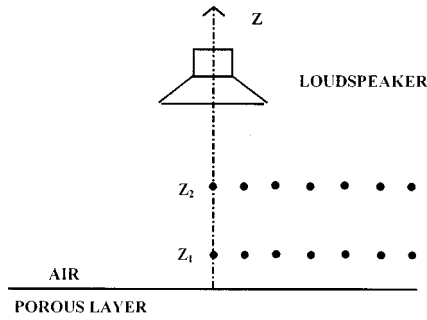


FIG. 3. The experimental setup for the Tamura method. Dots indicate the successive locations of the microphone.

even function of  $\cos \theta_1$ , and  $R \rightarrow 1/R$  when  $\cos \theta \rightarrow -\cos \theta$ . The reflected wave related to a pole can be considered as an incident wave related to a zero. This change of denomination corresponds to a change of sign for  $\cos \theta$ . Each pole is related to a zero for the same  $\sin \theta$ . When  $|k_1| \ll 1$ , only one pole, located close to  $\sin \theta = 1$  in the complex  $\sin \theta$  plane, is related to a surface wave similar to the one represented in Fig. 2(b). At the second-order approximation in  $|k_1|$ , this pole is located at  $\theta_p$  given by<sup>15</sup>

$$\cos \theta_p = k_z/k = \frac{i\phi kl}{m}(n^2 - 1), \quad (17)$$

$$\sin \theta_p = k_x/k = 1 + \frac{\phi^2 k^2 l^2}{2m^2}(n^2 - 1)^2. \quad (18)$$

For a thin grounded dielectric and TM waves, the pole is located at [see Eq. (27), p. 463 of Ref. 13, with  $R_s = X_s = 0$ ]

$$\cos \theta_p = ikl \left( 1 - \frac{1}{n^2} \right), \quad (19)$$

where  $n^2 = \varepsilon/\varepsilon_0$ . Equation (19) can be obtained by setting  $\phi = 1$  and  $m = n^2$  in Eq. (17). The real part of  $\sin \theta_p$  is larger than 1. The wave propagates in the  $x$  direction with a velocity  $c_x = \omega/(k \operatorname{Re}(\sin \theta_p))$  smaller than  $c$ . In order to obtain a precise evaluation of  $\sin \theta_p$  when  $|k_1| \ll 1$  does not hold, Eq. (18) can be used for a very small thickness and the solution modified progressively by an iterative process.<sup>15</sup> If  $\operatorname{Im}(\sin \theta_p) = 0$ ,  $|R|$  has a singularity in the complex  $\sin \theta$  plane on the  $\operatorname{Re}(\sin \theta)$  axis at  $\sin \theta = \sin \theta_p$ . If  $\operatorname{Im}(\sin \theta_p)$  is sufficiently small,  $|R|$  presents a peak on the  $\operatorname{Re}(\sin \theta)$  axis close to  $\sin \theta = \operatorname{Re}(\sin \theta_p)$ . The NAH method<sup>7,8</sup> can be used to measure the reflection coefficient on the  $\operatorname{Re}(\sin \theta)$  axis for  $\sin \theta > 1$ . A source creates an axisymmetric pressure field over a porous layer (see Fig. 3). This field can be written<sup>8</sup>

$$p(r, z) = \int_0^\infty k_r F(k_r) J_0(k_r r) (\exp(-ik_z z) + R(k_r) \exp(ik_z z)) dk_r, \quad (20)$$

where  $r$  is the horizontal distance to the source and  $z$  the vertical distance to the reflecting surface,  $k_r$  and  $k_z$  are the radial and the  $z$  wave number components ( $k_r^2 + k_z^2 = k^2$ ),  $J_0$  is the zero-order Bessel function, and  $F(k_r)$  depends on the nature of the source. For an ideal dipole,  $F(k_r)$  is constant. The reflection coefficient  $R$  for the axisymmetric field is the same as for a plane wave with  $k_x = k_r$  and the same  $k_z$  [see,

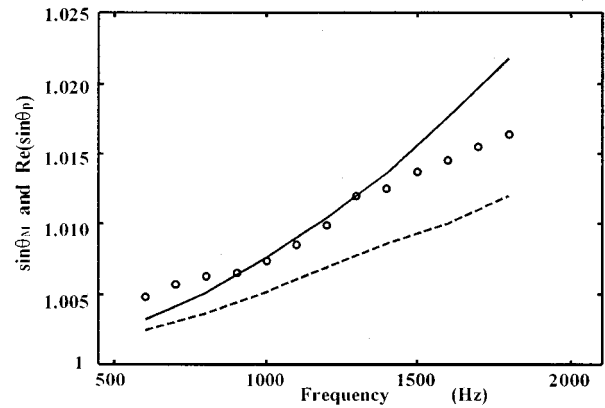


FIG. 4. Location of the maximum of the reflection coefficient  $\sin \theta_M$ , and  $\operatorname{Re}(\sin \theta_p)$ , for a layer of beads of thickness  $l = 2$  cm.  $\sin \theta_M$ : measurement  $\circ \circ \circ$ , prediction —,  $\operatorname{Re}(\sin \theta_p)$ : prediction ---.

for instance, Eq. (1.1.7), p. 3 of Ref. 14]. The amplitudes of the incident and reflected components of the field related to a couple  $k_r, k_z$  are evaluated from a Hankel spatial transform of the field at two heights,  $z_1$  and  $z_2$ . The Hankel transform can be performed with  $k_r > k$  or  $k_r < k$ . When  $k_r > k$ , the related  $\sin \theta$  real and larger than 1 corresponds to  $\theta = \pi/2 - i\psi$ ,  $\psi > 0$ ,  $\sin \theta = ch\psi$ , and  $\cos \theta = ish\psi$ . When  $k_r < k$ ,  $\sin \theta \leq 1$  corresponds to an ordinary incident plane wave with  $0 \leq \theta \leq \pi/2$ . The method has been used previously to detect the peak related to the surface wave for thin layers of fibrous material.<sup>6</sup> In this work, the location of the maximum of the peak on the  $\operatorname{Re}(\sin \theta)$  axis is measured for a layer of beads of thickness  $l = 2$  cm from 600 to 1800 Hz. The experimental setup is represented in Fig. 3. The source is a loudspeaker located 5.5 cm from the layer. The signal is a stretched pulse.<sup>7</sup> The pressure is measured on two axial lines beginning under the center of the loudspeaker and located at  $z_1 = 3$  mm and  $z_2 = 15$  mm from the layer, measurements being performed every 10 mm. The measured locations of the maximum of  $|R|$  on the  $\operatorname{Re}(\sin \theta)$  axis, denoted as  $\sin \theta_M$ , are compared in Fig. 4 to the predicted values obtained from Eq. (9). Measurements of the reflection coefficient have also been performed at 15 kHz for a layer of beads of thickness  $l = 5$  mm. In that case, the source is a tube of diameter equal to 1.4 mm, and the two lines are located 1.5 and 3 mm from the layer, measurements being performed every 1 mm. Measurements and predictions are represented in Fig. 5. The predicted maximum on the  $\operatorname{Re}(\sin \theta)$  axis is located at  $\sin \theta_M = 1.05$ , related to a predicted pole at  $\sin \theta_p = 1.037 + i4.7 \times 10^{-2}$ . The origin of the surface wave related to this pole is a trapped mode inside the porous layer, which experiences total reflection on the backing and the air-porous layer boundary.<sup>15</sup> The surface wave is the evanescent wave in air related to the total reflection at the air-porous layer boundary. A similar phenomenon exists for TM waves over thin grounded dielectrics.<sup>13</sup> It may be noted that, as in the case of the fibrous material previously studied,<sup>6</sup> this surface wave cannot be detected easily by classical methods like time-of-flight measurements. At low frequencies,  $\operatorname{Re}(\sin \theta_p)$  is very close to 1. The velocity of the surface wave in the  $x$  direc-

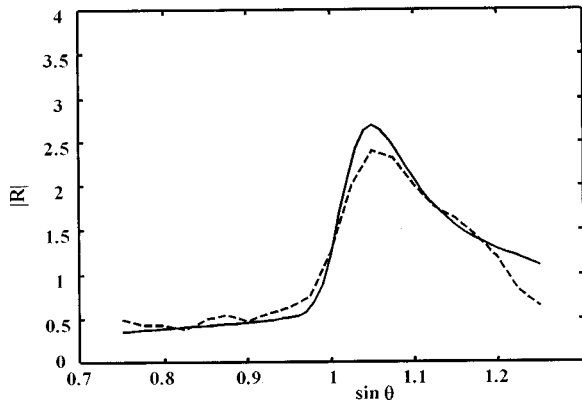


FIG. 5.  $|R|$  as a function of  $\sin \theta$  on the  $\text{Re}(\sin \theta)$  axis at 15 kHz for a layer of beads of thickness  $l=5$  mm: measurement ---, prediction —.

tion, given by  $c/\text{Re}(\sin \theta_p)$ , is very close to  $c$ . When frequency increases,  $\text{Re}(\sin \theta_p)$  becomes noticeably different from 1, but the damping in the  $x$  direction increases rapidly. Detection of the surface wave by time-of-flight measurements over a sufficiently large distance becomes difficult, due to the damping.

### III. POLES AND ZEROS OF THE REFLECTION COEFFICIENT FOR A SEMI-INFINITE LAYER OF BEADS

For a semi-infinite layer, zeros and poles of the reflection coefficient are located in the complex  $\sin \theta$  plane at

$$\sin \theta = \pm \left[ \left( \frac{m^2}{\phi^2} - n^2 \right) / \left( \frac{m^2}{\phi^2} - 1 \right) \right]^{1/2}, \quad (21)$$

$$\cos \theta = \pm \left[ (n^2 - 1) / \left( \frac{m^2}{\phi^2} - 1 \right) \right]^{1/2}, \quad (22)$$

$$\cos \theta_1 = \pm \frac{m}{\phi n} \left[ (n^2 - 1) / \left( \frac{m^2}{\phi^2} - 1 \right) \right]^{1/2}. \quad (23)$$

Poles correspond to  $\cos \theta_1 = -(m/\phi n)\cos \theta$  and zeros to  $\cos \theta_1 = (m/\phi n)\cos \theta$ . At high frequencies, the damping in the porous layer decreases with the viscous skin depth and the parameters  $m$  and  $n$  have a small imaginary part. At 15 kHz, for a semi-infinite layer of the beads previously studied, a zero exists at  $\sin \theta = 0.98 - i2.1 \times 10^{-3}$ ,  $\cos \theta = 0.21 + i9.6 \times 10^{-3}$ .

This zero corresponds to a weakly inhomogeneous wave which experiences total refraction ( $-k \cos \theta$  is the  $z$ -wave number component of the incident wave). The angle of incidence, close to 78 deg with a small imaginary part, is the Brewster angle of total refraction, which has a similar expression for TM waves over a semi-infinite dielectric [Eq. (21) can be compared with Eq. (5a), p. 455 of Ref. 13, when  $\phi=1$  and  $m=n^2$ , as indicated in Sec. I]. The reflection coefficient, measured with the Tamura method, is represented in Fig. 6 and compared with the reflection coefficient predicted with the correct physical choice for  $\cos \theta$  and  $k_1 \cos \theta_1$  [i.e.,  $\text{Im}(\cos \theta) \geq 0$ ,  $\text{Im}(k_1 \cos \theta_1) > 0$ ,  $\text{Re}(k_1 \cos \theta_1) > 0$ ]. The sharp minimum close to  $\sin \theta = 0.98$  indicates the presence of the zero close to the  $\text{Re}(\sin \theta)$  axis. The Brewster angle for

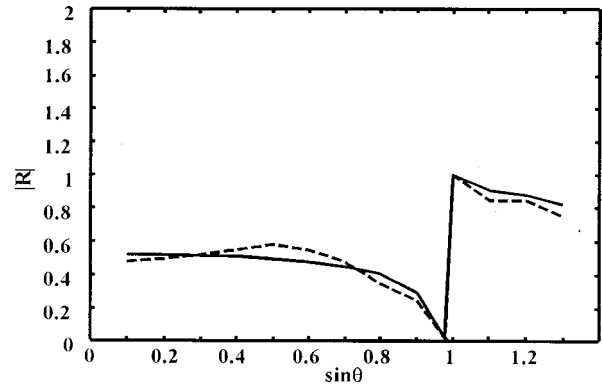


FIG. 6.  $|R|$  as a function of  $\sin \theta$  for a semi-infinite layer of glass beads at 15 kHz: Measurement ---, Prediction —.

porous media has been observed previously at higher frequency by B. Castagnède *et al.*<sup>17</sup> As long ago as 1907, it was shown by Zenneck that a guided electromagnetic wave could travel over the plane interface between two media of the different conductivities and dielectric constants. A short review of the literature concerning the Zenneck wave has been performed by Wait.<sup>18</sup> This wave, like surface waves previously studied, can be considered as a reflected wave for an angle of incidence which corresponds to a pole of the reflection coefficient. Mathematically, substituting  $-\cos \theta$  for  $\cos \theta$  in Eq. (13) leads to  $R \rightarrow 1/R$ , a zero becomes a pole. The observable zero of the reflection coefficient at the Brewster angle defined by a couple  $(\sin \theta, \cos \theta)$  is related to a pole for the couple  $(\sin \theta, -\cos \theta)$ . The change  $(\cos \theta \rightarrow -\cos \theta)$  is obtained by permuting the role of the incident and the reflected wave. Then, the reflected Zenneck wave with  $R$  infinite is identified by Collin<sup>13</sup> to the incident wave at the Brewster angle such as  $R=0$  which experiences total refraction. The unusual geometry of the new incident and the non-modified refracted field has been noticed by Deschamps<sup>19</sup> for similar cases, but the association of the previous refracted wave, and the previous incident (now reflected) wave satisfies all the physical constraints. The acoustic Zenneck wave at 15 kHz has a  $z$ -wave number vector component  $k_z = -k(0.21 + i9.6 \times 10^{-3})$  and its amplitude tends to infinity with  $z$ . This wave is an improper reflected wave, and as indicated by Tamir and Oliner<sup>16</sup> it cannot be viewed as an excited mode in a reflected field.

It may be shown that the situation is different at low frequencies. It has not been possible to measure the reflection coefficient of a semi-infinite layer of beads at low frequencies, due to the large thickness needed to simulate a semi-infinite layer. Predictions are presented at 50 Hz without experimental results. For a semi-infinite layer of the beads previously studied, the predicted location for one of the two zeros is  $\sin \theta = 0.997 + i7.7 \times 10^{-3}$ ,  $\cos \theta = 1.046 \times 10^{-1} - i7.42 \times 10^{-2}$ .

The  $z$ -wave number vector component is now  $k$  ( $-1.046 \times 10^{-1} + i7.42 \times 10^{-2}$ ) for the Zenneck wave, which is loosely bound to the surface. The Zenneck wave is not an improper reflected wave, and could be present, at low frequencies, in acoustic fields reflected by thick layers of

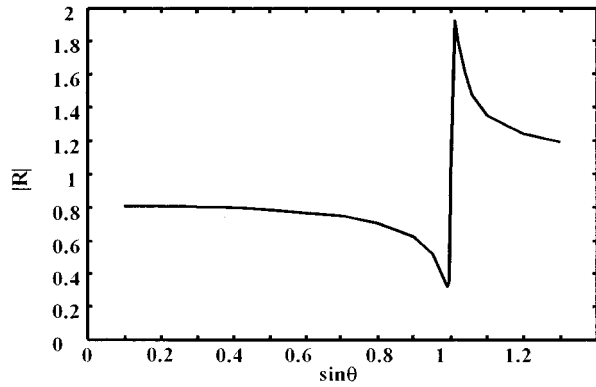


FIG. 7. Predicted  $|R|$  as a function of  $\sin \theta$  for a semi-infinite layer of glass beads at 50 Hz.

granular materials, at least at sufficiently small distances from the source (see p. 24 of Ref. 14).

The phase velocity in the  $z$  direction,  $c/\text{Re}(\sin \theta)$ , is larger than the sound speed. Simple calculations show that this wave is similar to the one represented in Fig. 2(b) with an angle  $\varphi=6\text{deg}$ . The  $x'$  and  $z'$  components of the wave number vector are  $k'_x=1.0028k$ ,  $k'_z=i.7.46\times 10^{-2}k$ . The component  $k'_x$  is very close to  $k$  and  $\text{Re}(k_x)=k'_x \cos \varphi$  is smaller than  $k$ . The case of electromagnetic Zenneck waves is similar. The predicted modulus of the reflection coefficient at 50 Hz is represented in Fig. 7. There is a minimum at  $\sin \theta < 1$  which is larger than at 15 kHz. A maximum now appears for  $\sin \theta > 1$ , which is not related to a pole at  $\text{Re}(\sin \theta) > 1$ . The peak can be predicted by rewriting Eq. (10) with  $\cos \theta_1=1$  [ $n=3.31+i2.75$ , and Eq. (11) indicates that  $\cos^2 \theta_1$  is close to 1 in a large range of variation of  $\sin \theta$ ]

$$R = \frac{-\phi n/m + \cos \theta}{\phi n/m + \cos \theta}, \quad (24)$$

where  $\phi n/m=0.103-i0.077$ . The minimum is reached at  $\sin \theta < 1$  for  $\cos \theta$  close to  $\text{Re}(\phi n/m)$ , and the maximum at  $\sin \theta > 1$  for  $\cos \theta$  close to  $-\text{Im}(\phi n/m)$ .

#### IV. CONCLUSION

The model for sound propagation presented in a previous paper is used to predict poles and zeros of the reflection coefficient of thin and semi-infinite bead layers. These predictions are in good agreement with measurements of the reflection coefficient performed at oblique incidence and for inhomogeneous waves with the near-field acoustical holog-

raphy method. Moreover, this model predicts a specific behavior of the reflection coefficient at low frequencies around the grazing incidence, related to the Brewster angle and the possible presence of a Zenneck reflected wave.

- <sup>1</sup>J. F. Allard, M. Henry, J. Tizianel, L. Kelders, and W. Lauriks, "Sound Propagation in Air Saturated Random Packings of Beads," *J. Acoust. Soc. Am.* **104**, 2525–2528 (1998).
- <sup>2</sup>D. L. Johnson, J. Koplik, and R. Dashen, "Theory of Dynamic Permeability and Tortuosity in Air-Saturated Porous Media," *J. Fluid Mech.* **176**, 379–402 (1987).
- <sup>3</sup>D. Lafarge, P. Lemarinier, J. F. Allard, and V. Tarnow, "Dynamic compressibility of Air in Porous Structures at Audible Frequencies," *J. Acoust. Soc. Am.* **102**, 1995–2006 (1997).
- <sup>4</sup>K. Attenborough, "Review of Ground Effects on Outdoor Sound Propagation from Continuous Broadband Sources," *Appl. Acoust.* **24**, 289–319 (1988).
- <sup>5</sup>M. A. Delany and E. N. Bazley, "Acoustic Properties of Fibrous Absorbent Materials," *Appl. Acoust.* **3**, 105–116 (1970).
- <sup>6</sup>J. Tizianel, J. F. Allard, W. Lauriks, and L. Kelders, "Experimental Localization of a Pole of the Reflection Coefficient of Porous Layers," *J. Sound Vib.* **204**, 600–604 (1997).
- <sup>7</sup>M. Tamura, "Spatial Fourier Transform Method of Measuring Reflection Coefficients at Oblique Incidence," *J. Acoust. Soc. Am.* **88**, 2259–2264 (1990).
- <sup>8</sup>B. Brouard, D. Lafarge, J. F. Allard, and M. Tamura, "Measurement and Prediction of the Reflection Coefficient of Porous Layers at Oblique Incidence and for Inhomogeneous Waves," *J. Acoust. Soc. Am.* **99**, 100–107 (1996).
- <sup>9</sup>S. Feng and D. L. Johnson, "High-Frequency Acoustic Properties of a Fluid/Porous Solid Interface. I. New Surface Mode," *J. Acoust. Soc. Am.* **74**, 906–914 (1983).
- <sup>10</sup>S. Feng and D. L. Johnson, "High-Frequency Acoustic Properties of a Fluid/Porous Solid Interface. II. The 2D Reflection Green's Function," *J. Acoust. Soc. Am.* **74**, 915–924 (1983).
- <sup>11</sup>P. B. Nagy, "Observation of a New Surface Mode on a Fluid-Saturated Permeable Solid," *Appl. Phys. Lett.* **60**, 2735–2737 (1992).
- <sup>12</sup>M. A. Biot, "The Theory of Propagation of Elastic Waves in a Fluid-Saturated Porous Solid," I. Low Frequency Range. II. Higher Frequency Range," *J. Acoust. Soc. Am.* **28**, 168–191 (1956).
- <sup>13</sup>R. E. Collin, *Field Theory of Guided Waves* (McGraw-Hill, New York, 1960).
- <sup>14</sup>L. M. Brekhovskikh and O. A. Godin, *Acoustics of Layered Media II. Point Source and Bounded Beams*, Springer Series on Wave Phenomena (Springer, New York, 1992).
- <sup>15</sup>W. Lauriks, L. Kelders, and J. F. Allard, "Surface Waves and Leaky Waves Above a Porous Layer," *Wave Motion* **28**, 59–67 (1998).
- <sup>16</sup>T. Tamir and A. A. Oliner, "The Spectrum of Electromagnetic Waves Guided by a Plasma Layer," *Proc. IEEE* **51**, 317–332 (1963).
- <sup>17</sup>B. Castagnede, A. Aknine, M. Melon, and C. Depollier, "Ultrasonic Characterization of the Anisotropic Behavior of Air-Saturated Porous Materials," *Ultrasonics* **36**, 323–341 (1998).
- <sup>18</sup>J. R. Wait, "Excitation of Surface Waves on Conducting, Stratified, Dielectric-Clad, and Corrugated Surfaces," *J. Res. Natl. Bur. Stand.* **59**, 365–377 (1957).
- <sup>19</sup>M. Deschamps, "Reflection and Refraction of the Inhomogeneous Plane Wave," in *Acoustic Interaction with Submerged Elastic Structures*, edited by A. Guran, J. Ripoche, and F. Ziegler (World Scientific, Singapore, 1996), Part I, Series B, Vol. 5.



# Experimental verification of the opposite effect of fluid loading on the velocity of dilatational waves in thin plates and rods

Waled Hassan

*Air Force Research Laboratory, Materials and Manufacturing Directorate, Metals, Ceramics, and NDE Division, WPAFB, Ohio 45433-7817*

Peter B. Nagy

*Department of Aerospace Engineering and Engineering Mechanics, University of Cincinnati, Cincinnati, Ohio 45221-0070*

(Received 14 September 1998; accepted for publication 12 January 1999)

In a recent paper [J. Acoust. Soc. Am. **102**, 3478 (1997)], it was demonstrated by analytical means that radiation loading increases the velocity of dilatational waves in immersed thin plates, but decreases it in thin rods. The main goal of this paper is to verify experimentally the predicted opposite effect, which is particularly interesting because in almost every respect, the lowest-order dilatational modes of wave propagation in thin plates and rods are very similar. Experimental verification of the predicted small radiation-induced velocity change is rather difficult, partly because of the accompanying strong attenuation effect caused by radiation losses, and partly because of the presence of an additional velocity change caused by viscous drag even in low-viscosity fluids like water. In spite of these inherent difficulties, the presented experimental results provide unequivocal verification of the earlier theoretical predictions. © 1999 Acoustical Society of America. [S0001-4966(99)03005-2]

PACS numbers: 43.20.Jr, 43.20.Tb [DEC]

## INTRODUCTION

The problem of wave propagation in fluid-loaded plates was extensively studied by numerous investigators. Schoch,<sup>1</sup> Osborne and Hart,<sup>2</sup> Merkulov,<sup>3</sup> Viktorov,<sup>4</sup> Pitts *et al.*,<sup>5</sup> and Selezov *et al.*<sup>6</sup> studied the effect of water loading on plates under a variety of conditions. Their results indicated that most modes are strongly attenuated by leaking into the fluid, but their phase velocity is essentially unaffected. More recent studies by Nayfeh and Chimenti<sup>7</sup> and Rokhlin *et al.*<sup>8</sup> showed that in some special cases of significant importance in non-destructive characterization of composite laminates, when the density ratio between the solid and the fluid is relatively low, the change in the phase velocity of the fundamental (lowest-order) symmetric mode is not negligible and, in extreme cases, even the topology of the mode structure might change. Most of the theoretical works on wave propagation in loaded rods considered the more general case of clad rods consisting of an isotropic core and an arbitrary number of isotropic coatings. A comprehensive review of these studies was published by Thurtson.<sup>9</sup>

Longitudinal guided-wave propagation in thin plates and rods is of great practical importance. Such guided modes are commonly used to evaluate the material properties of thin foils and plates, metal wires, optical fibers, and reinforcement filaments used in epoxy, metal, and ceramic matrix composites. The feasibility of using such longitudinal guided waves to evaluate the interface properties between the fiber and the surrounding solid matrix was investigated by both theoretical and experimental means.<sup>10,11</sup> Another interesting application is when an imbedded fiber is used to monitor the properties of the surrounding material during poly-

merization.<sup>12</sup> In this case, the fiber properties are known and the interface conditions are assumed to be perfect. Any change in the velocity or the attenuation of the guided mode in the fiber can be attributed to the surrounding material that changes from a viscous fluid to an elastic solid during the curing process to be monitored. Guided modes producing mainly tangential displacement can be readily used to measure fluid viscosity. From this point of view, torsional modes are the best,<sup>13,14</sup> but they are more difficult to generate and detect than extensional modes.<sup>15</sup> Nagy and Kent recently utilized the fundamental symmetric mode in thin wires and fibers to assess their Poisson's ratio through measuring the leaky attenuation of these modes in such structures.<sup>16</sup>

At low frequencies, only the fundamental symmetric longitudinal and antisymmetric flexural modes can propagate in thin plates. In rods, a dispersion-free torsional mode can also propagate in addition to the longitudinal and flexural modes. The fundamental longitudinal modes in thin plates and rods are physically very similar. Figure 1 shows the schematic diagram of a thin plate and a rod along with the common deformation pattern caused by the fundamental dilatational mode. At sufficiently low frequencies, the displacement along the propagation ( $x$ ) direction dominates the vibration. In a thin plate, the Poisson effect causes a weaker transverse vibration perpendicular to the surface of the plate, which is less than the longitudinal displacement by a factor proportional to  $\omega d\nu/(1-\nu)$ , where  $\omega$  denotes the angular frequency,  $d$  is the thickness of the plate, and  $\nu$  is the Poisson ratio. The fundamental axisymmetric mode propagating in a thin rod exhibits a similar behavior. The displacement along the axis of the rod dominates the vibration, while the

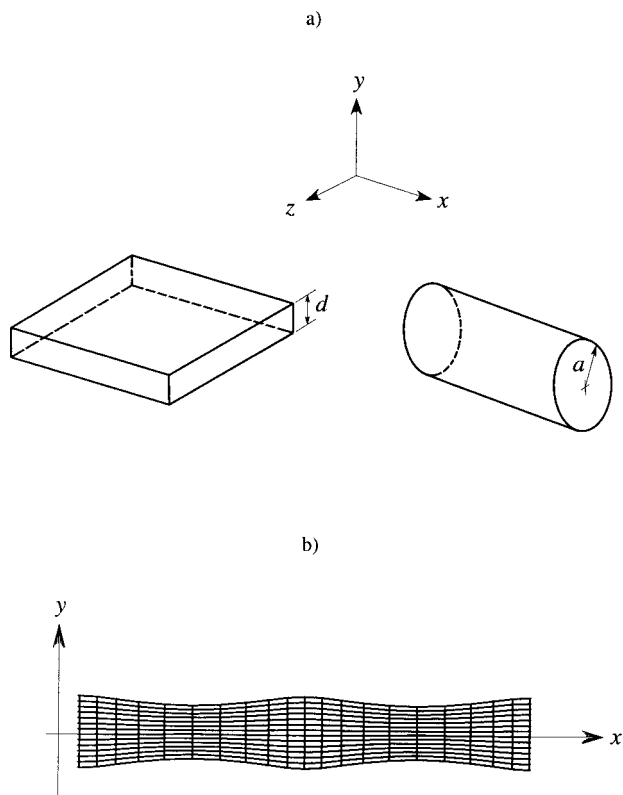


FIG. 1. The schematic diagram of a plate and a rod with the associated coordinate system (a) and the deformation pattern caused by the fundamental dilatational mode (b).

radial displacement caused by Poisson's effect is less than the longitudinal displacement by a factor proportional to  $\omega a \nu$ , where  $a$  is the radius of the rod. The only minor difference between these two modes is that in the case of a plate, the Poisson effect is restricted to only one dimension normal to the plate; therefore, the dilatational wave velocity is  $\sqrt{1/(1-\nu^2)} \approx 1.05$  times higher in a plate than in a rod for a typical value of  $\nu=0.3$ . As the frequency increases, in both cases the phase velocity first drops proportionally to the square of frequency and then asymptotically approaches the Rayleigh velocity. In spite of these common features of the dilatational vibrations in thin plates and rods, the effect of fluid loading on the phase velocity of these two modes is quite different.<sup>17</sup> Although the magnitude of the relative change in the phase velocity due to radiation loading is very small and practically negligible compared to the much stronger attenuation effect in both cases, the very fact that the signs of these velocity changes are opposite deserves some attention. This fundamental difference in the effect of radiation loading on the velocity of the dilatational mode in thin plates and rods was thoroughly investigated by analytical means in Ref. 17. The phenomenon was first demonstrated by both numerical and asymptotic inspection of the well-known exact dispersion equations. Then, simplified approximate models were introduced to facilitate a better understanding of the physical differences between the two cases that lead to such an opposite behavior. It was found that the somewhat unexpected opposite effect of fluid loading on immersed rods and plates is caused by the different nature of their radiation loading. In the case of the plate, the radiation

impedance exerted by the fluid is purely real, whereas it is dominantly imaginary in the case of thin rods. Our main goal in this paper is to verify experimentally these theoretical predictions by measuring the relative change in the velocity of the dilatational waves in thin plates and rods immersed in water.

## I. THEORETICAL CONSIDERATIONS

Generally, guided-wave propagation in plates and rods immersed in ordinary low-viscosity fluids like water is mainly affected by radiation loading due to the normal component of the surface vibration. The fundamental dilatational modes in thin plates and rods are somewhat different in that, due to the dominantly longitudinal vibration in the solid, the tangential component of the surface vibration can cause significant viscous drag. In this case, the total fluid-loading effect can be separated into comparable radiation and viscous parts. First, the normal component of the surface vibration causes radiation losses via energy leakage into the fluid, and slightly increases or decreases the velocity depending on whether we consider a plate or a rod. Second, the tangential component of the surface vibration causes additional losses via viscous dissipation in the fluid and slightly decreases the velocity in both plates and rods. Since the change in the velocity of this mode in thin plates and rods due to leakage into the fluid is very small, on the order of a fraction of a percent,<sup>17</sup> careful considerations must be taken of all extraneous effects to successfully measure it. Detailed analysis of the various parameters affecting this phenomenon is needed to establish the physical limits imposed on such measurements and to design the optimal experiment. It is also necessary to pay proper attention to other usually negligible effects such as inherent temperature variations caused by immersion in the fluid bath, remnant wetness, and evaporation cooling when the specimen is removed from the bath, etc. In the following discussion of the various factors affecting our ability to successfully measure the sought radiation-induced velocity change in immersed plates and rods, we separate the relevant parameters into two categories. The primary factors include the frequency, the geometrical dimensions, and the acoustical properties of the solid and the fluid used in the experiment. The secondary factors are those due to viscosity and temperature variations. In the following sections we proceed to examine these different factors, their contributions, and their importance. The results of the analysis are then implemented in the design of the experiments that are aimed at measuring the effect of radiation-induced fluid loading on the velocity of the dilatational mode in thin plates and rods.

### A. Radiation loading

In our experiment, the change in the velocity of the dilatational wave in thin plates and rods is determined at a given frequency by measuring the phase change in the received ultrasonic signal between the free case and the fluid-loaded case. To examine the effect of frequency, the dimension of the specimen, and the acoustical properties of the solid and the fluid on the accuracy of the measurement, and to establish their optimal values, we need to maximize the

radiation-induced change in the phase angle of the received signal against the attenuation of the signal. Too much attenuation will cause a serious degradation of the received ultrasonic signal and would result in significant errors in the phase measurement. On the other hand, too little attenuation is inherently accompanied by only a slight change in the phase of the received signal making it virtually impossible to measure the accurate phase variation. Towards this end, we define a “quality” factor  $Q$  as the ratio between the change in the phase angle  $\Delta\Phi$  and the total loss of the signal  $\Delta L$  over the fluid-loaded portion of the structure. The change in the phase angle can be approximated as

$$\Delta\Phi \approx \frac{\Delta c}{c_o^2} \omega \ell, \quad (1)$$

where  $\Delta c$  is the change in the velocity of the dilatational mode due to fluid loading,  $c_o$  is the low-frequency asymptotic limit of the velocity of the dilatational mode in the free structure,  $\omega$  is the angular frequency, and  $\ell$  is the length of the fluid-loaded portion of the structure. For the plate,  $c_o = \sqrt{E/[\rho_s(1-\nu^2)]} = c_s \sqrt{2/(1-\nu)}$ , and for the rod it is given as  $c_o = \sqrt{E/\rho_s} = c_s \sqrt{2(1+\nu)}$ , where  $E$  denotes Young’s modulus,  $\nu$  is the Poisson ratio,  $c_s$  is the shear velocity in the solid, and  $\rho_s$  is its density. The total attenuation of the mode due to energy leakage into the fluid is calculated as  $\Delta L = \alpha \ell$ , where  $\alpha$  is the attenuation coefficient of the leaky mode. Combining the above equations, we can express the quality factor as

$$Q = \frac{\omega \Delta c}{\alpha c_o^2}. \quad (2)$$

The low-frequency, weak-loading asymptotic expressions for the relative change in the velocity of the dilatational wave in thin plates and rods can be found in Ref. 17 and are repeated here for convenience. For the case of the plate, the relative velocity change due to radiation loading is given as

$$\left. \frac{\Delta c}{c_o} \right|_{\text{plate}} \approx \frac{1}{32} \left[ \frac{\rho d \omega \nu c_f^2}{c_s^2} \right]^2, \quad (3)$$

whereas for the rod it can be expressed as

$$\left. \frac{\Delta c}{c_o} \right|_{\text{rod}} \approx \frac{\rho a^2 \omega^2 \nu^2}{2 c_s^2 (1+\nu)} \ln \left( \frac{a \omega}{c_o} \sqrt{\frac{c_o^2}{c_f^2} - 1} \right). \quad (4)$$

Here,  $\rho = \rho_f/\rho_s$ , is the density ratio between the fluid and the solid,  $d$  is the thickness of the plate,  $a$  is the radius of the rod, and  $c_f$  is the sound velocity in the fluid. The opposite sign of the radiation-induced velocity change in plates and rods is quite obvious from Eqs. (3) and (4). The quantity given for a plate in Eq. (3) is always positive, while the corresponding quantity given for a rod in Eq. (4) is negative for low values of the normalized frequency  $a\omega/c_o$ . The low-frequency asymptotic expressions for the leaky attenuation coefficient of this mode are listed in Ref. 16 and are also repeated here for convenience. For the plate, we have

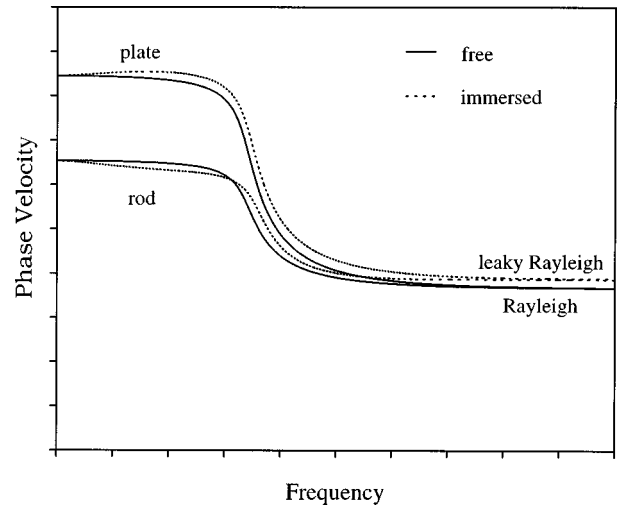


FIG. 2. Schematic diagram showing the fluid-loading induced changes in the topology of the dispersion curves for the fundamental dilatational modes in thin plates and rods (the changes in the phase velocity have been exaggerated for clarity).

$$\alpha_{\text{plate}} \approx \frac{\rho d \omega^2 \nu^2}{4 c_o^2 (1-\nu)^2 \sqrt{\frac{c_o^2}{c_f^2} - 1}}, \quad (5)$$

while for the rod we have

$$\alpha_{\text{rod}} \approx \frac{\pi \rho a^2 \omega^3 \nu^2}{2 c_o^3}. \quad (6)$$

Combining Eqs. (2) through (6) we arrive at the low-frequency asymptotic expressions for the above defined quality factor in the case of the plate and the rod. For the plate, we obtain

$$Q_{\text{plate}} \approx \frac{\rho d \omega c_f^2}{2 c_o^3} \sqrt{\frac{c_o^2}{c_f^2} - 1}. \quad (7)$$

Similarly, for the rod we can write the quality factor as follows:

$$Q_{\text{rod}} \approx \frac{2}{\pi} \ln \left( \frac{a \omega}{c_o} \sqrt{\frac{c_o^2}{c_f^2} - 1} \right). \quad (8)$$

To design a successful experiment, we need to maximize the magnitude of the quality factor (it is positive for the plate and negative for the rod). Several parameters can be changed to achieve this goal, including the characteristic dimension and velocity of the specimen, the frequency, and the density and velocity ratios between the solid and the fluid. Among all these parameters, frequency seems to be the one that can be most easily varied. Sweeping the frequency over a carefully chosen and sufficiently wide frequency range will be shown to achieve the desired goal in the case of the plate but has serious limitations in the case of the rod. To help illustrate the difference between the plate and the rod we refer to Fig. 2, which shows a schematic diagram of the dispersion curves for the fundamental dilatational modes in free and fluid-loaded plates and rods. The solid lines represent the phase velocity of the modes in the free case, and the dotted

lines give the velocity in the fluid-loaded case. In this schematic figure, the fluid loading-induced changes in the velocity of this mode have been greatly exaggerated for clarity. It is very well known that at high frequencies the velocity of the fundamental dilatational modes approach the velocity of the Rayleigh wave in both the free plate and the rod. Similarly, in the fluid-loaded case, those two modes approach the leaky Rayleigh mode which has a slightly higher velocity than the true Rayleigh mode (approximately 0.5% higher for water-loaded aluminum). As shown in Fig. 2, in the case of the plate, the velocity of the fundamental dilatational mode increases due to fluid loading throughout the entire frequency range, whereas in the case of the rod, the velocity first drops at low frequencies due to fluid loading and then switches over to be higher than the velocity in the free case. For typical Poisson ratios, the switch between negative and positive radiation-loading effects occurs at around  $a\omega/c_s \approx 1.3$ . Our measurements in the case of the rod should be carried out at frequencies well below this switch-over frequency. As a matter of fact, the negative radiation-loading effect peaks at around  $a\omega/c_s \approx 0.6$ . More importantly, the magnitude of the above-defined quality factor, i.e., the ratio between the radiation-induced velocity and attenuation effects, becomes higher and higher as the frequency is decreased. In a plate, the measurements should be carried out at relatively high normalized frequencies ( $d\omega/c_s$ ) over a short propagation distance in order to avoid excessive attenuation upon immersion. In contrast, in a rod, the corresponding measurements should be carried out at relatively low normalized frequencies ( $a\omega/c_s$ ) over a long propagation distance to assure significant phase variation upon immersion. This trend was expected, as it has been shown previously that the radiation loading in inviscid fluids is a first-order effect in the case of a rod, whereas it is a second-order effect for a plate.<sup>17</sup> At first sight, this result could suggest that the radiation-induced velocity change is easier to observe in the rod than it is in the plate. This fact is true in a strictly inviscid fluid up to the point where the negative effect of fluid loading on the velocity of the fundamental dilatational mode in the rod reaches its maximum. However, this point is reached at a very low frequency. Because the effect in the case of the plate can be maintained, and actually increased, to much higher frequencies, it is possible to measure a much bigger effect in the plate by going to higher frequencies, where the otherwise inevitable viscosity effects become negligible. Theoretically, the quality factor can also be increased in the case of the rod by lowering the frequency, but the logarithmic nature of the increase makes this approach very ineffective. More importantly, as we shall show in the next section, lowering the frequency increases the otherwise negligible effects of the inevitable viscous drag between the solid and the fluid, thereby rendering the detection of radiation-induced velocity change less feasible.

## B. Viscous drag

The most important extraneous effect that can interfere with the experimental verification of the previously described opposite effect of radiation loading on the velocity of the fundamental dilatational modes in immersed thin plates

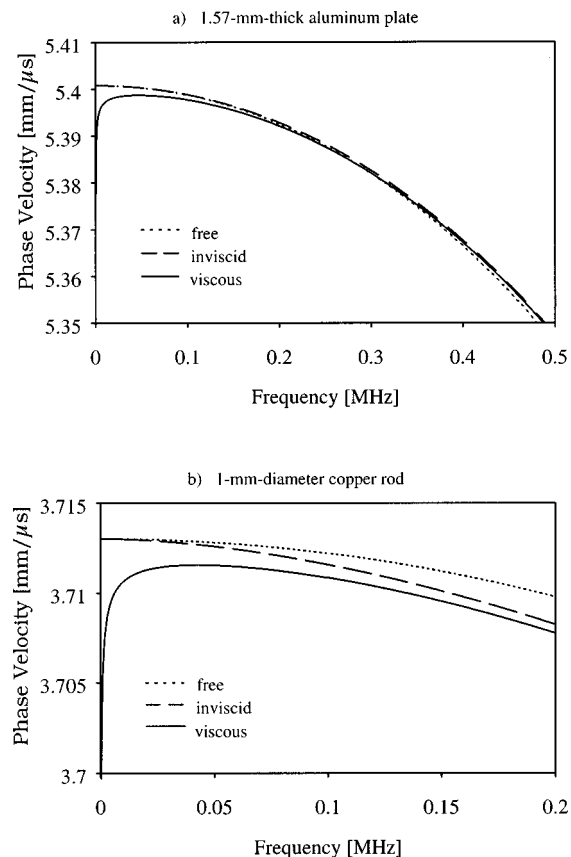


FIG. 3. Dispersion curves in an aluminum plate (a) and copper rod (b) immersed in water.

and rods is the inherent viscous drag presented by real fluids. Because the surface vibration of these modes at low frequencies is almost entirely tangential, even in low-viscosity fluids like water, fluid loading is dominated by viscous effects below a certain frequency. The viscosity-induced velocity change in thin plates and rods can be easily calculated by modeling the viscous fluid as a hypothetical isotropic solid having a frequency-dependent rigidity of  $-i\omega\eta$ , where  $\eta$  denotes the viscosity of the fluid.<sup>18,19</sup> In this way, the vorticity mode associated with the viscosity of the fluid is formally described as the shear mode in the fictitious solid. Figure 3 shows the dispersion curves in an aluminum plate (a) and copper rod (b) immersed in ideal inviscid and ordinary viscous water along with the corresponding dispersion curves in the free specimens. The density, dilatational velocity, shear velocity, and thickness of the aluminum plate were taken as  $\rho_{sa} = 2700 \text{ kg/m}^3$ ,  $c_{da} = 6323 \text{ m/s}$ ,  $c_{sa} = 3100 \text{ m/s}$ , and  $d = 1.57 \text{ mm}$ , respectively, to model the actual specimen used in our experiments. Similarly, the density, dilatational velocity, shear velocity, and radius of the copper rod were taken as  $\rho_{sc} = 8900 \text{ kg/m}^3$ ,  $c_{dc} = 4700 \text{ m/s}$ ,  $c_{sc} = 2260 \text{ m/s}$ , and  $a = 0.5 \text{ mm}$ , respectively. For water, the density was taken as  $\rho_f = 1000 \text{ kg/m}^3$ , the sound velocity as  $c_f = 1000 \text{ m/s}$ , and the viscosity as  $\eta = 10^{-3} \text{ kg/ms}$ .

The theoretically predicted effect of the viscosity of the fluid on the velocity of the fundamental dilatational wave in a thin plate is illustrated in Fig. 3(a). The dotted line shows the dispersion curve of this mode in the fluid-free plate; the dashed line shows the velocity of the same mode as a func-

tion of frequency in the plate as it is immersed in ideal inviscid water to represent the separate effect of radiation loading, whereas the solid line shows the dispersion curve for this mode as the plate is immersed in ordinary viscous water to represent the total fluid loading. Below 300 kHz, fluid loading actually decreases the phase velocity in the plate as a result of viscous drag. However, at higher frequencies, fluid loading is dominated by the radiation effect and the velocity increases with respect to the free plate. The effect of fluid loading on the velocity of the fundamental dilatational wave in a thin rod is illustrated in Fig. 3(b). Again, the dotted line shows the dispersion curve of this mode in the fluid-free specimen, and the dashed line shows the velocity of the same mode as a function of frequency in the rod as it is immersed in ideal inviscid water to represent the separate effect of radiation loading, whereas the solid line shows the dispersion curve for this mode as the rod is immersed in ordinary viscous water to represent the total fluid loading. We notice that as the copper wire is loaded by inviscid water, the velocity of the dilatational wave decreases. This reduction in the velocity increases as the frequency is increased. When the wire is immersed in viscous water, there is an additional drop in the velocity of the dilatational mode caused by the viscosity of the fluid. As the frequency increases, the effect of viscosity diminishes and the effect of leakage into the fluid becomes dominant.

Since our main interest is to measure the radiation part of fluid loading on this mode of wave propagation, and since the always-negative viscosity effect dominates at low frequencies while the radiation effect diminishes, it is necessary to limit our experimental measurements to a minimum frequency where the viscosity effect becomes negligible. However, the highest possible frequency in the case of the rod was limited by the fact that the reduction in the velocity of this mode is a low-frequency effect, and it reverses sign at relatively modest frequencies after which the velocity of this mode in the rod actually increases. In the case of the plate, such maximum does not exist as radiation loading always increases the phase velocity of the fundamental dilatational mode, regardless of the frequency. It should be mentioned, however, that once the dispersion becomes excessive we can no longer use the approximations given in Eqs. (5) and (7). In conclusion, viscosity will adversely affect our ability to measure separately the radiation component of the fluid-loading induced velocity drop in thin rods, while the leaky attenuation limits our measurements to low frequencies. In comparison, viscosity will have negligible effect on the measurement of the radiation component of the fluid-loading induced velocity increase in thin plates, when the leaky attenuation limits our measurements to high frequencies anyway. The experimental results and the appropriate frequency ranges for the two cases will be presented and fully discussed in the following section.

## II. EXPERIMENTAL METHOD

The experimental setup used in measuring the relative change in the velocity of the fundamental dilatational waves in thin plates and rods due to fluid loading is shown in Fig. 4. Both transmission and reception were accomplished by sym-

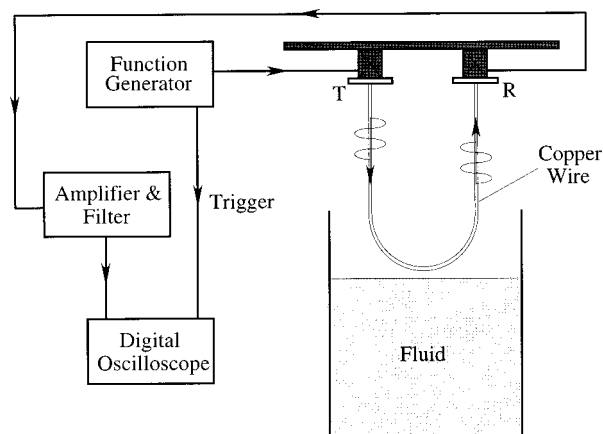


FIG. 4. The experimental setup used in investigating the fluid-loading effect on the dilatational wave velocity in thin rods and plates.

metrically mounted longitudinal transducers to reject all flexural modes. However, the much lower group velocity of these modes in thin plates and rods made it possible to further eliminate spurious flexural vibrations by simply gating out the faster fundamental longitudinal mode. The arrangement shown in Fig. 4 is for the case of a thin rod and is also used for the thin plate with minor modifications. The negative effect of fluid loading on the velocity of the lowest-order dilatation mode in a thin rod can be observed only in a narrow window limited from below by viscosity effects and also from above by a turning point beyond which the effect changes sign. Therefore, we had to use a very long and very thin wire which was bent as shown in Fig. 4. The leaky wave produced by the wire is diverging and will not interfere with the guided wave propagating in the wire even if it hits part of it. We made sure that the radius of curvature of the bent wire is high enough not to cause any perceivable attenuation or dispersion at the inspection frequencies used. In contrast, the positive effect of fluid loading on the velocity of the lowest-order dilatation mode in a thin plate can be observed in a wide frequency range anywhere above a lower limit set by viscosity effects. Therefore, we could use a relatively short plate without bending it like the wire shown in Fig. 4 (the sealed receiver was underwater). One additional difference between the wire and plate was the difference in material. A longitudinal contact transducer is very difficult to mount on the end of a thin rod, therefore, a copper wire was used which could first be soldered to a very thin copper foil which in turn was glued to the face of the transducer as described in Ref. 16. In the case of the plate, coupling is much easier, but the necessarily higher inspection frequency requires a less attenuating medium; therefore, an aluminum plate was used (at three times higher frequency, the grain-scattering induced attenuation in copper would be almost 100 times higher). However, it should be emphasized that our main goal is to experimentally demonstrate the opposite sign of the fluid-loading effect on rods and plates; therefore, the difference in material properties is of no particular importance. It should also be mentioned that the large dimensions of the water tank combined with the relatively short excitation and propagation times of the tone burst signals allowed us to completely disregard the acoustic modes generated in the fluid itself.

The basic arrangement uses two ultrasonic contact transducers in a pitch–catch mode of operation. One of the transducers acts as a transmitter and the other as a receiver. The transmitter is driven by a tone burst excitation generated by DS345 Stanford Research Systems synthesized function generator at the desired frequency. The number of cycles in the tone burst is adjusted according to the frequency, where a small number of cycles ( $\approx 3\text{--}5$ ) is used at low frequencies, and a large number of cycles ( $\approx 10\text{--}30$ ) is used in the high-frequency range. The transmitted signal is detected by the receiver and passed through an amplifier and bandpass filter, then displayed using a LeCroy 9310 digital oscilloscope. The function generator is also used to trigger the oscilloscope. The specific experimental arrangement and procedure for the plate and the rod are detailed later.

It should be mentioned that the velocity of the dilatational wave in thin plates and rods perceptibly decreases with increasing temperature. This extraneous effect was found to be more serious in thin rods which have low thermal capacity against temperature variations and require more accurate measurements at the same time. For example, the velocity of the fundamental dilatational mode in a copper wire decreased in a linear fashion with the temperature of the water bath at a rate of approximately  $6 \times 10^{-4}/^\circ\text{C}$ . With relatively simple means, the water temperature can be maintained at the ambient air temperature within better than  $\pm 1^\circ\text{C}$ , which effectively eliminates any gross temperature variation in the specimen as it is immersed in the water bath. However, an additional variation occurs when the specimen is subsequently taken out of the bath due to evaporation of the water from the surface, which is much more difficult to control. This evaporation causes the structure to cool temporarily, thereby increasing the velocity of the dilatational wave. Fortunately, this effect can be kept under control by waiting long enough for the temperature to stabilize as indicated by the received ultrasonic signal's returning to its original state recorded before immersion.

### A. Thin plate

A  $1.57 \times 267 \times 127$ -mm aluminum plate was used in this experiment (the density and dilatational and shear velocities of the specimen were given in the previous section). Two 1.0-MHz, 0.5-in.-diameter transducers were mounted on the two opposite edges normal to the longer sides of the plate and operated in pitch–catch mode. This arrangement proved to be better than the pulse–echo configuration as it reduced adverse cross talk, better separated the multiple reflections, and minimized the spurious reflections from the edges of the plate within the time window of interest. Figure 5(a) shows typical rf signals recorded at a carrier frequency of 950 kHz with a 200-kHz filter bandwidth. To reduce the effects of wave interaction with the surface of the water bath, approximately 5 mm of the plate is immersed in water even in the “free” state [top signal in Fig. 5(a)]. In the “immersed” state [bottom signal in Fig. 5(a)] the immersed part is increased to a larger value, e.g., 25 mm, i.e., by  $\ell = 20$  mm, using a computer-controlled translation stage. The actual distance  $\ell$  was varied at different frequencies based on the attenuation of the signal. Figure 5(b) shows the correspond-

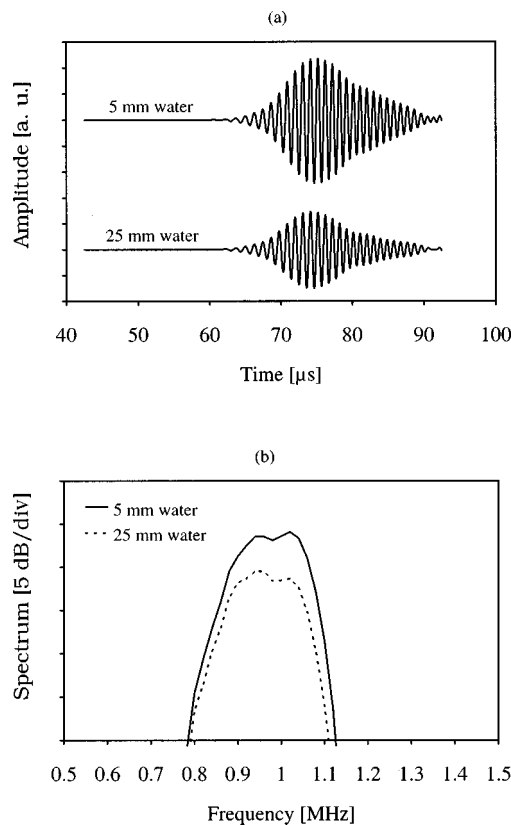


FIG. 5. Typical (a) rf signals in a 1.57-mm-thick aluminum plate immersed in water at 950 kHz, and (b) the corresponding spectrum.

ing spectra of the signals displayed in Fig. 5(a). An average fluid-loading induced attenuation of approximately  $\Delta L \approx 5$  dB was maintained during these measurements by varying the distance  $\ell$  from 30 mm at low frequencies to 10 mm at high frequencies.

It is possible to estimate the fluid-loading induced velocity change by simply monitoring a given zero crossing in the detected rf signal. However, careful inspection of the recorded rf signals revealed that some of the zero crossings exhibited higher or lower than average shift as a result of perceptible dispersion in spite of the relatively narrow bandwidth. This remnant dispersion, which is mainly caused by inevitable spurious standing waves interfering with the propagating wave, produces a localized phase variation that can be eliminated by averaging over the whole length of the signal. Therefore, the relative change in the velocity of the dilatational wave in the plate was determined by measuring the phase of the total rf signal recorded in the free and fluid-loaded states at the desired frequency. Calculating the phase of the total digitized rf signal by FFT for the free and the fluid-loaded cases gives a more realistic estimate of the change in the velocity than localized zero-crossing measurements. The phase difference  $\Delta\Phi$  between the loaded and unloaded spectra was then converted to the sought relative velocity change  $\Delta c/c_o$  via Eq. (1).

### B. Thin rod

A 1.0-mm-diameter, 500-mm-long copper wire was used in the measurements (the density and dilatational and

shear velocities of the specimen were given in the previous section). Two 2.25-MHz, 0.25-in.-diameter transducers were attached to the copper wire through thin copper foils that were soldered to the wire to ensure stable and good mechanical coupling. The two transducers were glued on the other side to an aluminum plate through two rubber cushions to create a stable structure that can be easily handled during the experiment. The two rubber cushions acted as vibration isolators that were necessary to isolate the transducers from the elastic aluminum plate, thereby minimizing any structural vibrations that can couple from the transmitter to the receiver through the plate. The copper wire was bent into a “U” shape and approximately 95% of its length was loaded by water.

The relative velocity change in this case was determined by monitoring a zero crossing in the rf signal with and without fluid loading. Unlike the case of the plate, all the zero crossings in this case seemed to be moving in the same direction with approximately the same amount indicating the absence of strong standing-wave interference patterns within the structure. The time shift of the monitored zero crossing was recorded and converted to relative velocity change using the following simple equation:

$$\frac{\Delta c}{c_o} \approx \frac{\Delta t c_o}{\ell}, \quad (9)$$

where  $\Delta t$  is the time delay of the zero crossing,  $\ell$  is the portion of the length of the wire that is loaded by water, and  $c_o$  is the low-frequency asymptotic value of the phase velocity in the free rod. Due to the finite length of the wire, and possibly to the finite curvature of the bent specimen, a weak standing-wave pattern is established in the wire as the longitudinal wave propagates from the transmitter side to the receiver side and reflects back to the transmitter side. This weak standing-wave pattern was reduced by averaging the results obtained at more than one zero crossing in the rf signal.

### III. EXPERIMENTAL RESULTS AND DISCUSSION

In this section, we present our experimental results on the effect of fluid loading on the velocity of the fundamental dilatational modes in thin plates and rods and compare them with the theoretical predictions obtained by numerically solving the corresponding exact dispersion equations for the cases of inviscid and viscous fluid loading. We will start by presenting the results for the plate and follow that with the results for the rod.

#### A. Thin plate

Figure 6 shows the relative change in the velocity of the fundamental dilatational mode in a 1.57-mm-thick aluminum plate immersed in water as a function of frequency. The change in the velocity is normalized to the low-frequency asymptotic value ( $c_o = 5400$  m/s) of the velocity of the dilatational mode in the free plate. The solid circles represent the average of five experimental measurements of the relative change in the velocity at the same excitation frequency, and the bars superimposed on them represent the estimated error

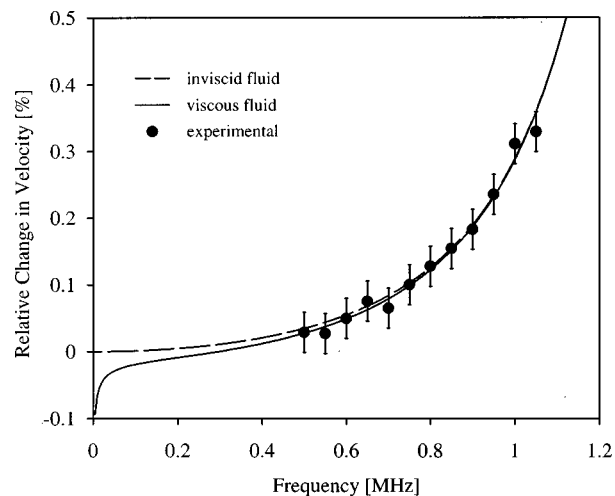


FIG. 6. Relative change in the velocity of the dilatational wave in a 1.57-mm-thick aluminum plate immersed in water.

in the measurements ( $\approx \pm 0.03\%$ ), which was based on the largest scatter in the measured data. The dashed line represents the theoretically predicted relative change in the velocity of this mode for the plate immersed in ideal inviscid water. This includes only the effect due to energy leakage into the fluid. The solid line shows the results for the case of viscous water which includes both the component due to leakage into the fluid and the component due to viscosity which is relatively weaker at high frequencies where the measurements were made. The frequency range in which the experimental measurements were performed extends from 520 kHz to 1.0 MHz. Excessive attenuation and dispersion of this mode restricted the highest frequency at which we could measure this effect. The lowest frequency at which the measurements were carried out was limited mainly by the bandwidth of the ultrasonic transducers, by the viscosity effect, which becomes more dominant below about 300 kHz, and by the reflections from the side edges of the plate which significantly distorted our signal below 500 kHz.

We can conclude from Fig. 6 that there is an excellent quantitative agreement between the experimentally measured relative change in the velocity of the fundamental dilatational mode in a thin plate and the corresponding predictions from theoretical models. The experimental results clearly verify that radiation loading indeed increases the velocity of the fundamental dilatational mode in a thin plate.

#### B. Thin rod

Figure 7 shows the relative change in the velocity of the fundamental dilatational wave in a 1.0-mm-diameter copper rod immersed in water. The change in the velocity is normalized to the low-frequency asymptotic value ( $c_o = 3710$  m/s) of the velocity of the dilatational mode in the free rod. The solid circles again represent the average of five experimental measurements of the relative change in the velocity at the same excitation frequency, and the bars superimposed on them represent the estimated error in the measurements ( $\approx \pm 0.015\%$ ), which was again based on the largest scatter in the measured data. In this case, due to the lack of strong interference patterns that could distort the detected signal,

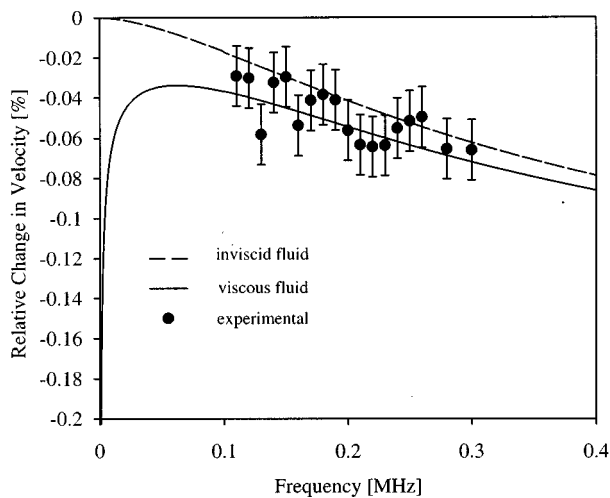


FIG. 7. Relative change in the velocity of the dilatational wave in a 1.0-mm-diameter copper wire immersed in water.

the simpler time-delay measurement of a zero crossing in the rf signal was used to determine the relative velocity change. The solid line in the figure represents the theoretical prediction in the case of a thin rod immersed in ordinary viscous water, and the dashed line represents the case of idealized inviscid water. The frequency range over which the experimental measurements were taken extends from 100 to 300 kHz. The highest measuring frequency in this case was limited by the fact that the negative radiation-loading effect in the case of the rod occurs only up to a certain limit, beyond which this effect becomes positive, and that even below this switch-over frequency the fluid-loading induced attenuation becomes excessively high for precise velocity measurements. This behavior was previously illustrated in Fig. 2 and discussed in a previous section. This measurement is particularly difficult since the measuring frequency range is also limited from below at about 100 kHz by inevitable viscosity effects which exert a much stronger influence on thin rods than plates. Although the radiation loading-induced velocity change in a thin rod is a first-order effect, its actual value is very small, which makes it extremely difficult to produce good agreement with theoretical predictions.

Considering the experimental difficulties associated with this measurement, the agreement between the experimentally obtained results and the theoretical predictions for the relative change in the velocity of the fundamental dilatational wave in a thin rod is good. These experimental results verify that radiation loading indeed decreases the velocity of the fundamental dilatational wave in a thin rod.

#### IV. CONCLUSIONS

In spite of the basic similarities between the fundamental dilatational wave modes in thin plates and rods, radiation loading exerts an opposite effect on their respective phase velocities. It was shown recently that this somewhat unexpected opposite effect is caused by the different nature of radiation loading in the two cases, as the radiation impedance exerted by the fluid on thin plates is purely real while it is dominantly imaginary in the case of thin rods.<sup>17</sup> In this paper, our main goal was to experimentally verify these ear-

lier analytical predictions. The relative change in the velocity of the dilatational mode in thin plates and rods was measured over a substantially wide frequency range. Simple theoretical considerations showed that, in the case of the rod, the frequency range over which the negative radiation-loading effect can be unequivocally observed is very narrow. The problem is less serious in the case of the plate, as the positive radiation-loading effect is maintained to much higher frequencies where viscosity effects are negligible. Because of these limitations on the frequency range over which the measurements could be carried out, the agreement between the experimental results and the theoretical predictions was better in the case of the plate as compared to the rod. Although the accuracy of the measurements in the case of the rod ( $\approx \pm 0.015\%$ ) is slightly better than that in the case of the plate ( $\approx \pm 0.03\%$ ), the absolute value of the detectable velocity change is much less for the rod than for the plate; therefore, positive verification of the predicted velocity change is more difficult in the case of the rod.

Considering the fact that the velocity effect is much smaller than the accompanying attenuation effect, and that its relative value is only on the order of a fraction of a percent, we can only conclude that such measurements are very demanding and require the minimization of all other effects that can mask the radiation-loading effect. Such effects mainly include the effect of viscosity of the fluid and to some degree, temperature variations. In our measurements, both of these effects were minimized to a level where the radiation-loading effect was dominant. Generally, the agreement between the experimental results and the theoretical predictions in both cases of thin plates and rods was good. The experiments unequivocally verified that radiation loading decreases the velocity of the dilatational wave in thin rods and increases it in thin plates. To our knowledge, these measurements constitute the first experimental verification of the previously reported opposite effect that radiation loading exerts on the velocity of the fundamental dilatational modes in immersed thin plates and rods.

<sup>1</sup>A. Schoch, "The transmission of waves through plates," (in German) *Acustica, Akust. Beih.* **2**, 1–17 (1952).

<sup>2</sup>M. F. M. Osborne and S. D. Hart, "Transmission, reflection, and guiding of an exponential pulse by a steel plate in water. I. Theory," *J. Acoust. Soc. Am.* **17**, 1–18 (1945).

<sup>3</sup>L. G. Merkulov, "Damping of normal modes in a plate immersed in a liquid," *Sov. Phys. Acoust.* **10**, 169–173 (1964).

<sup>4</sup>I. A. Viktorov, *Rayleigh and Lamb Waves* (Plenum, New York, 1967).

<sup>5</sup>L. E. Pitts, T. J. Plona, and W. G. Mayer, "Theoretical similarities of Rayleigh and Lamb modes of vibration," *J. Acoust. Soc. Am.* **16**, 374–377 (1976).

<sup>6</sup>I. T. Selezov, V. V. Sorokina, and V. V. Yakovlev, "Wave propagation in a thickness-inhomogeneous elastic layer bounded by fluid media," *Sov. Phys. Acoust.* **31**, 220–223 (1985).

<sup>7</sup>A. H. Nayfeh and D. E. Chimenti, "Propagation of guided waves in fluid-coupled plates of fiber-reinforced composite," *J. Acoust. Soc. Am.* **83**, 1736–1743 (1988).

<sup>8</sup>S. I. Rokhlin, D. E. Chimenti, and A. H. Nayfeh, "On the topology of the complex wave spectrum in a fluid-coupled elastic layer," *J. Acoust. Soc. Am.* **85**, 1074–1080 (1989).

<sup>9</sup>R. N. Thurston, "Elastic waves in rods and clad rods," *J. Acoust. Soc. Am.* **64**, 1–37 (1978).



- <sup>10</sup>J. A. Simmons, E. D. Krasicka, and H. N. G. Wadley, "Leaky axisymmetric modes in infinite clad rods. I," *J. Acoust. Soc. Am.* **92**, 1061–1090 (1992).
- <sup>11</sup>P. B. Nagy, "Leaky guided wave propagation along imperfectly bonded fibers in composite materials," *J. Nondestruct. Eval.* **13**, 137–145 (1994).
- <sup>12</sup>R. T. Harrold and Z. N. Sanjana, "Acoustic waveguide monitoring of the cure and structural integrity of composite materials," *Polym. Eng. Sci.* **26**, 367–372 (1986).
- <sup>13</sup>J. O. Kim, Y. Wang, and H. H. Bau, "The effect of an adjacent viscous fluid on the transmission of torsional stress waves in a submerged wave guide," *J. Acoust. Soc. Am.* **89**, 1414–1422 (1991).
- <sup>14</sup>M. Viens, Y. Tsukahara, C. Y. Yen, and J. D. N. Cheeke, "Leaky torsional acoustic modes in infinite clad rods," *J. Acoust. Soc. Am.* **95**, 701–707 (1994).
- <sup>15</sup>Y. Liu, M. Viens, and D. Cheeke, "Behavior of extensional modes in a thin rod immersed in a viscous medium," *J. Acoust. Soc. Am.* **99**, 706–712 (1996).
- <sup>16</sup>P. B. Nagy and R. M. Kent, "Ultrasonic assessment of Poisson's ratio in thin rods," *J. Acoust. Soc. Am.* **98**, 2694–2701 (1995).
- <sup>17</sup>W. Hassan and P. B. Nagy, "Why fluid loading has an opposite effect on the velocity of dilatational waves in thin plates and rods," *J. Acoust. Soc. Am.* **102**, 3478–3483 (1997).
- <sup>18</sup>P. B. Nagy and A. H. Nayfeh, "Excess attenuation of leaky Lamb waves due to viscous fluid loading," *J. Acoust. Soc. Am.* **101**, 2649–2658 (1997).
- <sup>19</sup>P. B. Nagy and A. H. Nayfeh, "Viscosity-induced attenuation of longitudinal guided waves in fluid-loaded rods," *J. Acoust. Soc. Am.* **100**, 1501–1508 (1996).

# Active impedance control within a cylindrical waveguide for generation of low-frequency, underwater plane traveling waves

James J. Finneran<sup>a)</sup> and Mardi C. Hastings<sup>b)</sup>

Department of Mechanical Engineering, The Ohio State University, Columbus, Ohio 43210

(Received 10 February 1998; accepted for publication 25 January 1999)

A cylindrical, water-filled acoustic waveguide with an active termination was used to generate constant-frequency, plane traveling waves. The waveguide was constructed of acrylic tubing with an underwater sound projector flanged to each end. One projector acted as the primary source and generated continuous harmonic waves. The active control system measured the transfer function between two hydrophones located inside the waveguide and used a pattern search algorithm to adjust the secondary source amplitude and phase in order to drive the measured transfer function to that for a plane traveling wave. The active control system was able to reduce the reflection coefficient to below 0.05 within the frequency range 12.5–400 Hz and adjust the acoustic pressure/particle velocity ratio to match that of a plane progressive wave. © 1999 Acoustical Society of America. [S0001-4966(99)02905-7]

PACS numbers: 43.20.Mv, 43.60.Qv [DEC]

## INTRODUCTION

A formidable problem in low-frequency underwater acoustic experiments is the design of a suitable testing environment in the laboratory. For many applications, it is important to know the nature of the acoustic field, including the relationship between pressure and particle velocity, in order to meaningfully interpret experimental data. A plane progressive or traveling wave is often a desirable stimulus because the relationship between the acoustic pressure and particle velocity is well known. Also, waves from any real source approach plane progressive waves at distances far from the source in open water.

In this study, a cylindrical, water-filled acoustic waveguide, with an active termination was used to generate constant-frequency, plane traveling waves within the 10–400 Hz frequency range. The waveguide was constructed of acrylic tubing, which is relatively flexible compared to water. This resulted in much larger attenuation and slower sound speed (and hence shorter wavelengths) than normally encountered in open water or in a rigid duct. These conditions helped achieve anechoic end conditions by reducing the amplitude of waves and the reflection coefficient at the waveguide exit, respectively. In a previous study using a 12-cm-inside diameter (i.d.) acrylic tube with a wall thickness of 3.2 mm, the sound speed was about 20% of the value for open water, and anechoic conditions were approximately realized at frequencies above 300 Hz (Finneran *et al.*, 1995; Hastings *et al.*, 1996).

The objectives of this study, however, required anechoic conditions at frequencies down to 10 Hz. Because passive absorption techniques typically rely on viscous dissipation, which requires an absorber thickness to be a sizable part of a wavelength (Parvulescu, 1967), they do not generally work at the low frequencies considered here. A passive termina-

tion was therefore considered impractical for this study and an active termination, featuring a second acoustic source, was necessary.

Although active control of sound in ducts is relatively common today (Elliott and Nelson, 1993), most applications involve active noise control in rigid, air-filled ducts. In these applications the acoustic pressure is minimized at some location or the sound propagating down a duct is canceled or reflected back toward the source, so that it does not propagate downstream. For this study, the objective was not to minimize pressure, but to use the secondary source to absorb the incident sound by matching the impedance at the waveguide exit to the medium characteristic impedance. This is known as active impedance control.

The use of a secondary source to absorb sound was originally proposed by Olson and May (1953) and has been demonstrated within ducts by various authors (e.g., Beatty, 1964; Guicking and Karcher, 1984; Orduña-Bustamante and Nelson, 1992); however, the majority of these studies have been performed in air. The few studies of active impedance control in water-filled tubes have been restricted to frequencies above 1000 Hz.

## I. BACKGROUND

### A. Attenuated plane waves in a flexible waveguide

For an acrylic tube in water, the attenuation is not negligible, thus some of the acoustic relations commonly used for rigid ducts were re-examined. For plane wave propagation within the waveguide, the acoustic pressure is

$$p(x,t) = [Ae^{jk(L-x)} + Be^{-jk(L-x)}]e^{j\omega t}, \quad (1)$$

and the axial particle velocity is

$$v(x,t) = \frac{k}{\rho\omega} [Ae^{jk(L-x)} - Be^{-jk(L-x)}]e^{j\omega t}, \quad (2)$$

where  $A$  represents the forward wave,  $B$  represents the reflected wave,  $j = \sqrt{-1}$ ,  $x$  is the distance from the source,  $t$  is

<sup>a)</sup>Current address: SPAWAR SYS CEN, D3503 PL-BS, 49620 Beluga Rd., San Diego, CA 92152.

<sup>b)</sup>Corresponding author: Dept. of Mechanical Engineering, 206 W. 18th Ave., The Ohio State University, Columbus, OH 43210.

time,  $\rho$  is the density of the fluid within the waveguide,  $\omega$  is the circular frequency,  $L$  is the length of the waveguide, and

$$\mathbf{k} = k - j\alpha = \frac{\omega}{c} - j\alpha \quad (3)$$

is the complex wave number ( $c$  is the sound speed within the waveguide and  $\alpha$  is the attenuation coefficient). From Eqs. (1) and (2), the pressure reflection coefficient  $R(x)$  is

$$R(x) = \frac{B}{A} e^{-j2\mathbf{k}(L-x)}, \quad (4)$$

and the acoustic impedance ratio  $\zeta(x)$  is

$$\zeta(x) = \frac{k}{\mathbf{k}} \left[ \frac{1+R(x)}{1-R(x)} \right]. \quad (5)$$

Measurement of the reflection coefficient or impedance ratio normally requires two closely spaced hydrophones. The locations of hydrophone 1 and hydrophone 2 are defined as  $x_1$  and  $x_2$ , respectively. The spacing between the hydrophones is defined as  $s = x_2 - x_1$  and the midpoint between the two hydrophones is  $\bar{x} = (x_1 + x_2)/2$ . The transfer function between hydrophones 1 and 2 is defined as

$$H(\bar{x}, \omega) = \frac{P(x_2, \omega) P^*(x_1, \omega)}{P(x_1, \omega) P^*(x_2, \omega)}, \quad (6)$$

where  $P(x, \omega)$  is the Fourier transform of  $p(x, t)$  and  $*$  indicates the complex conjugate. From Eqs. (1)–(3) and Eq. (6), the transfer function may be written as

$$H(\bar{x}, \omega) = \frac{R(\bar{x}, \omega) + e^{-j\mathbf{k}s}}{R(\bar{x}, \omega) e^{-j\mathbf{k}s} + 1}, \quad (7)$$

or as a function of the impedance ratio:

$$H(\bar{x}, \omega) = \frac{\zeta(\bar{x}, \omega) \mathbf{k} / k - j \tan(\mathbf{k}s/2)}{\zeta(\bar{x}, \omega) \mathbf{k} / k + j \tan(\mathbf{k}s/2)}. \quad (8)$$

### 1. No reflected wave

For this study the desired acoustic stimulus is a plane traveling wave within the waveguide. This is achieved in theory by having an infinite length duct or perfectly absorbing termination, therefore no reflected wave is generated. For no reflected wave,  $B=0$  and  $R(x)=0$ . The expression for the pressure amplitude along the length of the waveguide is

$$|P(x, \omega)| = P_0(\omega) e^{-\alpha x}, \quad (9)$$

where  $P_0(\omega)$  is the pressure amplitude at  $x=0$ . Substituting  $R(x)=0$  into Eq. (5) yields

$$\zeta(\bar{x}, \omega) = \frac{k}{\mathbf{k}} = \frac{k^2}{k^2 + \alpha^2} + j \frac{k\alpha}{k^2 + \alpha^2}. \quad (10)$$

Normally, a plane traveling wave is described as having  $\zeta(x, \omega) = 1 \angle 0^\circ$ ; however, Eq. (10) reveals that if the attenuation is nonnegligible, the impedance ratio has a nonzero imaginary part and a real part with magnitude less than one. If  $\alpha^2 \ll k^2$ , then Eq. (10) reduces to

$$\zeta(\bar{x}, \omega) = 1 + j \frac{\alpha}{k}. \quad (11)$$

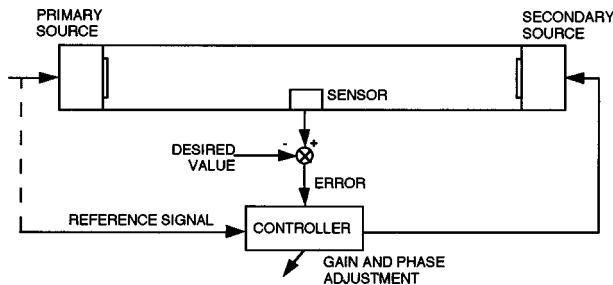


FIG. 1. Block diagram for active sound absorption within a duct.

With no reflected wave the transfer function between hydrophones 1 and 2 reduces to

$$H(\bar{x}, \omega) = e^{-j\mathbf{k}s}. \quad (12)$$

Equation (12) is identical to the form for no attenuation, with the wave number  $k$  replaced with the complex wave number  $\mathbf{k}$ .

### B. Active control principles

Active impedance control systems generally follow the block diagram shown in Fig. 1. The desired value is a mathematical expression representing the sensor output if the incident sound is completely absorbed at the secondary source. The controller acts upon the error signal, along with a reference signal correlated to the primary source input, to generate the secondary source input. Essentially, the controller provides a gain and phase adjustment to the secondary source so that the impedance at the secondary source is equal to the medium characteristic impedance (Parvulescu, 1967).

To examine how a second acoustic source can be used to absorb sound, the waveguide is assumed to be a cylindrical duct with cross-sectional area  $A_c$  and length  $L$ , as shown in Fig. 2. The primary source is located at  $x=0$  and the secondary source is located at  $x=L$ . The volume velocities of the primary and secondary sources are  $Q_1$  and  $Q_2$ , respectively, and their electrical inputs are  $E_1$  and  $E_2$ , respectively.

Assuming plane wave propagation in the waveguide, the pressure and particle velocity are given by Eqs. (1) and (2). Applying the boundary conditions

$$v(0, t) = \frac{Q_1}{A_c} e^{j\omega t} \quad (13)$$

and

$$v(L, t) = \frac{Q_2}{A_c} e^{j\omega t} \quad (14)$$

to Eq. (2) allows one to solve for  $A$  and  $B$ . Equations (1) and (2) are then used to find the acoustic impedance ratio at any point within the tube:

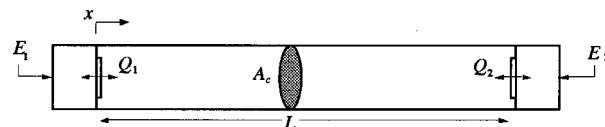


FIG. 2. One-dimensional model of active impedance control in the waveguide.

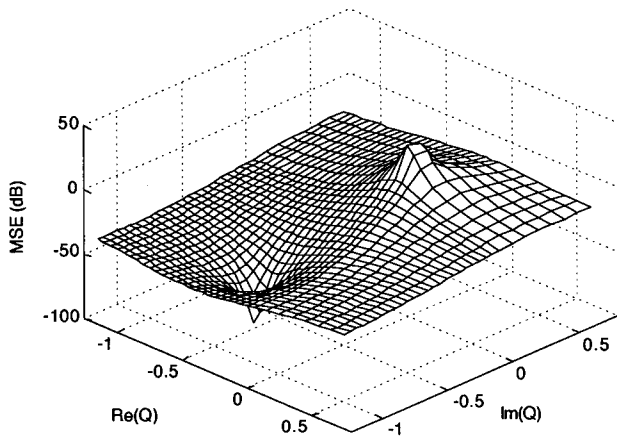


FIG. 3. Example of a performance surface for active impedance control in a cylindrical duct. The vertical axis is  $20 \log_{10} |E(\bar{x}, \omega)|^2$ ; the horizontal axes are the real and imaginary parts of the source input ratio  $Q$ .

$$\zeta(x, \omega) = \frac{-jk \cos \mathbf{k}(L-x) - Q \cos \mathbf{k}x}{\mathbf{k} \sin \mathbf{k}(L-x) + Q \sin \mathbf{k}x}, \quad (15)$$

where  $Q = Q_2/Q_1$ . At the waveguide exit,

$$\zeta(L, \omega) = \frac{-jk \frac{1-Q \cos \mathbf{k}L}{Q \sin \mathbf{k}L}}{\mathbf{k}}. \quad (16)$$

A one-to-one correspondence exists between  $Q$  and  $\zeta(L, \omega)$ , so manipulation of  $Q$  allows any desired ratio to be obtained (Bobber, 1962). For an anechoic termination,  $\zeta(L, \omega) = k/\mathbf{k}$ , thus

$$Q_{\text{anechoic}} = e^{-j\mathbf{k}L}. \quad (17)$$

If the attenuation is zero, then Eq. (17) requires the secondary source to have the same motion as the primary source, delayed by the time it takes the sound to travel from the primary source to the secondary source.

Rather than manipulate the acoustic impedance ratio directly, it is more useful to define an error signal so that when the error signal is zero, anechoic conditions exist. For active impedance control a suitable error is:

$$E(\bar{x}, \omega) = H(\bar{x}, \omega) - H_{12}(\bar{x}, \omega), \quad (18)$$

where  $E(\bar{x}, \omega)$  is the error,  $H(\bar{x}, \omega)$  is the transfer function measured between two hydrophones located at  $x_1$  and  $x_2$ , and  $H_{12}(\bar{x}, \omega)$  represents the *ideal* transfer function between hydrophones 1 and 2 (Orduña-Bustamante and Nelson, 1992).

For a plane traveling wave,

$$H_{12}(\bar{x}, \omega) = e^{-j\mathbf{k}s}. \quad (19)$$

Manipulation of Eqs. (8), (15), (18), and (19) allows one to solve for the error in terms of  $Q$ :

$$E(\bar{x}, \omega) = \frac{\cos \mathbf{k}(L-x_2) - Q \cos \mathbf{k}x_2}{\cos \mathbf{k}(L-x_1) - Q \cos \mathbf{k}x_1} e^{j\mathbf{k}s}. \quad (20)$$

If the mean squared error (MSE), defined as  $|E(\bar{x}, \omega)|^2$ , is plotted as a function of the real and imaginary parts of  $Q$ , a hypersurface, known as the performance surface, is generated. Figure 3 shows an example of a performance surface for active impedance control in a duct. The objective of the

controller in Fig. 1 is to adjust  $E_2$  so as to move the error toward the minimum point on the performance surface, that is, to minimize the MSE. This has been done by various authors using different techniques.

## 1. Control methodologies

Beatty (1964) used a theoretical relationship analogous to Eq. (17) to derive  $Q_{\text{anechoic}}$  for a water-filled steel tube. Using the theoretical value, anechoic conditions were obtained at 1000 Hz and 1600 Hz.

Guicking and Karcher (1984), Guicking *et al.* (1985), and Karcher (1982) used a two-microphone technique to separate the standing wave field into incident and reflected wave components. A signal proportional to the incident wave was passed through a control filter and used to drive the secondary source. Reflection coefficients less than 0.05 were obtained from 100 to 800 Hz in air. For best results, however, the control filter required manual adjustment at each frequency.

Curtis *et al.* (1990) manually adjusted the amplitude and phase of a secondary source to achieve anechoic end conditions in an air-filled duct.

A substantial improvement over earlier techniques was presented by Orduña-Bustamante and Nelson (1992), who used a time-domain adaptive control scheme featuring the filtered-x LMS algorithm to provide an absorbing termination in an air-filled tube. The adaptive nature of the controller allowed use with periodic, random, and transient signals. Anechoic end conditions were achieved from 30 to 330 Hz.

## II. ACOUSTIC WAVEGUIDE

### A. Design

The waveguide was fabricated from 18-cm i.d., 6.35-mm-thick acrylic tube. Acrylic flanges were cut and attached to individual sections of tube, allowing the sections to be bolted together for a total length of approximately 14 m. At 0.1-m intervals along the length of the waveguide, 9.5-mm holes were drilled to allow hydrophone access. Naval Research Laboratory (NRL) Type J13 underwater sound projectors were flanged to each end of the waveguide. The assembled waveguide was filled with water and submerged in a water-filled channel measuring approximately  $15 \times 1.2 \times 1$  m.

### B. Experimental characterization

Experimental measurements were performed to evaluate the waveguide behavior without using active control. These experiments were specifically designed to provide information regarding (1) the highest frequency at which plane wave propagation was valid, (2) the lowest frequency, if any, at which traveling waves existed without active control, (3) the phase velocity, and (4) the attenuation. During these tests a second J13 was located at the far end of the waveguide, but was used as only a passive termination.

#### 1. Radial behavior

To determine the highest frequency at which the plane wave assumption was valid, a B&K 8103 miniature hydro-

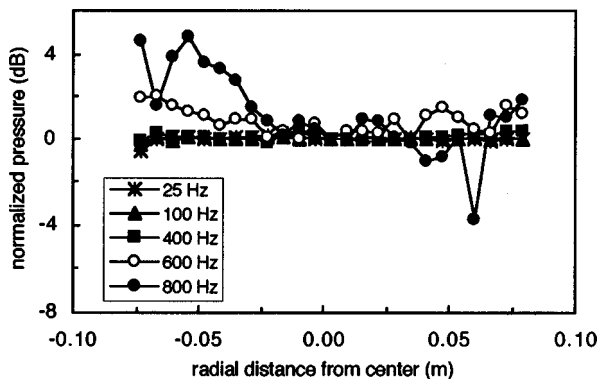


FIG. 4. Experimentally measured radial pressure for 25, 100, 400, 600, and 800 Hz.

phone was used to experimentally measure the acoustic pressure along the waveguide cross section at several different frequencies.

Figure 4 shows the normalized pressure measured across the waveguide diameter for frequencies of 25, 100, 400, 600, and 800 Hz. At each frequency, the data have been normalized to give 0 dB at  $r=0$ . Any deviation from the 0-dB level indicates departure from the plane wave assumption. At low frequencies the curves are flat, indicating plane waves. As the frequency increases, the measured pressure begins to deviate from the 0-dB baseline; at 600 Hz and above the plane wave assumption is no longer valid.

## 2. Axial behavior

The acoustic pressure was measured along the longitudinal axis of the waveguide in order to determine whether traveling waves or standing waves were present. If a traveling wave exists, the pressure is given by Eq. (9). If a standing wave exists, the axial pressure data will consist of a series of nodes and antinodes, with the distance between adjacent nodes or antinodes equal to  $\lambda/2$ , where  $\lambda$  is the acoustic wavelength.

The axial pressure was measured using two B&K 8103 miniature hydrophones which were moved in a point-by-point fashion along the length of the waveguide. At each location, the acoustic pressures were measured at several frequencies. The particle velocity in the axial direction was also estimated from the measured (complex) pressures using a discretized equation of conservation of momentum for an ideal fluid.

Figure 5 shows the measured axial sound pressure level (SPL) at 12.5, 75, and 300 Hz. At high frequencies the data begin to approach the ideal behavior for a traveling wave. At low frequencies, standing waves are present. At 12.5 Hz, however, the data are somewhat misleading. When the acoustic wavelength is greater than twice the waveguide length, the standing wave pattern may no longer be present. Therefore, at very low frequencies, the acoustic impedance ratio is a better indicator of whether or not a traveling wave exists.

Figure 6 shows the acoustic impedance ratio measured along the waveguide. The data at 12.5 Hz now clearly show that a traveling wave does not exist; the impedance ratio at

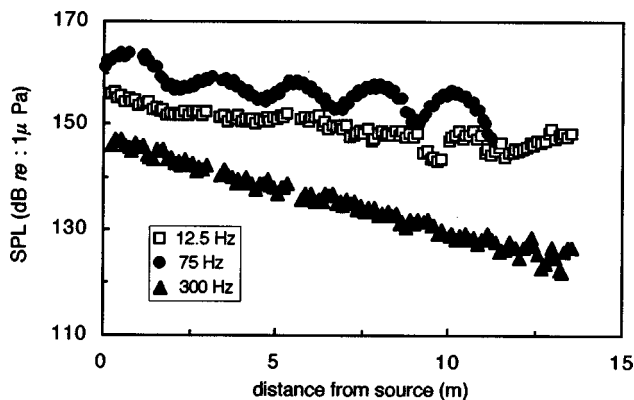


FIG. 5. SPL measured along the waveguide longitudinal axis.

12.5 Hz is primarily imaginary. At 350 Hz a traveling wave is approached, except at locations close to either the source or the termination.

## 3. Phase velocity

Two methods were used to experimentally measure the phase velocity. At low frequencies, with standing waves present, the distance between adjacent pressure nodes was measured. This distance equals  $\lambda/2$ . Since the wavelength, phase speed, and frequency are related by  $\lambda = c/f$ , the phase velocity in the waveguide could be calculated since the frequency was known. At higher frequencies, when reflections

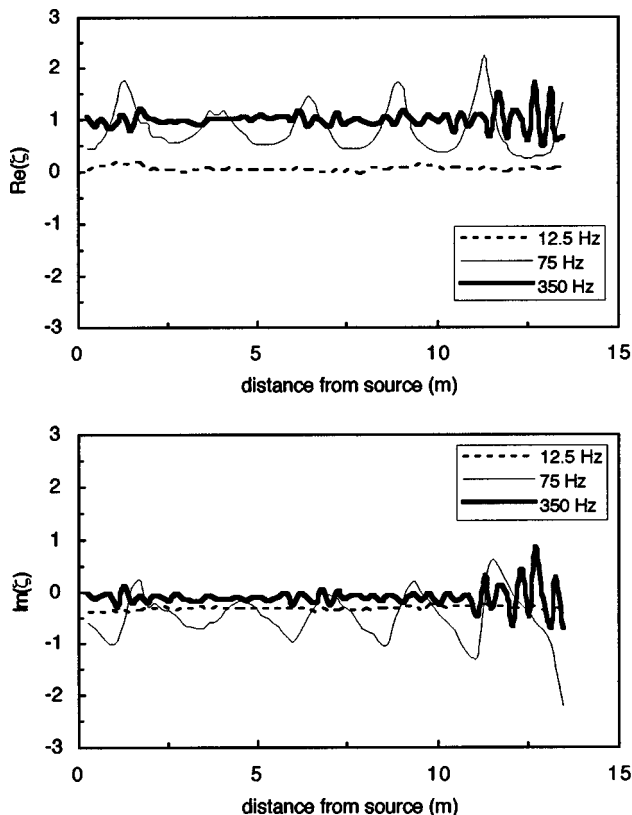


FIG. 6. Acoustic impedance ratio measured along the waveguide longitudinal axis. Top—real part; bottom—imaginary part.

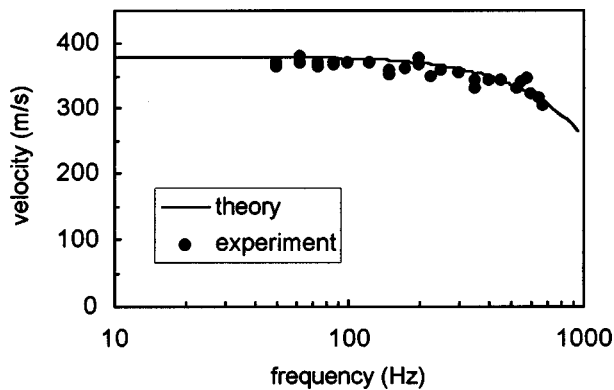


FIG. 7. Phase velocity as a function of frequency within the waveguide. The theoretical curve is based on the model of Junger (1955).

are less significant, the time delay between the pressure incident on two closely spaced hydrophones was measured and used to estimate the phase velocity.

Figure 7 shows good agreement between the experimentally measured phase velocity and the theoretical phase velocity based on Junger's analysis of the propagation of waves in a fluid-filled, elastic shell submerged in a fluid (Junger, 1955). This theory assumes an axisymmetric, thin shell (the actual ratio of waveguide diameter to wall thickness is 28). The material properties for acrylic were obtained from Read and Dean (1978). The phase velocity in the waveguide is approximately 370 m/s at frequencies below 200 Hz; this is about 25% of the sound speed in open water.

#### 4. Attenuation

Attenuation in the waveguide was also measured using two different methods, depending on whether standing waves or traveling waves were present. When traveling waves were present, a function having the form of Eq. (9) was fit (using a linear regression technique) to the axial pressure data. At low frequencies, the attenuation was estimated by fitting the equation

$$P_{\min}(x) = P_L \alpha (L - x), \quad (21)$$

where  $P_L$  is the pressure at the antinode closest to  $x=L$ , to the measured pressure minima (Kinsler *et al.*, 1982).

Figure 8 shows the measured attenuation as a function of frequency. The symbols represent the experimental data and

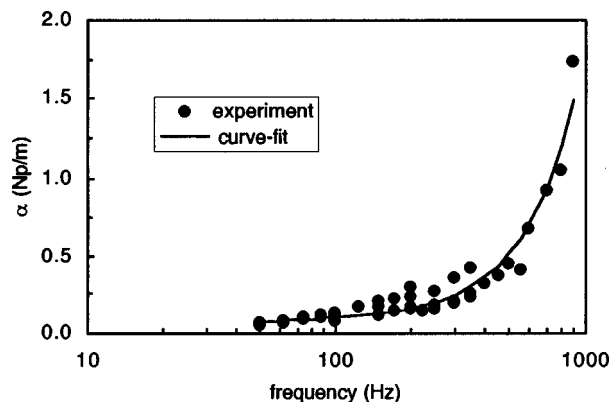


FIG. 8. Attenuation as a function of frequency in the waveguide.

the solid line is a curve fit to the data. It should be noted that the attenuation values here are much larger than those typically encountered in open water.

### III. ACTIVE IMPEDANCE CONTROL SYSTEM

#### A. Pattern search method

Although the adaptive controller used by Orduña-Bustamante and Nelson (1992) is well-suited for creating absorbing terminations in ducts, it was considered more complicated than necessary for this study. Here the primary source input is well-known and consists of a single frequency. The controller is only required to adjust the amplitude and phase of the secondary source to minimize the error signal. This can be done without resorting to adaptive methods if the problem is treated as one of multivariable optimization. In this case the independent variables are the real and imaginary parts of  $Q$  (or  $E_2$  if  $E_1$  is fixed), while the dependent variable is the MSE.

Optimization routines may be classified as search methods, which require function evaluation only, or gradient methods, which also require the Jacobian gradient vector (Adby and Dempster, 1974). For this study the pattern search technique was chosen for several reasons. It is robust in the presence of noise and tends to follow the line of steepest descent (Elliott *et al.*, 1987). Also, the pattern search technique has been successfully used in experimental studies to minimize the sound pressure of single-frequency fields in enclosures (Elliott *et al.*, 1987; Nelson and Elliott, 1987; Bullmore *et al.*, 1985).

The pattern search takes incremental steps, called pattern moves, after suitable directions have been found by local exploration. If the search progresses well, the step size is increased; otherwise the step size is reduced. When the step size is reduced below a set value, the search is ended (Adby and Dempster, 1974).

Figure 9 is a flowchart illustrating the operation of the pattern search. The successive base points are represented by the vector  $\mathbf{n} = [\text{Re}(Q) \text{Im}(Q)]$ , where  $\mathbf{n}_0$  is the initial base point and  $\mathbf{n}_i$  is the base point after  $i$  iterations. The variables  $\delta$  and  $m$  are the step size and the number of times the step size has been halved, respectively. The local exploration for the  $i$ th iteration defines a new local minimum  $\mathbf{n}_{A,i}$  and updates the search direction vector  $\mathbf{N}$ . If  $\mathbf{N} \neq \mathbf{0}$  and the function evaluation at the new local minimum is less than the value at the previous minimum,  $f(\mathbf{n}_{A,i}) < f(\mathbf{n}_{A,i-1})$ , a pattern move is made to a new base. Otherwise the increments are halved and a new pattern begun. After halving  $M$  times, the search is terminated and the answer given by the current value of  $\mathbf{n}$ .

#### B. Experimental implementation

Figure 10 shows the individual elements of the active control system. Two NRL Type J-13 underwater sound projectors acted as the primary and secondary sources. Two hydrophones (B&K 8103) were located near the test section. The hydrophone outputs passed into separate charge amplifiers (Kistler Type 5010) and then into the analog inputs of a National Instruments AT-DSP2200 DSP board within a 486 DX2/66 PC. The measured transfer function between the two

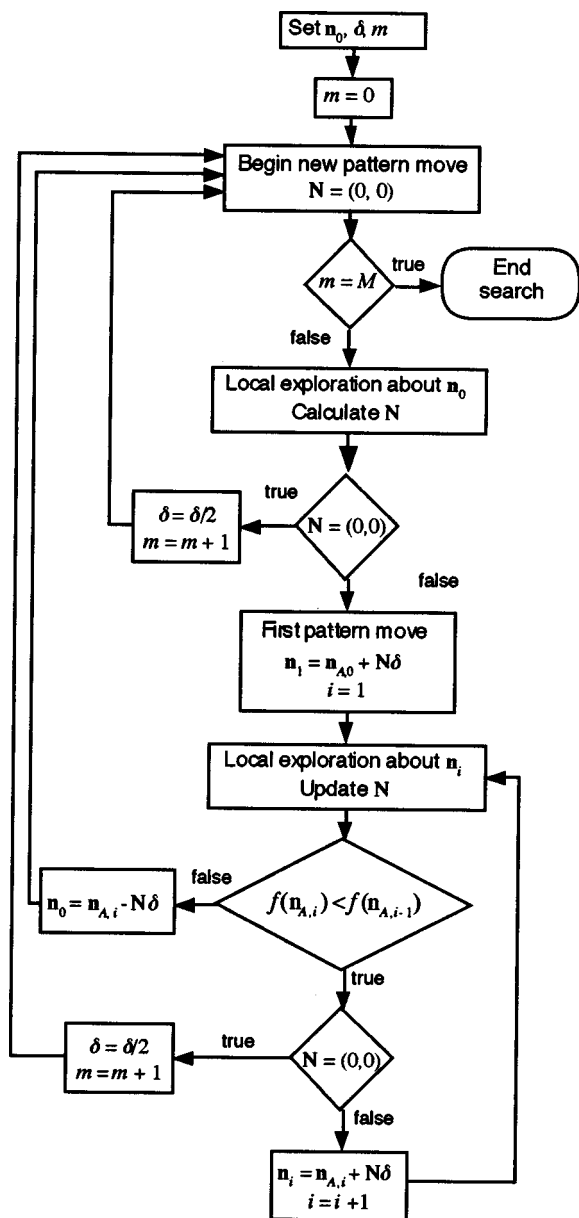


FIG. 9. Flowchart for the pattern search method.

hydrophones was corrected for phase mismatch according to ASTM E 1050-90 (ASTM, 1990). The two analog outputs of the DSP board went into separate channels of an audio amplifier (Crown Power Tech 1) and then to the primary and secondary sources. At frequencies above 150 Hz, audio transformers (Altech model 15300) were used to improve the impedance mismatch between the audio amplifiers and the J13's.

The pattern search algorithm was implemented in software using a custom virtual instrument written in LabVIEW®. The search was performed at a number of discrete frequencies between 12.5 and 500 Hz. At each individual frequency, the primary source was driven with continuous waves at a constant amplitude. The transfer function between hydrophones 1 and 2 was continuously measured and used to calculate the MSE. The secondary source amplitude and phase were then manipulated using the pattern search algorithm in order to minimize the MSE. Figure 11 is

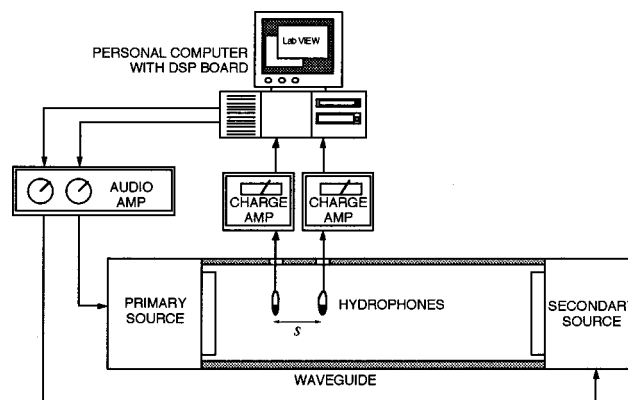


FIG. 10. Active control system hardware.

an example of the typical progress of the optimization routine, in this case at a frequency of 25 Hz.

At the conclusion of the pattern search, the impedance ratio  $\zeta(\bar{x}, \omega)$  and reflection coefficient  $R(\bar{x})$  at the test section were calculated from the measured transfer function. The primary source frequency was then adjusted to the next value and the process repeated. After repeating at all desired frequencies, the data for  $\zeta(\bar{x}, \omega)$  were compared to the theoretical prediction (for a traveling wave), given by Eq. (11), in order to assess the success of the optimization. The optimized values for the secondary source amplitude (at each frequency) were saved and could be used to later reproduce the traveling wave as long as environmental conditions had not changed dramatically. Changes in the channel water depth, ambient temperature, or barometric pressure normally required the optimization to be repeated every few days. Conditions within the waveguide could always be quickly checked by measuring the impedance ratio at each frequency using the optimized values for the secondary source input.

The performance of the active control system was limited at high frequencies by several factors. For a given hydrophone spacing, the maximum frequency was limited (according to ASTM E 1050-90) by the relation

$$s \ll \frac{c}{2f}. \quad (22)$$

For  $s=0.1$  m, the maximum frequency was experimentally observed to be near 400–500 Hz. More significant was the fact that, at higher frequencies, the impedance mismatch and

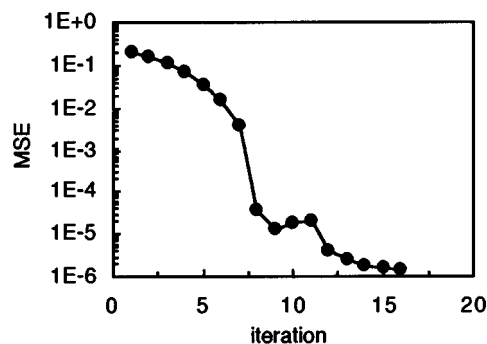


FIG. 11. Progress of the optimization routine at 25 Hz.

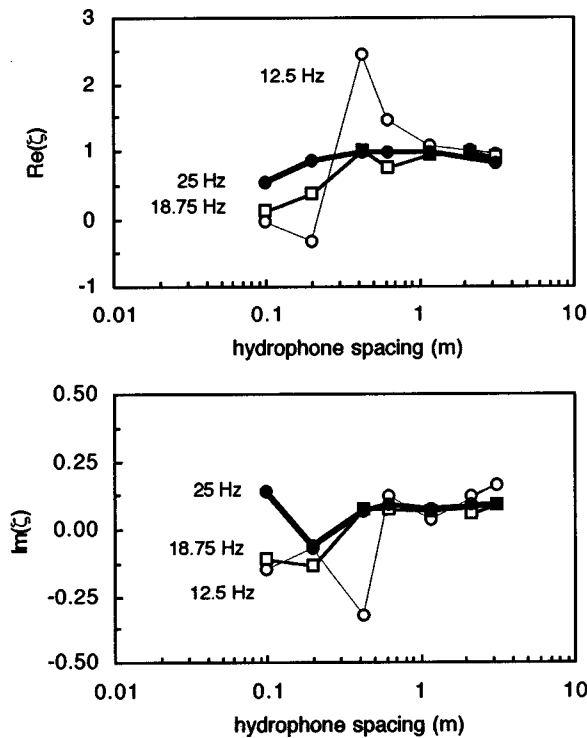


FIG. 12. Effect of hydrophone spacing on active control system performance at low frequencies.

attenuation increased dramatically. These conditions had the net effect of preventing the sound generated by the secondary source from having sufficient amplitude to cancel the primary source component. For these reasons the active control system behaved poorly above 400 Hz and was unstable above 500 Hz.

To investigate the low-frequency performance, the impedance ratio was measured (after optimization) for several hydrophone spacings. Figure 12 shows the results for 12.5 Hz, 18.75 Hz, and 25 Hz. These data indicate that as the hydrophone spacing decreases, the low-frequency performance suffers. At sufficiently large hydrophone spacings the results are acceptable, even at 12.5 Hz. Using Fig. 12 the range of useful hydrophone spacings may be estimated for each frequency: at 25 Hz the minimum spacing is 0.8 m, while at 12.5 Hz a spacing of more than 1 m is required. Overall, three different hydrophone spacings were used; Table I lists the hydrophone spacing used for each frequency range.

Figures 13 and 14 show some representative results for the active control system using the hydrophone spacings listed in Table I. Figure 13 compares the impedance ratios measured (at the test section) with and without the active control system. Figure 14 compares the reflection coefficient

TABLE I. Active control system hydrophone spacings used for each frequency range.

Frequency range (Hz)	Spacing (m)
$f < 38$	1.8
$38 \leq f \leq 112$	0.4
$f > 112$	0.1

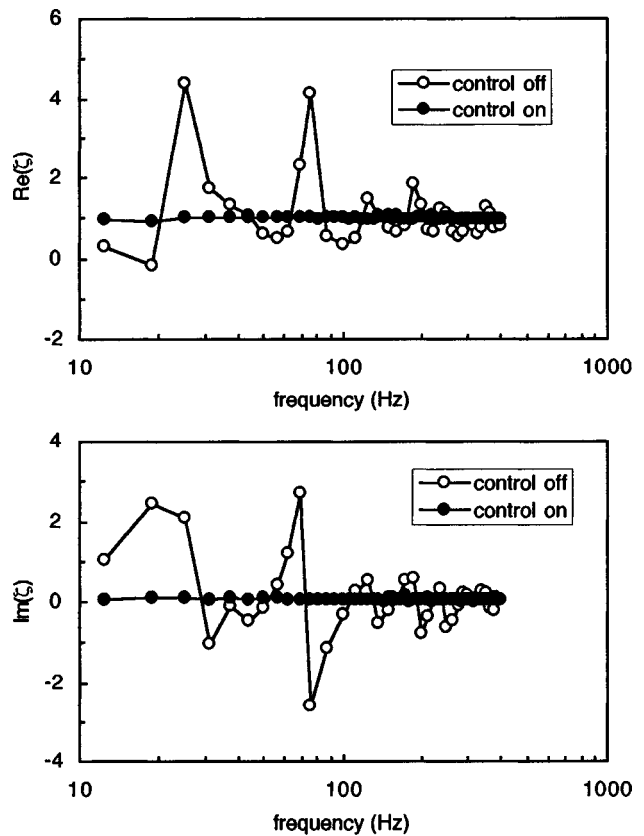


FIG. 13. Acoustic impedance ratio measured with and without active control. Top—real part; bottom—imaginary part.

(at the test section) measured with and without active control. Overall, the system was very successful in absorbing the incident sound and providing a traveling wave. The measured reflection coefficient was less than 0.05 between 12.5 and 400 Hz. It should be noted that this is the reflection coefficient at the test section, not at the waveguide exit. The reflection coefficient measured at the exit will always be large because of the presence of the secondary source.

### C. Comparison with an unbounded plane progressive wave

After successful optimization of the secondary source, or at frequencies where a traveling wave exists without active control, the impedance ratio within the waveguide is

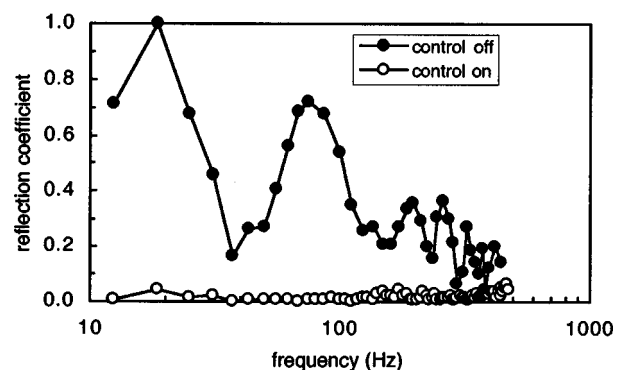


FIG. 14. Reflection coefficient measured with and without active control.



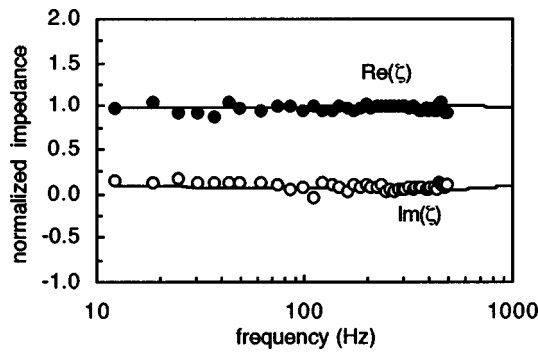


FIG. 15. Real and imaginary parts of the normalized acoustic impedance measured using the desired value from Eq. (26) in the active control system. The solid lines indicate the ideal impedance ratio,  $k/k$ .

given by Eq. (10), which indicates that the acoustic particle velocity for a traveling wave within the waveguide is

$$V(\bar{x}, \omega) = \frac{P(\bar{x}, \omega)}{\rho c} \left( \frac{\mathbf{k}}{k} \right). \quad (23)$$

For a plane progressive wave in open water the attenuation is negligible below 10 kHz, thus the particle velocity is

$$V_0(\bar{x}, \omega) = \frac{P(\bar{x}, \omega)}{\rho c_0}, \quad (24)$$

where  $c_0$  is the unbounded sound speed in water. The phase velocity in the waveguide is approximately one-fourth of the unbounded sound speed; therefore, for the same acoustic pressure, a traveling wave in the waveguide will have a particle velocity magnitude about *four times larger* than that of a traveling wave in open water.

To resolve this problem, a new desired value may be used in place of Eq. (19). The goal of the controller is now to generate a wave with

$$\frac{P(\bar{x}, \omega)}{\rho c_0 V(\bar{x}, \omega)} = \frac{k}{\mathbf{k}}. \quad (25)$$

The desired transfer function  $H_{12}(\bar{x}, \omega)$  is then

$$H_{12}(\bar{x}, \omega) = \frac{c_0/c - j \tan(\mathbf{k}s/2)}{c_0/c + j \tan(\mathbf{k}s/2)}. \quad (26)$$

Using this new value for  $H_{12}(\bar{x}, \omega)$ , the active control experiments were repeated, again using the setup shown in Fig. 10. To assess the controller's performance the acoustic impedance was measured and normalized by dividing by  $\rho c_0$ , rather than  $\rho c$ , as in Fig. 13. Figure 15 shows the measured normalized impedance as a function of frequency. The symbols represent the experimental data; the solid lines indicate the ideal performance, that is, the normalized impedance with zero error ( $=k/k$ ). The experimental results are very close to the ideal behavior. This indicates that the secondary source amplitude and phase have been successfully optimized in order to generate a traveling wave with a  $p/v$  ratio equivalent to that of a plane progressive wave in open water.

#### IV. SUMMARY AND CONCLUSIONS

This work has demonstrated an active control technique for generating harmonic acoustic traveling waves within a water-filled acrylic tube. The active control system measured the transfer function between two hydrophones and used the pattern search algorithm to adjust the secondary source amplitude and phase in order to drive the measured transfer function to some desired value. The pattern search was well-suited for this study, where the primary source generated continuous waves at a constant (known) frequency. For this reason, the use of newer, more complicated control methodologies was deemed unnecessary. The active control system was able to reduce the reflection coefficient to below 0.05 within the frequency range 12.5–400 Hz.

Operation at low frequencies was facilitated by increasing the spacing between the hydrophones. Operation at high frequencies was limited by the high attenuation within the waveguide and the large impedance mismatch between the audio amplifier and secondary source. Although a more powerful audio amplifier and better impedance match between the amplifier and the J13's would likely extend the useful range of the system, it is unlikely, because of the high attenuation in the waveguide and its radial pressure distribution, that the current configuration could be used at frequencies above 500 Hz.

The walls of the acrylic waveguide constitute a nonrigid boundary and slow the acoustic wave propagating within the waveguide. The sound speed within the waveguide is about 25% of the sound speed in open water, thus the traveling wave obtained using the desired value of Eq. (12) will have a particle velocity four times larger than that of a plane progressive wave with the same pressure. To compensate for this, the desired value was adjusted to that given by Eq. (26). The use of this desired value results in the creation of a traveling wave with a  $p/v$  ratio identical to that of a plane progressive wave.

In a manner similar to the derivation of Eq. (26), the desired value could be adjusted to create any  $p/v$  relationship. Work similar to this has been done by several authors in order to manipulate standing wave fields and create pressure nodes or antinodes at a desired location (van den Berg and Schuijff, 1985; Buwalda, 1981; Cahn *et al.*, 1969; Weiss, 1967; Parvulescu, 1961). The active control system described here could be easily adapted to perform such tasks.

#### ACKNOWLEDGMENTS

This work was supported by the Office of Naval Research, Grant No. N00014-94-1-0337. The authors would also like to thank Peter H. Rogers for suggesting that this system could be used to generate an unbounded plane progressive wave as described in Part C.

Aaby, P. R., and Dempster, M. A. H. (1974). *Introduction to Optimization Methods* (Chapman and Hall, London).

ASTM (1990). ASTM E 1050-90, "Standard test method for impedance and absorption of acoustical materials using a tube, two microphones, and a digital frequency analysis system" (American Society for Testing and Materials, Philadelphia).

- Beatty, L. G. (1964). "Acoustic impedance in a rigid-walled cylindrical sound channel terminated at both ends with active transducers," *J. Acoust. Soc. Am.* **36**, 1081–1089.
- Bobber, R. J. (1962). "Active load impedance," *J. Acoust. Soc. Am.* **34**, 282–288.
- Bullmore, A. J., Elliott, S. J., and Nelson, P. A. (1985). "The active minimization of periodic sound in enclosures," *Proc. Inter-Noise '85*, 575–578.
- Buwalda, R. J. A. (1981). "Segregation of directional and nondirectional acoustic information in the cod," in *Hearing and Sound Communication in Fishes*, edited by W. N. Tavolga, A. N. Popper, and R. R. Fay (Springer-Verlag, NY), pp. 139–171.
- Cahn, P. H., Siler, W. A., and Wodinsky, J. (1969). "Acoustico-lateralis system of fishes: Tests of pressure and particle velocity sensitivity in grunts, *Heamulon sciurus* and *Heamulon parrais*," *J. Acoust. Soc. Am.* **46**, 1572–1578.
- Curtis, A. R. D., Nelson, P. A., and Elliott, S. J. (1990). "Active reduction of a one-dimensional enclosed sound field: An experimental investigation of three control strategies," *J. Acoust. Soc. Am.* **88**, 2265–2268.
- Elliott, S. J., and Nelson, P. A. (1993). "Active noise control," *IEEE Trans. Signal Process.* **41**(10), 12–35.
- Elliott, S. J., Curtis, A. R. D., Bullmore, A. J., and Nelson, P. A. (1987). "The active minimization of harmonic enclosed sound fields, part III: Experimental verification," *J. Sov. Laser Res.* **117**, 35–58.
- Finneran, J. J., Hastings, M. C., Popper, A. N., and Lanford, P. J. (1995). "Effects of manmade underwater sound on the auditory organs and lateral line of the oscar (*Astronotus ocellatus*)—Preliminary results," *Proc. Inter-Noise '95*, 975–978.
- Guicking, D., and Karcher, K. (1984). "Active impedance control for one-dimensional sound," *J. Vib. Acoust. Stress Rel. Design* **106**, 393–396.
- Guicking, D., Karcher, K., and Rollwage, M. (1985). "Coherent active methods for applications in room acoustics," *J. Acoust. Soc. Am.* **78**, 1426–1434.
- Hastings, M. C., Popper, A. N., Finneran, J. J., and Lanford, P. J. (1996). "Effects of low-frequency underwater sound on hair cells of the inner ear and lateral line of the teleost fish *Astronotus ocellatus*," *J. Acoust. Soc. Am.* **99**, 1759–1766.
- Junger, M. C. (1955). "The effect of a surrounding fluid on pressure waves in a fluid-filled elastic tube," *J. Appl. Mech.* **22**, 227–231.
- Karcher, K. (1982). "Active modification of the acoustic wall impedance for normal incidence," Ph.D. thesis, Drittes Physikalisches Institut, University of Göttingen.
- Kinlser, L. E., Frey, A. R., Coppens, A. B., and Sanders, J. V. (1982). *Fundamentals of Acoustics*, 3rd ed. (Wiley, New York), pp. 206–210.
- Nelson, P. A., and Elliott, S. J. (1987). "The active minimization of acoustic fields," *J. Theor. Appl. Mech.* **6** (suppl.), 39–98.
- Olson, H. F., and May, E. G. (1953). "Electronic sound absorber," *J. Acoust. Soc. Am.* **25**, 1130–1136.
- Orduña-Bustamante, F., and Nelson, P. A. (1992). "An adaptive controller for the active absorption of sound," *J. Acoust. Soc. Am.* **91**, 2740–2747.
- Parvulescu, A. (1961). "Active electroacoustic networks and the synthetic absorber," *J. Acoust. Soc. Am.* **33**, 1668 (abstract).
- Parvulescu, A. (1967). "Problems of propagation and processing," in *Marine Bioacoustics*, edited by W. N. Tavolga (Pergamon, Oxford), pp. 87–100.
- Read, D. E., and Dean, G. D. (1978). *The Determination of Dynamic Properties of Polymers and Composites* (Wiley, New York).
- van den Berg, A. V., and Schuijf, A. (1985). "Acoustics of a standing wave tank for studying the hearing capacity of fish," *J. Acoust. Soc. Am.* **78**, 12–16.
- Weiss, B. A. (1967). "Sonic sensitivity in the goldfish (*Carassius auratus*)," in *Lateral Line Detectors*, edited by P. H. Cahn (Indiana University Press, Bloomington, Indiana), pp. 249–264.

# Wavefront and group velocity in relaxing and bubbly fluids

Johan L. Leander

Department of Military Technology, National Defence College, P.O. Box 27805, S-11593 Stockholm, Sweden

(Received 17 April 1997; accepted for publication 25 January 1999)

In this paper it is demonstrated that a theoretical model for wave propagation may indeed correspond to a well-posed transient problem although the group velocity for finite frequencies becomes greater than the high frequency limit of the phase velocity, negative or even infinite. Sufficient conditions for causality are derived and the particular cases of relaxing and bubbly fluids are considered so as to show some of the properties of the group velocity concept. © 1999 Acoustical Society of America. [S0001-4966(99)01505-2]

PACS numbers: 43.20.Px [DEC]

## INTRODUCTION

In an earlier work by this author it was shown how the group velocity is related to the wavefront speed of a pulse.<sup>1</sup> That work was an outgrowth of a long debate concerning interpretations of the group velocity concept for the case of transient wave propagation. However, due to the response to that paper, it seems to be convenient to expand on the transient problem. We here first look at sinusoidal acoustic waves in dissipative media in general. The quantities of special interest are the attenuation, the phase, and the group velocities, respectively. To simplify the transient analysis we restrict the attention to the regime of small dissipative effects which here means that the velocity dispersion and attenuation per wavelength are small compared to unity. This is of no limitation with respect to the purpose of the paper and will simplify the mathematics in the way that we can avoid branch cuts and Riemann sheets in the complex plane.<sup>2</sup> The special case of relaxing and bubbly fluids is considered so as to demonstrate the severe difficulties with the group velocity concept that may occur depending on the media of consideration.<sup>3,4</sup> For the relaxing fluid, the phase velocity is a monotonically increasing but bounded function of frequency. The corresponding group velocity has a maximum for a finite frequency. For the bubbly fluid, the phase velocity is at maximum for finite frequencies. Furthermore, the group velocity becomes negative for a certain frequency region and it is also found that it becomes infinite. We then turn to the transient problem of pulse propagation in dissipative fluids in general. A mathematical method handling arbitrary input pulses is suggested which thus is independent on the frequency content of the excitation. It is used to show that the wavefront speed is given by the high frequency limit of the phase velocity although the values of the phase and group velocities might be higher for finite frequencies. Sufficient conditions concerning whether a transient problem will yield a wavefront are discussed. It is demonstrated how the overall complex wave number is related to the causality condition. This is analogous to the Kramers–Kronig dispersion relations which connect the attenuation and the dispersion for a causal medium model.<sup>5,6</sup> We then again look at the particular

case of relaxing and bubbly fluids, and find that these models are consistent with the requirement of existence of a wavefront.

## I. MONOCHROMATIC WAVE PROPAGATION

### A. Complex wave number and group velocity

We consider here waves propagating in the positive  $x$ -direction where the amplitude is proportional to  $\exp[i(\omega t - kx)]$  where  $i$ ,  $t$ , and  $x$  are the imaginary unit, the time, and the propagation distance, respectively. The quantity  $\omega$  is the angular frequency and the complex wave number for the dissipative medium is denoted  $k(\omega)$ . It is convenient to decompose  $k(\omega)$  into its real and imaginary parts as

$$k(\omega) = \frac{\omega}{c(\omega)} - i\alpha(\omega), \quad (1)$$

where  $\alpha(\omega)$  and  $c(\omega)$  are the attenuation and the phase velocity.<sup>7</sup> With respect to the purpose of this paper it is convenient to consider the regime of small effects of the dissipative processes in the sense that the velocity dispersion and the attenuation per wavelength for all frequencies are both small compared to unity

$$\left| \frac{c(\omega) - c_0}{c_0} \right| \ll 1, \quad (2)$$

and

$$\frac{\alpha(\omega)c(\omega)}{\omega} \ll 1, \quad (3)$$

where  $c_0 = \omega/k_0$  denotes a reference speed of sound which for the relaxing fluid to be considered in the following subsection, B, is the low frequency limit of the phase velocity. For the bubbly fluid to be considered in subsection C,  $c_0$  is the speed of sound in the fluid surrounding the bubbles. The group velocity can be found from the real part of the complex wave number as<sup>1</sup>

$$u(\omega) = \left\{ \frac{d[\omega/c(\omega)]}{d\omega} \right\}^{-1} = \frac{c^2(\omega)}{c(\omega) - \omega \frac{dc(\omega)}{d\omega}}. \quad (4)$$

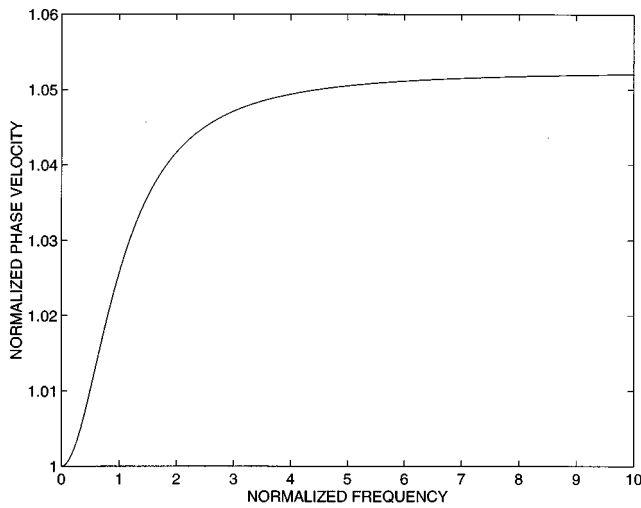


FIG. 1. Normalized phase velocity for relaxing fluid  $c_R(\omega)/c_0$  as a function of normalized frequency  $\omega\tau_s$ . Relaxation amplitude  $f_s=0.1$ .

### B. Relaxing fluids

In order to discuss the concepts of wavefront speed and group velocity it is convenient to consider specific models. In this work we have chosen relaxing and bubbly fluids since they are very different in terms of the behavior of the phase velocity and the attenuation and thus the distortion.<sup>8,9</sup> A relaxing fluid according to the Maxwell model is given by the dispersion relation

$$k_R^2(\omega) = k_0^2 \left[ 1 - \frac{i\omega\tau_s f_s}{1 + i\omega\tau_s} \right], \quad (5)$$

for the case of a single step relaxation process where  $f_s$  is the relaxation amplitude and  $\tau_s$  is the relaxation time. From Eqs. (1)–(3) and (5) the phase velocity and absorption are given by

$$c_R(\omega) = c_0 \left[ 1 - \frac{f_s \omega^2 \tau_s^2}{2(1 + \omega^2 \tau_s^2)} \right]^{-1}, \quad (6)$$

and

$$\alpha_R(\omega) = \frac{f_s \omega^2 \tau_s}{2c_0(1 + \omega^2 \tau_s^2)}. \quad (7)$$

In Figs. 1 and 2 the normalized phase velocity  $c_R(\omega)/c_0$  and the normalized attenuation  $\alpha_R(\omega)c_0\tau_s$  are shown as functions of normalized frequency  $\omega\tau_s$  for an relaxation amplitude  $f_s=0.1$ . It is seen that the phase velocity and the attenuation are monotonically increasing but bounded functions of frequency. For the group velocity we find from Eqs. (4) and (6)

$$u_R(\omega) = \frac{c_0}{1 - \frac{f_s \omega^2 \tau_s^2}{2(1 + \omega^2 \tau_s^2)} - \frac{f_s \omega^2 \tau_s^2}{(1 + \omega^2 \tau_s^2)^2}}. \quad (8)$$

It is to be noted that the above expression for the group velocity is not exactly the same as used in our earlier work.<sup>1</sup> This is so because we there used a Taylor expansion of the phase velocity. This fact, however, does not change the results and conclusions in the earlier work nor in this investi-

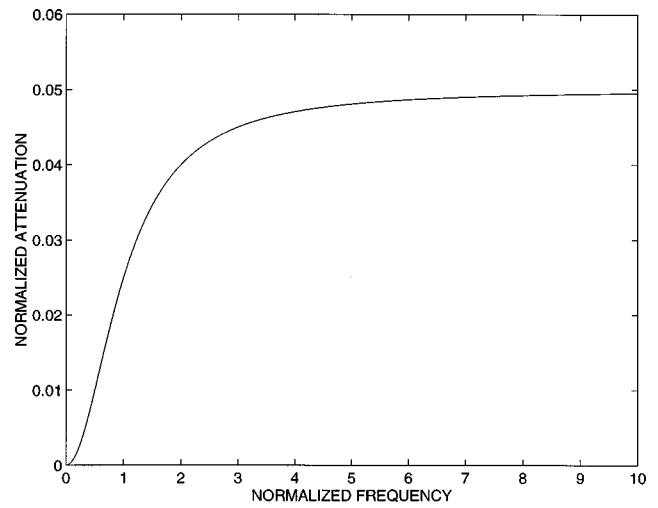


FIG. 2. Normalized attenuation for relaxing fluid  $\alpha_R(\omega)c_0\tau_s$  as a function of normalized frequency  $\omega\tau_s$ . Relaxation amplitude  $f_s=0.1$ .

gation. The normalized group velocity,  $u_R(\omega)/c_0$ , as function of normalized frequency,  $\omega\tau_s$ , is shown in Fig. 3 for  $f_s=0.1$ . It is seen that it attains its maximum value for a finite frequency. Moreover, this maximum is larger than the high frequency limit of the phase velocity.

### C. Bubbly fluids

In this subsection we look at sinusoidal waves in a bubbly fluid. The model originates from work by Kennard and has been derived and improved in various ways by many investigators over the years.<sup>5,10–12</sup> The dispersion relation is

$$k_B^2(\omega) = k_0^2 \left[ 1 + \frac{4\pi c_0^2 N R_s}{\omega_b^2 - \omega^2 + 2i\omega\beta_b} \right], \quad (9)$$

where  $R_s$  is the equilibrium bubble radius and  $N$  is the number of bubbles per unit volume and unit radius. The bubble resonance frequency is  $\omega_b^2 = \omega_i^2 + \omega_r^2$  and the total damping function is  $\beta_b = \beta_v + \beta_i + \beta_r$ .<sup>12</sup> From Eqs. (1)–(3) and (9) we find

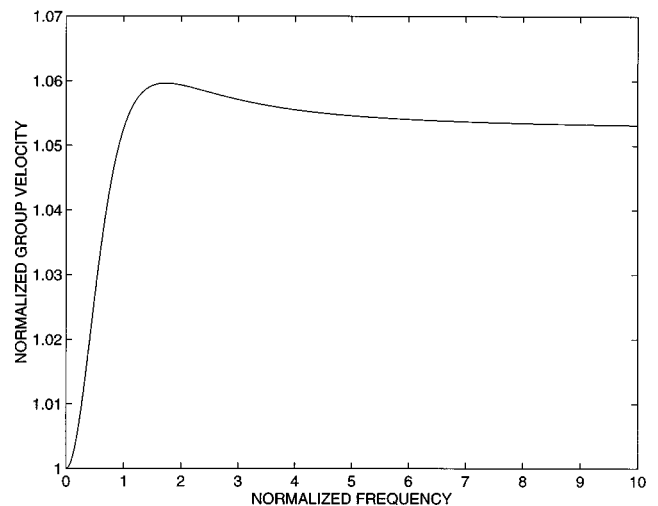


FIG. 3. Normalized group velocity for relaxing fluid  $u_R(\omega)/c_0$  as a function of normalized frequency  $\omega\tau_s$ . Relaxation amplitude  $f_s=0.1$ .

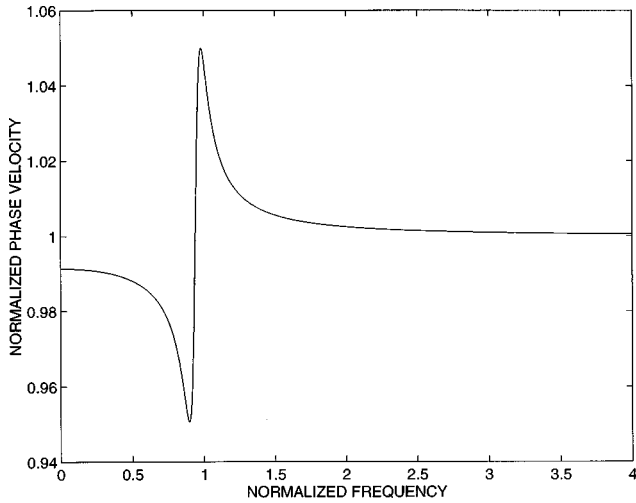


FIG. 4. Normalized phase velocity for bubbly fluid  $c_B(\omega)/c_0$  as a function of normalized frequency  $\omega/\omega_a$ . Gas volume fraction  $V=10^{-6}$ . Equilibrium bubble radius  $R_s=50$  microns.

$$c_B(\omega) = c_0 \left[ 1 + \frac{2\pi c_0^2 R_s N (\omega_b^2 - \omega^2)}{(\omega_b^2 - \omega^2)^2 + 4\omega^2 \beta_b^2} \right]^{-1}, \quad (10)$$

and

$$\alpha_B(\omega) = \frac{4\pi c_0 \omega^2 R_s \beta_b N}{(\omega_b^2 - \omega^2)^2 + 4\omega^2 \beta_b^2}. \quad (11)$$

For the purpose of this investigation it is convenient to introduce the following approximations;  $\omega_r(\omega) \approx \omega_r(\omega_a)$  and  $\beta_r(\omega) \approx \beta_r(\omega_a)$ , where  $\omega_a$  is the adiabatic resonance frequency.<sup>13</sup> We also neglect effects of liquid compressibility on the bubble motion so that  $\omega_r = \beta_r = 0$ . The result of the above assumptions is that the quantities  $\omega_b$  and  $\beta_b$  are constants.<sup>12</sup> These approximations will be used from here on and they are not in conflict with our coming conclusions concerning the wavefront speed and the group velocity. Rather, however, the mathematical machinery will be significantly reduced. In Figs. 4 and 5, the normalized phase velocity  $c_B(\omega)/c_0$  and the normalized attenuation  $\alpha_B(\omega)c_0/\omega_a$

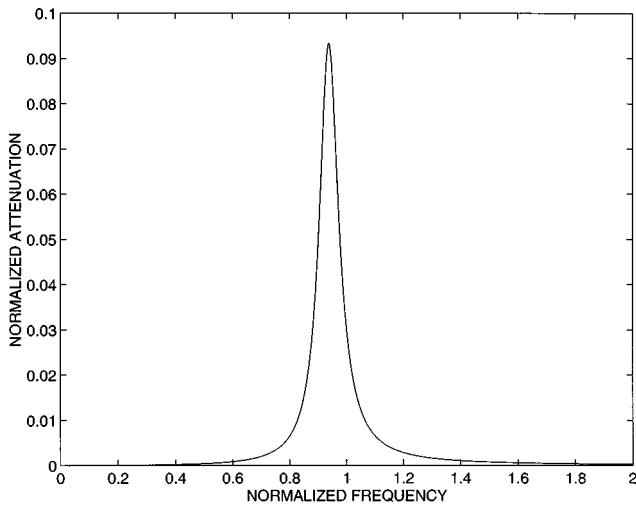


FIG. 5. Normalized attenuation for bubbly fluid  $\alpha_B(\omega)c_0/\omega_a$  as a function of normalized frequency  $\omega/\omega_a$ . Gas volume fraction  $V=10^{-6}$ . Equilibrium bubble radius  $R_s=50$  microns.

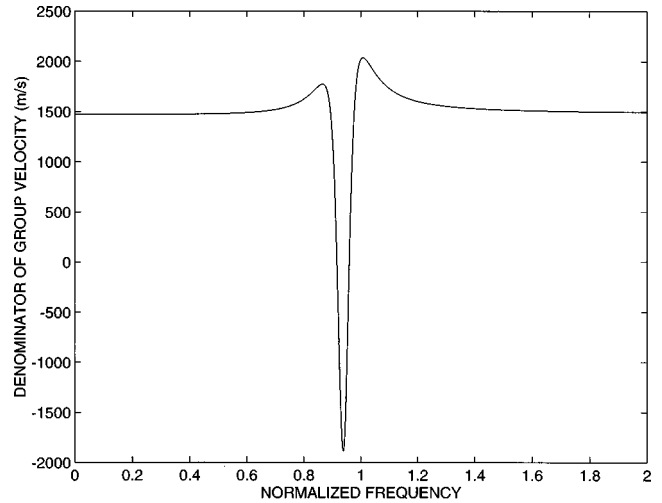


FIG. 6. Denominator of group velocity for bubbly fluid  $c_B^2/u_B$  as a function of normalized frequency  $\omega/\omega_a$ . Gas volume fraction  $V=10^{-6}$ . Equilibrium bubble radius  $R_s=50$  microns.

are shown as functions of normalized frequency  $\omega/\omega_a$ . The gas volume fraction,  $V=4\pi R_s^3 N/3$ , is  $10^{-6}$  and the equilibrium bubble radius,  $R_s$ , is 50 microns. It is seen that the phase velocity has a minimum and a maximum for finite frequencies. The attenuation has a sharp peak corresponding to the resonance region. The group velocity is found from Eqs. (4) and (10)

$$u_B(\omega) = \frac{c_0}{1 + C(\omega) + \frac{2\pi c_0^2 R_s N (\omega_b^2 - \omega^2)}{(\omega_b^2 - \omega^2)^2 + 4\omega^2 \beta_b^2}}, \quad (12)$$

where

$$C(\omega) = \frac{4\pi c_0^2 \omega^2 R_s N [(\omega_b^2 - \omega^2)^2 - 4\omega_b^2 \beta_b^2]}{[(\omega_b^2 - \omega^2)^2 + 4\omega^2 \beta_b^2]^2}. \quad (13)$$

The denominator of the group velocity,  $c_B^2/u_B$ , is shown in Fig. 6 for  $V=10^{-6}$  and  $R_s=50$ . It is seen that it becomes negative in a certain frequency region. Moreover, it is zero at two frequencies indicating that the group velocity becomes infinite.

## II. TRANSIENT WAVE PROPAGATION

### A. The mathematical method

A wavefront is here defined as a surface beyond which, at a given instant of time, the medium is completely at rest.<sup>3</sup> In this section we suggest a method for finding and defining the speed of the wavefront of a pulse; that is, the wavefront speed. The method presented is based on linear systems theory and the principles used are significantly influenced by the book by Papoulis.<sup>2</sup> In traveling through a dissipative medium, an acoustic pulse  $p(x,t)$  is distorted and its form and properties can be found by evaluating the Fourier integral

$$p(x,t) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} P(0,\omega) M(x,\omega) \exp(i\omega t) d\omega, \quad (14)$$

where

$$M(x, \omega) = \exp[-ik(\omega)x], \quad (15)$$

is the medium function and  $P(0, \omega)$  is the Fourier transform of the input pulse  $p(0, t)$ . In order to analyze the speed of the wavefront it is convenient to study the propagation of an initial dirac pulse for which  $P(0, \omega) = 1$  and the integral of interest is

$$p(x, t) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} M(x, \omega) \exp(i\omega t) d\omega. \quad (16)$$

The objective now is to demonstrate that the wavefront speed is given by the high frequency limit of the phase velocity. In order to do that we rewrite  $M(x, \omega)$  as

$$M(x, \omega) = M_{\infty}(x, \omega) + M_1(x, \omega), \quad (17)$$

where

$$M_{\infty}(x, \omega) = \exp(-ik_{\infty}x), \quad (18)$$

and

$$k_{\infty}(\omega) = \frac{\omega}{c_{\infty}} - i\alpha_{\infty}. \quad (19)$$

Further,

$$M_1(x, \omega) = M_{\infty}(x, \omega) \{\exp[Q(\omega)] - 1\}, \quad (20)$$

where the kernel of the whole transient problem,  $Q(\omega)$ , is given by

$$Q(\omega) = -ix[k(\omega) - k_{\infty}(\omega)]. \quad (21)$$

The high frequency limits of the phase velocity and the attenuation are here assumed to exist and are denoted  $c_{\infty}$  and  $\alpha_{\infty}$ , respectively. The distorted pulse is given by a corresponding decomposition as

$$p(x, t) = p_{\infty}(x, t) + p_1(x, t), \quad (22)$$

where from Eqs. (16)–(19)

$$p_{\infty}(x, t) = \exp(-\alpha_{\infty}x) \delta(t - t_{\infty}), \quad (23)$$

and from Eqs. (20) and (21)

$$p_1(x, t) = \frac{\exp(-\alpha_{\infty}x)}{2\pi} \int_{-\infty}^{\infty} \{\exp[Q(\omega)] - 1\} \times \exp[i\omega(t - t_{\infty})] d\omega, \quad (24)$$

where  $t_{\infty} = x/c_{\infty}$ . From Eq. (23) it is concluded that

$$p_{\infty}(x, t) = 0, \quad t < t_{\infty}. \quad (25)$$

In order to examine  $p_1(x, t)$ , the complex  $z$ -plane is introduced where  $z = \nu + i\omega$ . Let us study the contour integral

$$I = \frac{\exp(-\alpha_{\infty}x)}{2\pi i} \int_{CR} \{\exp[Q(z/i)] - 1\} \exp[z(t - t_{\infty})] dz, \quad (26)$$

in the right hand  $z$ -plane where  $\nu \geq 0$ . The contour,  $CR$ , is from *minus*  $i\Omega$  to *plus*  $i\Omega$  along the  $i\omega$ -axis and then connected with a semicircle in the right hand  $z$ -plane. For  $\Omega$  tending to infinity it is found that the part along the  $i\omega$ -axis corresponds to  $p_1(x, t)$ , Eq. (24), that is, the inverse Fourier transform of  $M_1(x, \omega)$ . The residue theorem will be used and let us first consider the requirements of convergence of the

integral for  $p_1(x, t)$ . Along the  $i\omega$ -axis  $\nu = 0$  so the factor in the integrand,  $\exp[i\omega(t - t_{\infty})]$ , represents sinusoidal oscillations. This means that if the magnitude,  $|\exp[Q(\omega) - 1]|$ , is square integrable, there is convergence and  $p_1(x, t)$  exists for every  $t$ .<sup>2</sup> Along the semicircle,  $\nu > 0$ , we have that if  $\{\exp[Q(z/i)] - 1\}$  tends to zero as  $z$  tends to infinity, and if  $t < t_{\infty}$ , the contribution will be zero. This is due to the Jordan's lemma.<sup>2</sup> Finally, if the sum of the residues of  $\{\exp[Q(z/i)] - 1\}$  in the right hand  $z$ -plane are zero, we find from the residue theorem that

$$p_1(x, t) = 0, \quad t < t_{\infty}. \quad (27)$$

Equations (22), (25), and (27) tell us that, under the circumstances presented,  $p(x, t)$  be zero for  $t < t_{\infty}$  and the model of concern is causal.<sup>1,2,5,6,8,9</sup> No information can propagate faster than with the high frequency limit of the phase velocity,  $c_{\infty}$ . This is the definition of the wavefront speed. For  $t > t_{\infty}$  an analogous contour is used in the left hand  $z$ -plane and there will, in general, be a nonzero contribution due to singularities of  $Q(z/i)$  for  $\nu < 0$ . For  $t = t_{\infty}$  we formally have<sup>2</sup>

$$p(x, t_{\infty}) = \frac{p(x, t_{\infty}^+) + p(x, t_{\infty}^-)}{2}. \quad (28)$$

In the above it is shown that if a medium model obey certain conditions, it will correspond to a transient problem with a wavefront propagating with a finite speed. The sufficient conditions are summarized as;

- (1) The existence of the high frequency limits of  $c(\omega)$  and  $\alpha(\omega)$ . The former as a positive, nonzero real constant. For the attenuation, the limit can be zero.
- (2) A behavior of  $Q(z/i)$  for  $z$  tending to infinity so that  $|\exp[Q(\omega) - 1]|$  is square integrable and that  $\{\exp[Q(z/i)] - 1\}$  tends to zero.
- (3) The sum of the residues of  $\{\exp[Q(z/i)] - 1\}$  is zero in the right hand  $z$ -plane. This is for example the case if  $Q(z/i)$  is analytic there.

In the next two subsections we will apply these conditions on relaxing and bubbly fluids to demonstrate that the wavefront speed may indeed be lower than the maximum of both the phase and group velocities at finite frequencies.

## B. Relaxing fluids

A relaxing fluid is considered here with respect to the findings in the earlier section. The high frequency limit of the phase velocity is found from Eq. (6) as

$$c_R(\infty) = c_{R\infty} = c_0 [1 - f_s/2]^{-1}. \quad (29)$$

The limit thus exists as a positive bounded real constant. For the attenuation we have from Eq. (7)

$$\alpha_R(\infty) = \alpha_{R\infty} = \frac{f_s}{2c_0\tau_s}, \quad (30)$$

which is in concordance with the requirements. The quantity  $Q_R(z/i)$  is from Eqs. (1), (6), (7), (21), and (29), (30),

$$Q_R(z/i) = \frac{xf_s}{2c_0\tau_s(1 + z\tau_s)}. \quad (31)$$

It is readily shown from Eq. (31) that  $|\exp[Q_R(\omega)-1]|$  is square integrable and that  $\{\exp[Q_R(z/i)]-1\}$  tends to zero as  $z$  tends to infinity. Furthermore,  $Q_R(z/i)$  has no singularities in the right hand  $z$ -plane but one in the left hand  $z$ -plane for  $z = -(1/\tau_s)$ . The relaxing fluid thus gives us an example where the group velocity, at finite frequencies, is higher than the wavefront speed.

### C. Bubbly fluids

The bubbly fluid is considered here and the high frequency limit of the phase velocity is found from Eq. (10) as

$$c_B(\infty) = c_{B\infty} = c_0. \quad (32)$$

The high frequency of the attenuation yields from Eq. (11)

$$\alpha_B(\infty) = \alpha_{B\infty} = 0. \quad (33)$$

The above results are thus consistent with the requirements. The quantity  $Q_B(z/i)$  is from Eqs. (1), (10), (11), (21), and (32), (33)

$$Q_B(z/i) = \frac{x2\pi c_0 R_s N z}{z^2 + 2\beta_b z + \omega_b^2}, \quad (34)$$

so we find from Eq. (34) that  $|\exp[Q_B(\omega)-1]|$  is square integrable and that  $\{\exp[Q_B(z/i)]-1\}$  tends to zero as  $z$  tends to infinity. Moreover,  $Q_B(z/i)$  has no singularities in the right hand  $z$ -plane but two in the left hand  $z$ -plane

$$z_1 = -\beta_b + i\sqrt{\omega_b^2 - \beta_b^2}, \quad (35)$$

and

$$z_2 = -\beta_b - i\sqrt{\omega_b^2 - \beta_b^2}. \quad (36)$$

The situation  $\omega_b > \beta_b$  corresponding to underdamped bubble oscillations is used here since it is of normal interest in bubble dynamics.<sup>13</sup> It is thus shown that the model for the bubbly fluid represents a causal transient problem. This gives us an example where the both the phase and group velocities, at finite frequencies, are higher than the wavefront speed. Furthermore, the presence of negative or even infinite group velocities are thus of no problem.

### III. CONCLUSIONS

In this work we have demonstrated that the wavefront of an acoustic pulse can propagate with a speed which is lower

than the maximum values of both the phase and group velocities at finite frequencies. We have suggested and used a model for transient analysis so as to show that the wavefront speed is given by the high frequency limit of the phase velocity. The particular cases of pulse propagation in relaxing and bubbly fluids has been analyzed. For the relaxing fluid, the group velocity, at finite frequencies, is higher than the wavefront speed. For the bubbly fluid, both the phase and group velocities become larger, at finite frequencies, than the wavefront speed. Moreover, the group velocity becomes negative and also infinite for finite frequencies. It has been clearly demonstrated that the group velocity scale is of limited applicability and should be used with great care.

<sup>1</sup>J. L. Leander, "On the relation between the wavefront speed and the group velocity concept," *J. Acoust. Soc. Am.* **100**, 3503–3507 (1996).

<sup>2</sup>A. Papoulis, *The Fourier Integral and Its Applications* (McGraw-Hill, New York, 1962).

<sup>3</sup>L. Brillouin, *Wave Propagation and Group Velocity* (Academic, London, 1960).

<sup>4</sup>J. L. Leander, "Wavefront speed and group velocity in bubbly fluids," *Proc. 4th European Conference on Underwater Acoustics I*, 113–118 (1998).

<sup>5</sup>S. Temkin, "Attenuation and dispersion of sound in bubbly fluids via the Kramers-Kronig relations," *J. Fluid Mech.* **211**, 61–72 (1990).

<sup>6</sup>J. L. Leander, "Comments on attenuation and dispersion in many spherical scatterer systems and the Kramers-Kronig relations," *J. Acoust. Soc. Am.* **101**, 1111–1114 (1998).

<sup>7</sup>S. Temkin, *Elements of Acoustics* (Wiley, New York, 1981).

<sup>8</sup>J. L. Leander, "Acoustic pulse propagation in Maxwell fluids," *J. Acoust. Soc. Am.* **94**, 1643–1650 (1993).

<sup>9</sup>J. L. Leander, "Transient wave propagation through bubbly layers via the Foldy-Twersky integral equation," *J. Acoust. Soc. Am.* **95**, 2378–2386 (1994).

<sup>10</sup>E. H. Kennard, "Radial motion of water surrounding a sphere of gas in relation to pressure waves," in *Underwater Explosion Research*, Vol. II, The Gas Globe, Office of Naval Research (1943).

<sup>11</sup>L. L. Foldy, "The multiple scattering of waves," *Phys. Rev.* **67**, 107–119 (1945).

<sup>12</sup>K. W. Commander and A. Prosperetti, "Linear pressure waves in bubbly liquids: Comparison between theory and experiments," *J. Acoust. Soc. Am.* **85**, 732–746 (1989).

<sup>13</sup>J. L. Leander, "On the eigenfrequency of a gas bubble in a liquid," *J. Acoust. Soc. Am.* **102**, 1900–1903 (1997).

# Parabolic equations for gravity and acousto-gravity waves

Joseph F. Lingeitch and Michael D. Collins  
*Naval Research Laboratory, Washington, DC 20375*

William L. Siegmann  
*Rensselaer Polytechnic Institute, Troy, New York 12180*

(Received 13 July 1998; accepted for publication 26 February 1999)

Parabolic equations for gravity and acousto-gravity waves are derived and implemented. The wave equations for these problems contain singularities at depths at which the buoyancy frequency equals the forcing frequency. One of the advantages of the parabolic equation solution is that it is easy to avoid numerical problems associated with the singularities. Some problems involve an infinite number of propagating modes. This artifact of neglecting viscosity is handled by including stability constraints in the rational approximations used in the implementation of the parabolic equation. The parabolic equation is tested for idealized problems involving surface, internal, and interface gravity waves. Parabolic equation solutions are also presented for range-dependent problems involving internal waves in the ocean and acousto-gravity waves in the atmosphere. © 1999 Acoustical Society of America. [S0001-4966(99)04906-1]

PACS numbers: 43.20.Mv, 43.28.Bj [DLB]

## INTRODUCTION

The parabolic equation method has been extensively applied to problems in ocean acoustics and seismo-acoustics.<sup>1</sup> In this paper, we develop parabolic equation techniques for gravity and acousto-gravity waves. The acousto-gravity wave equation is a generalization of the acoustic wave equation that includes both compressibility and buoyancy effects.<sup>2</sup> Solutions that correspond to limiting cases include pure acoustic waves (in which buoyancy is neglected) and pure gravity waves (in which compressibility is neglected). Both of these waves are commonly observed in the ocean (where they are decoupled) and in the atmosphere. Parabolic equation techniques have been developed for surface gravity waves in the horizontal plane over gradually varying bathymetry.<sup>3,4</sup> We consider gravity and acousto-gravity waves in the vertical plane, with applications to internal and surface gravity waves in the ocean and acousto-gravity waves in the atmosphere.

Two complications arise in generalizing parabolic equation techniques to acousto-gravity waves. In contrast to the acoustic wave equation, the acousto-gravity wave equation has an infinite number of propagating modes for some problems. It is not possible to handle such problems with parabolic equation techniques, which are based on rational approximations that are valid over a finite band of horizontal wave numbers. Fortunately, only a finite number of the modes are physically meaningful because small scale waves are annihilated by viscosity. These waves can be handled properly by including viscosity in the analysis. An alternate approach is to design the rational approximations to annihilate small scale waves. The other complication is related to singularities in the acousto-gravity wave equation, which are regular but must be handled with care in numerical solutions. An advantage of the parabolic equation solution is that its form makes it easy to avoid difficulties associated with the singularities.

The acousto-gravity wave equation and interface condi-

tions are formulated in Sec. I. A parabolic equation for acousto-gravity waves and an approach for generating initial conditions are derived in Sec. II. Energy-conservation<sup>5-7</sup> and single-scattering<sup>8</sup> techniques for handling range dependence are derived in Sec. III. Numerical solutions of the gravity and acousto-gravity wave equations are illustrated in Sec. IV.

## I. THE WAVE EQUATION AND INTERFACE CONDITIONS

We present the acousto-gravity wave equation and horizontal interface conditions in this section. For simplicity, we work in the frequency domain, consider a two-dimensional problem, and neglect the effects of nonlinearity, viscosity, and the earth's rotation and curvature. We work in Cartesian coordinates, where the range  $x$  is the horizontal distance from a source or reference point and  $z$  is the depth in the ocean or altitude in the atmosphere. We assume a steady ambient state in which the ambient flow can be neglected. Horizontal variations in density are consistent with this assumption provided they are sufficiently gradual. Linearizing the equation of state and the equations for conservation of momentum, mass, and entropy, and performing some manipulations,<sup>2</sup> we obtain

$$\left(\frac{\partial^2}{\partial x^2} + k^2\right)p = -i\omega\rho\left(\frac{\partial}{\partial z} + \frac{g}{c^2}\right)w, \quad (1)$$

$$(N^2 - \omega^2)w = \frac{i\omega}{\rho}\left(\frac{\partial}{\partial z} - \frac{g}{c^2}\right)p, \quad (2)$$

$$N^2 \equiv \frac{g}{\rho} \frac{\partial \rho}{\partial z} - \frac{g^2}{c^2}, \quad (3)$$

where  $p$  and  $w$  are perturbations to the ambient pressure and vertical component of velocity,  $\rho$  is the ambient density,  $k = \omega/c$  is the acoustic wave number,  $c$  is the sound speed,  $\omega$



is the circular frequency,  $g$  is the acceleration due to gravity, and  $N$  is the buoyancy frequency. Using Eq. (2) to eliminate  $w$  in Eq. (1), we obtain the acousto-gravity wave equation,

$$\frac{\partial^2 p}{\partial x^2} + \rho \frac{\partial}{\partial z} \left( \frac{1}{\rho} \frac{\omega^2}{\omega^2 - N^2} \frac{\partial p}{\partial z} \right) + \left( \frac{\omega^2 k^2}{\omega^2 - N^2} - \frac{\partial}{\partial z} \left( \frac{g k^2}{\omega^2 - N^2} \right) \right) p = 0. \quad (4)$$

Since  $\rho$  varies by several orders of magnitude in the atmosphere, it is often desirable to use the dependent variable  $\rho^{-1/2} p$  rather than  $p$ . It is easy to include this change of variable in the parabolic equation techniques presented in Secs. II and III.

The quantities  $p + i\omega^{-1} g \rho w$  and  $w$  are conserved across a horizontal interface between two layers.<sup>9</sup> Using Eq. (2) to eliminate  $w$ , we obtain the interface conditions,

$$\left[ \left( 1 - \frac{g^2 c^{-2}}{\omega^2 - N^2} \right) p + \frac{g}{\omega^2 - N^2} \frac{\partial p}{\partial z} \right] = 0, \quad (5)$$

$$\left[ \frac{\omega^2}{\omega^2 - N^2} \left( \frac{1}{\rho} \frac{\partial p}{\partial z} - \frac{g}{\rho c^2} p \right) \right] = 0, \quad (6)$$

where  $[\cdot]$  denotes the jump of the indicated quantity across the interface. The quantity in Eq. (5) vanishes at a free boundary. The quantity in Eq. (6) vanishes at a rigid boundary. Limiting cases of Eq. (4) include the acoustic wave equation,

$$\frac{\partial^2 p}{\partial x^2} + \rho \frac{\partial}{\partial z} \left( \frac{1}{\rho} \frac{\partial p}{\partial z} \right) + k^2 p = 0, \quad (7)$$

which is obtained by neglecting gravity, and the gravity wave equation,

$$\frac{\partial^2 p}{\partial x^2} + \rho \frac{\partial}{\partial z} \left( \frac{1}{\rho} \frac{\omega^2}{\omega^2 - N^2} \frac{\partial p}{\partial z} \right) = 0, \quad (8)$$

which is obtained by assuming that  $c$  is much larger than a representative wave speed (this is equivalent to assuming that the medium is incompressible). Various parabolic equation techniques have been developed for solving Eq. (7). In Secs. II and III, we develop parabolic equation techniques for solving Eq. (4), including Eq. (8) as a special case.

Some of the terms in Eq. (4) are singular at depths at which  $\omega = N$ . Although these singularities are regular, they can cause difficulties in numerical solutions. As we show in Sec. II, it is easy to avoid this problem in the parabolic equation solution. Another difficulty that can arise in solving Eq. (4) is that this equation has an infinite number of propagating modes for some problems. To illustrate this, we consider an arbitrary mode  $\phi_j(z)$  and eigenvalue  $k_j^2$  of Eq. (8) for a problem involving constant  $N$  and rigid boundaries at  $z=0$  and  $z=H$ . The depth separated eigenvalue problem is

$$\frac{\rho \omega^2}{\omega^2 - N^2} \frac{d}{dz} \left( \frac{1}{\rho} \frac{d\phi_j}{dz} \right) = k_j^2 \phi_j, \quad (9)$$

where  $\phi_j'(0) = \phi_j'(H) = 0$ . Multiplying Eq. (9) by  $\rho^{-1} \phi_j$ , integrating over  $z$ , and performing integration by parts, we obtain

$$k_j^2 = \frac{\omega^2}{N^2 - \omega^2} \frac{\int_0^H \rho^{-1} |\phi_j'|^2 dz}{\int_0^H \rho^{-1} |\phi_j|^2 dz}. \quad (10)$$

Since propagating modes correspond to  $k_j^2 > 0$ , we conclude that all of the modes are propagating for  $\omega < N$  and that all of the modes are evanescent for  $\omega > N$ . It is not possible to handle an infinite number of propagating modes with parabolic equation techniques. Fortunately, this is not necessary because only a finite number of the modes are physically meaningful. The modes that correspond to large horizontal wave numbers are annihilated by viscosity, which is not included in Eq. (4). In Sec. II, we discuss an alternative to including viscosity in the model.

## II. THE PARABOLIC EQUATION AND SELF STARTER

In this section, we derive a parabolic equation for acousto-gravity waves and describe an approach for solving it. The parabolic equation is obtained by factoring the operator in Eq. (4) into a product of incoming and outgoing operators. This factorization is exact for range-independent problems. Range-dependent problems are handled by approximating the medium by a sequence of range-independent regions. The parabolic equation is applied to propagate the field through each region. Techniques for handling the vertical interfaces between regions are described in Sec. III.

The acousto-gravity wave equation is in the form,

$$\frac{\partial^2 p}{\partial x^2} + k_0^2 (1 + X) p = 0, \quad (11)$$

$$X \equiv k_0^{-2} \left( \rho \frac{\partial}{\partial z} \frac{1}{\rho} \frac{\omega^2}{\omega^2 - N^2} \frac{\partial}{\partial z} + \frac{\omega^2 k^2}{\omega^2 - N^2} - \frac{\partial}{\partial z} \left( \frac{g k^2}{\omega^2 - N^2} \right) - k_0^2 \right), \quad (12)$$

where  $k_0 = \omega/c_0$  is a representative horizontal wave number and  $c_0$  is a representative phase speed. Factoring the operator in Eq. (11) into a product of incoming and outgoing operators, we obtain

$$\left( \frac{\partial}{\partial x} + ik_0 (1 + X)^{1/2} \right) \left( \frac{\partial}{\partial x} - ik_0 (1 + X)^{1/2} \right) p = 0. \quad (13)$$

Assuming that outgoing energy dominates backscattered energy, we obtain the outgoing wave equation,

$$\frac{\partial p}{\partial x} = ik_0 (1 + X)^{1/2} p. \quad (14)$$

Solving Eq. (14) formally, we obtain

$$p(x + \Delta x, z) = \exp(ik_0 \Delta x (1 + X)^{1/2}) p(x, z). \quad (15)$$

Substituting an  $n$ -term rational approximation for the depth operator in Eq. (15), we obtain the split-step Padé solution,<sup>10</sup>

$$p(x + \Delta x, z) = \exp(ik_0 \Delta x) \prod_{m=1}^n \frac{1 + \alpha_{m,n} X}{1 + \beta_{m,n} X} p(x, z), \quad (16)$$

which provides higher-order accuracy in both  $X$  (the asymptotic parameter) and  $k_0 \Delta x$  (the numerical parameter).

The solution of Eq. (16) is obtained by solving the equations,

$$(1 + \beta_{m,n}X)p_m = (1 + \alpha_{m,n}X)p_{m-1}, \quad (17)$$

where  $p_0 = \exp(ik_0\Delta x)p(x,z)$  and  $p_n = p(x + \Delta x, z)$ . Simple poles in the operator  $X$  can be avoided using the factor,

$$\zeta \equiv \left( \frac{\omega^2 - N^2}{\omega^2} \right)^2. \quad (18)$$

Multiplying both sides of Eq. (17) by  $\zeta$ , we obtain the equation,

$$(\zeta + \beta_{m,n}\zeta X)p_m = (\zeta + \alpha_{m,n}\zeta X)p_{m-1}, \quad (19)$$

which is free of singularities. Several approaches have been developed<sup>11-13</sup> for obtaining the coefficients  $\alpha_{m,n}$  and  $\beta_{m,n}$ , which can be defined by specifying two types of constraints. Accuracy constraints are imposed to guarantee that the propagating modes are handled properly. Stability constraints are imposed to annihilate the evanescent portion of the spectrum. The stability constraints can be used to account for the effects of viscosity by annihilating solutions of Eq. (4) that correspond to large horizontal wave numbers. As discussed in Sec. I, these modes are artifacts of neglecting the effects of viscosity. Evanescent waves can also be annihilated by using rotated Padé approximations,<sup>14,15</sup> which are obtained by first generating a rational approximation using only accuracy constraints and then rotating the branch cut below the negative real axis.

Initial conditions for the parabolic equation can be generated using the self starter.<sup>13,16,17</sup> Placing a line source term on the right-hand side of Eq. (11), we obtain

$$\frac{\partial^2 p}{\partial x^2} + k_0^2(1 + X)p = 2i\delta(x)\delta(z - z_0). \quad (20)$$

Integrating Eq. (20) over an arbitrarily small interval centered on  $x=0$  and using the continuity of  $p$ , we obtain the jump condition,

$$\left[ \frac{\partial p}{\partial x} \right] = 2i\delta(z - z_0). \quad (21)$$

Since the pressure field is symmetric about  $x=0$ , we obtain

$$\lim_{x \rightarrow 0^+} \frac{\partial p}{\partial x} = i\delta(z - z_0). \quad (22)$$

Using Eq. (14) to eliminate the  $x$  derivative in Eq. (22), we obtain

$$\lim_{x \rightarrow 0^+} k_0(1 + X)^{1/2}p = \delta(z - z_0). \quad (23)$$

To avoid the singularity at the source location, we apply Eqs. (15) and (23) to obtain

$$p(x_0, z) = k_0^{-1}(1 + X)^{-1/2} \exp(ik_0x_0(1 + X)^{1/2}) \times \delta(z - z_0), \quad (24)$$

where  $x_0$  is on the order of a wavelength.

Although the solution of Eq. (24) is bounded, we rearrange this equation to avoid encountering singular intermediate results and obtain

$$p(x_0, z) = k_0^{-1}(1 + X)^{-1/2}(1 + iX)^2 \times \exp(ik_0x_0(1 + X)^{1/2})\xi(z), \quad (25)$$

$$(1 + iX)^2\xi(z) = \delta(z - z_0). \quad (26)$$

Since the eigenvalues of the self-adjoint operator  $X$  are real, the intermediate solution  $\xi(z)$  is bounded. The initial condition is implemented by substituting a rational approximation for the function of  $X$  in Eq. (25). For some problems, this function contains a singularity at  $X = -1$ , which is an artifact of neglecting viscosity. This difficulty can be avoided by using stability constraints in the rational approximation. The generalization of Eq. (25) to the case of a point source in cylindrical coordinates is given by

$$p(r_0, z) = (k_0r_0)^{-1/2}(1 + X)^{-1/4}(1 + iX)^2 \times \exp(ik_0r_0(1 + X)^{1/2})\xi(z). \quad (27)$$

### III. RANGE DEPENDENCE AND VERTICAL INTERFACES

A range-dependent waveguide can be approximated by a sequence of range-independent regions. The parabolic equation techniques described in Sec. II can be used to construct an initial condition and propagate the field through each of the range-independent regions. In this section, we describe energy-conservation<sup>5-7</sup> and single-scattering<sup>8</sup> techniques for approximating transmitted fields across the vertical interfaces between regions. These approaches have proven to be effective for acoustics problems in which outgoing energy dominates backscattered energy. Single-scattering techniques are not as efficient as energy-conservation techniques but are applicable to a wider class of problems.

We consider an arbitrary range-independent region that begins at  $x=0$  and use the modal representation,

$$p(x, z) = \sum_j a_j \phi_j(z) \exp(ik_jx), \quad (28)$$

$$k_0^2(1 + X)\phi_j = k_j^2\phi_j, \quad (29)$$

$$\int_0^H \rho^{-1}\phi_i\phi_j dz = \delta_{ij}, \quad (30)$$

$$a_j = \int_0^H \rho^{-1}\phi_j p(0, z) dz, \quad (31)$$

where  $\phi_j$  and  $k_j^2$  are the  $j$ th mode and eigenvalue. The energy flux across a vertical interface at the end of the range-independent region is

$$E = \text{Im} \int_0^H \rho^{-1} p^* \frac{\partial p}{\partial x} dz. \quad (32)$$

Substituting Eq. (14) into Eq. (32), we obtain

$$E = k_0 \text{Re} \int_0^H \rho^{-1} p^*(1 + X)^{1/2} p dz. \quad (33)$$

We assume that  $p$  is composed of propagating modes and apply the modal representation and the orthogonality of the modes to obtain

$$E = k_0 \int_0^H |\rho^{-1/2}(1+X)^{1/4}p|^2 dz. \quad (34)$$

Since the integrand in Eq. (34) is a perfect square, we obtain the energy-conservation interface condition,

$$[\rho^{-1/2}(1+X)^{1/4}p] = 0. \quad (35)$$

The energy-conserving solution should be accurate for a wide class of problems. This approach is not appropriate for problems involving sloping rigid boundaries, which can cause significant backscattering even when the slope is small. The single-scattering approximation is applicable to such problems and should provide accurate outgoing solutions in many cases. In a range-dependent waveguide that is approximated by a sequence of range-independent regions, a sloping interface becomes a sequence of stair steps. We consider a vertical interface between regions in which there is a rigid boundary at  $z=H_i$  on the incident side and  $z=H_t$  on the transmitted side. We define the incident, transmitted, and reflected fields  $p_i$ ,  $p_t$ , and  $p_r$ . The total field is  $p_i+p_r$  on the incident side and  $p_t$  on the transmitted side. After obtaining  $p_t$  and  $p_r$ , we discard  $p_r$  and propagate  $p_t$  through the next range-independent region with the parabolic equation.

We first consider the upslope case  $H_t < H_i$ . The interface conditions are  $p_i+p_r=p_t$  for  $z < H_t$  and

$$\rho_i^{-1} \frac{\partial}{\partial x} (p_i+p_r) = \begin{cases} \rho_t^{-1} \partial p_t / \partial x & 0 < z < H_t \\ 0 & H_t < z < H_i \end{cases}. \quad (36)$$

Applying Eq. (14) to eliminate the range derivatives in Eq. (36), we obtain

$$L_i(p_i-p_r) = L_t p_t, \quad (37)$$

$$L_i \equiv \rho_i^{-1}(1+X_i)^{1/2}, \quad (38)$$

$$L_t p_t \equiv \begin{cases} \rho_t^{-1}(1+X_t)^{1/2} p_t & 0 < z < H_t \\ 0 & H_t < z < H_i \end{cases}. \quad (39)$$

The negative sign in Eq. (37) corresponds to the fact that  $p_r$  is incoming. Substituting the pressure conservation condition into Eq. (37) and rearranging, we obtain the iteration formula,

$$p_r = \frac{\tau-2}{\tau} p_r + \frac{1}{\tau} (1-L_i^{-1}L_t)(p_i+p_r), \quad (40)$$

where the convergence factor  $\tau \geq 2$  can be set to a relatively large value in order to improve convergence when  $p_r$  is relatively strong. The evaluation of the right side of Eq. (40) involves the application of  $L_t$  to  $p_i+p_r$  restricted to  $0 < z < H_t$ .

To simplify the downslope case  $H_t > H_i$ , we artificially extend  $p_i$  and  $p_r$  to  $H_i < z < H_t$  and let  $p_i+p_r=p_t$  and  $p_i=0$  in this interval. The interface conditions are  $p_i+p_r=p_t$  for  $0 < z < H_i$  and

$$\rho_t^{-1} \frac{\partial p_t}{\partial x} = \begin{cases} \rho_i^{-1} (\partial p_i / \partial x + \partial p_r / \partial x) & 0 < z < H_i \\ 0 & H_i < z < H_t \end{cases}. \quad (41)$$

Applying Eq. (14) to eliminate the range derivatives in Eq. (41), we obtain

$$M_i p_t = M_i (p_i - p_r), \quad (42)$$

$$M_i \equiv \rho_i^{-1}(1+X_i)^{1/2}, \quad (43)$$

$$M_i (p_i - p_r) \equiv \begin{cases} \rho_i^{-1}(1+X_i)^{1/2}(p_i - p_r) & 0 < z < H_i \\ 0 & H_i < z < H_t \end{cases}. \quad (44)$$

Substituting the pressure conservation condition into Eq. (42) and rearranging, we obtain the iteration formula,

$$p_r = \frac{\tau-2}{\tau} p_r + \frac{1}{\tau} (1-M_i^{-1}M_i)(p_r - p_i). \quad (45)$$

The evaluation of the right side of Eq. (45) involves the application of  $M_i$  to  $p_r - p_i$  restricted to  $0 < z < H_i$ . The artificial values of  $p_i$  and  $p_r$  are required in the evaluation of the other terms on the right side of Eq. (45).

#### IV. EXAMPLES

We present parabolic equation solutions to gravity and acousto-gravity wave problems in this section. Examples A, B, and C are idealized problems designed to test the parabolic equation techniques and illustrate range-dependent propagation effects. Examples D, E, and F involve realistic oceanic and atmospheric waveguides. Each of the examples involves a rigid lower boundary and plane geometry. For examples A, B, C, and D, which involve gravity waves, the upper boundary is modeled as a free surface. For examples E and F, which involve acousto-gravity waves, the upper atmosphere is modeled as an absorbing layer by allowing  $\omega$  to be complex.<sup>18</sup> For each of the examples, we let  $g=9.81 \text{ m/s}^2$  and define the transmission loss to be  $-10 \log_{10} |\rho^{-1} p^2|$ . We approximate operators with the rotated Padé approximations of Refs. 14 and 15 where  $\theta$  is the rotation angle of the branch cut. We handle vertical interfaces using the single-scattering approach described in Sec. III, with five iterations and  $\tau=2$ .

Example A involves a 0.25-Hz source at  $z=8 \text{ m}$  in a waveguide consisting of two homogeneous layers that are 10-m thick. The density is  $1 \text{ g/cm}^3$  in the upper layer and  $8 \text{ g/cm}^3$  in the lower layer. The bottom boundary is at  $z=20 \text{ m}$  for  $x < 400 \text{ m}$ ,  $z=8 \text{ m}$  for  $400 \text{ m} < x < 800 \text{ m}$ , and  $z=20 \text{ m}$  for  $x > 800 \text{ m}$ . This waveguide supports surface and interface waves in the deep regions and a surface wave in the shallow region. All of the other modes are evanescent. We take  $c_0=5 \text{ m/s}$ ,  $n=6$ , and  $\theta=10^\circ$ . The outgoing and incoming fields appear in Fig. 1. Backscattering is strong from the first vertical interface and weak from the second vertical interface. The contribution of the interface wave is relatively strong in the incoming field. The contribution of the surface wave is relatively strong for  $x > 800 \text{ m}$ . Evidence of evanescent modes appears near the source and interfaces.

Example B is similar to example A, with the exception that the bottom boundary slopes from  $z=20 \text{ m}$  to  $z=8 \text{ m}$  over the interval  $200 \text{ m} < x < 1000 \text{ m}$ . We solved this problem by approximating the sloping bottom with a sequence of stair steps, applying the single-scattering approximation at

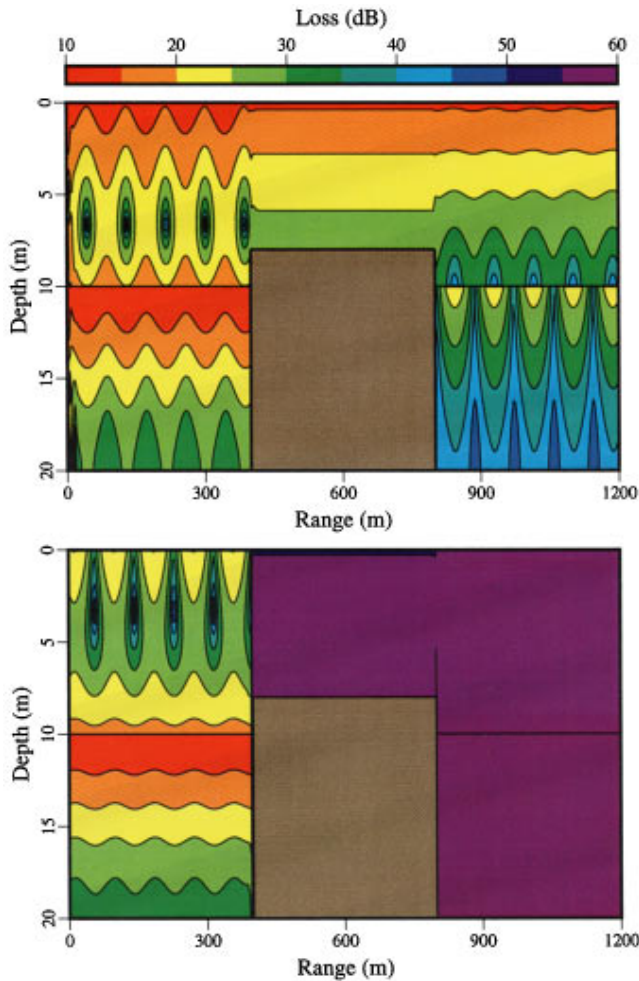


FIG. 1. Transmission loss for example A, which involves a 0.25-Hz source at  $z=8$  m. The waveguide contains two 10-m-thick layers of different density and vertical interfaces at  $x=400$  m and  $x=800$  m. In the deep regions, the outgoing field (top) contains an interference pattern between surface and interface waves. The backscattered field (bottom) is strong to the left of the first vertical interface. The small distortions of the two-mode interference patterns near the source and the vertical interfaces are due to contributions of evanescent modes.

each of the vertical interfaces, discarding the reflected fields, and using the transmitted fields to construct an outgoing solution. The outgoing component of the single-scattering solution is known to be accurate for some types of continuous range dependence (e.g., sloping interfaces in acoustics problems). Although the accuracy of the single-scattering approximation has not been confirmed for problems involving sloping rigid boundaries, the modal interference pattern in the outgoing solution in Fig. 2 appears to behave properly as depth decreases.

Example C involves 50-mHz gravity waves in a waveguide in which  $\rho=1$  g/cm<sup>3</sup> at  $z=0$  and  $N/2\pi\equiv 60$  mHz. The bottom boundary is at  $z=50$  m for  $x<100$  m,  $z=25$  m for  $x>200$  m, and sloping for  $100\text{ m}<x<200$  m. This waveguide supports an infinite number of propagating modes, including a mode that is independent of depth. In the Boussinesq approximation,<sup>19</sup> energy propagates along beams that make a constant angle with the horizontal. These features have been observed experimentally.<sup>20</sup> The purpose of this example is to confirm that the parabolic equation exhibits

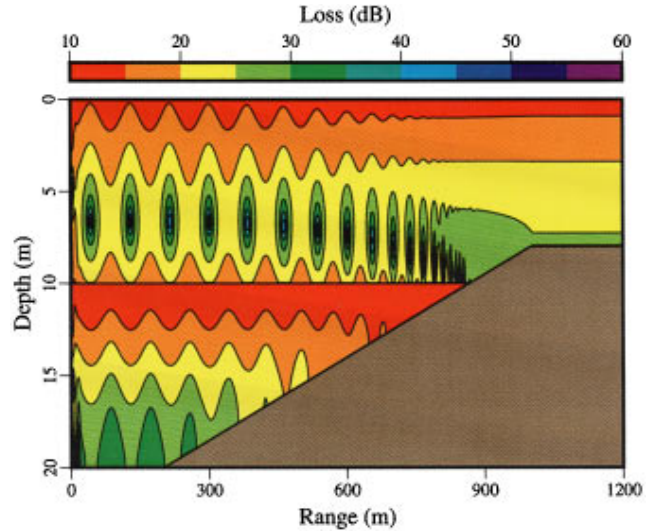


FIG. 2. Transmission loss for example B, which is similar to example A but has a sloping bottom boundary. The period of the interference pattern between the surface and interface waves decreases as the bathymetry decreases.

this behavior and to illustrate what happens when a beam reflects from a sloping boundary. To limit the number of modes that are excited, we use the initial condition,<sup>21,22</sup>

$$\rho^{-1/2}p(0,z) = \begin{cases} \sin(2\pi z/\lambda) + \frac{1}{2}\sin(4\pi z/\lambda) & |z-z_0| < \frac{1}{2}\lambda \\ 0 & |z-z_0| > \frac{1}{2}\lambda \end{cases}, \quad (46)$$

where  $\lambda=5$  m and  $z_0=25$  m. We take  $c_0=0.25$  m/s,  $n=12$ , and  $\theta=5^\circ$ . The outgoing solution appears in Fig. 3. The propagation angle agrees with the Boussinesq estimate and does not change when the beams reflect from the sloping boundary.

Example D involves 0.4-mHz gravity waves in an ocean with the  $N(z)$  profile appearing in Fig. 4, which was obtained from the North Pacific Ocean.<sup>23</sup> The bottom boundary

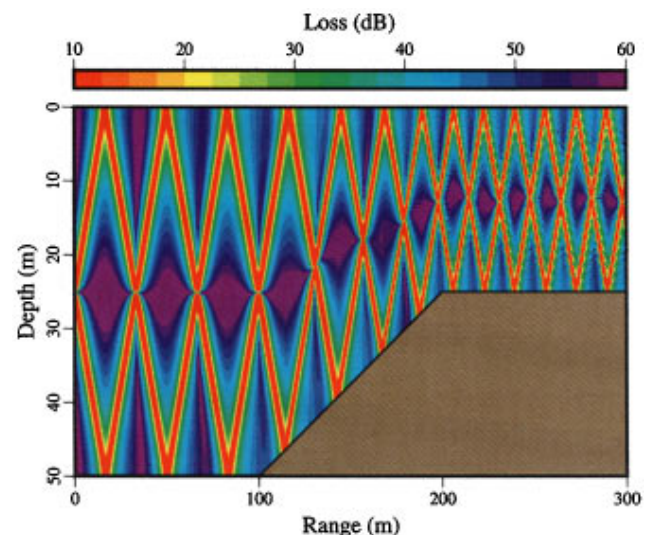


FIG. 3. Transmission loss for example C, which involves a 50-mHz wave in a medium with constant buoyancy frequency. The initial condition is defined in Eq. (46). In accordance with the Boussinesq approximation, the beams propagate at constant horizontal angles, even after reflecting from the sloping bottom.

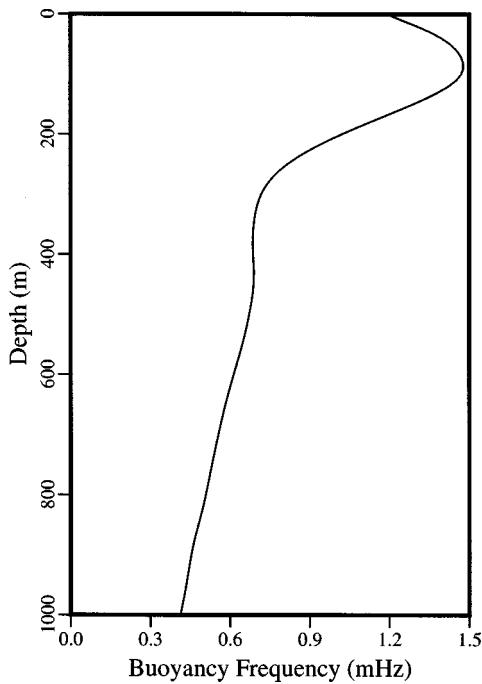


FIG. 4. The buoyancy frequency for example D, which was obtained in the North Pacific Ocean (Ref. 23). The displayed quantity is  $N/2\pi$ .

is at  $z = 1$  km for  $x < 2$  km,  $z = 500$  m for  $x > 8$  km, and sloping for  $2 \text{ km} < x < 8$  km. We obtained an initial condition for this problem by selecting a random combination of the modes according to the Garrett–Munk spectrum.<sup>24</sup> We selected the mode coefficients  $a_j$  in Eq. (28) from a Gaussian distribution with zero mean and the variance,<sup>25,26</sup>

$$\sigma_j = (j_*^2 + j^2)^{-1}, \quad (47)$$

where  $j_* = 3$  is an empirical parameter. We take  $c_0 = 0.25$  m/s,  $n = 6$ , and  $\theta = 10^\circ$ . The parabolic equation solution appearing in Fig. 5 corresponds to a particular realiza-

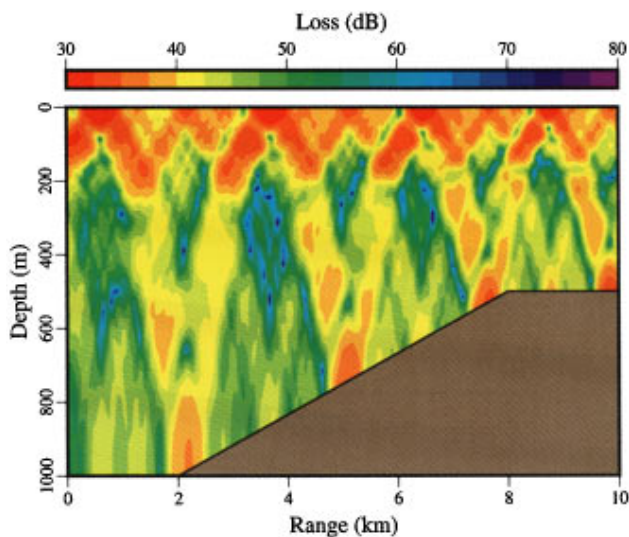


FIG. 5. Transmission loss for example D, which involves 0.4-mHz waves corresponding to a realization of the Garrett–Munk spectrum in an ocean with the buoyancy frequency appearing in Fig. 4. In contrast to the solution appearing in Fig. 3, energy propagates along curved paths because the buoyancy frequency varies with depth.

TABLE I. The coefficients of the rational functions that define the model atmosphere in Eqs. (48) and (49). The first four pairs are for the density. The second four pairs are for the temperature. These coefficients were obtained by performing least-squares fits to standard profiles (Ref. 27).

$i$	$A_i$	$B_i$
1	$-3.908\ 201\ 7 \times 10^{-2}$	$-4.924\ 463\ 7 \times 10^{-3}$
2	$-1.152\ 646\ 5 \times 10^{-3}$	$-1.298\ 414\ 2 \times 10^{-6}$
3	$3.289\ 193\ 7 \times 10^{-5}$	$-1.570\ 159\ 5 \times 10^{-6}$
4	$-2.049\ 495\ 8 \times 10^{-7}$	$1.553\ 597\ 4 \times 10^{-8}$
5	$-4.708\ 729\ 5 \times 10^{-2}$	$-2.722\ 176\ 9 \times 10^{-2}$
6	$1.250\ 638\ 7 \times 10^{-3}$	$4.247\ 473\ 3 \times 10^{-4}$
7	$-1.519\ 449\ 8 \times 10^{-5}$	$-3.958\ 318\ 1 \times 10^{-6}$
8	$6.581\ 887\ 7 \times 10^{-8}$	$1.729\ 579\ 5 \times 10^{-8}$

tion of the initial field. In contrast to example C, energy propagates along curved paths because  $N$  is depth dependent. There is a relatively large amount of energy in the upper water column because the coefficients of the lower modes (which decay with depth) are relatively large.

Examples E and F involve a model atmosphere spanning  $0 \leq z \leq 200$  km that we constructed by performing least-squares fits to standard density and temperature profiles<sup>27</sup> to obtain the rational approximations,

$$\log_{10} \left( \frac{\rho(z)}{\rho(0)} \right) \cong \frac{A_1 z + A_2 z^2 + A_3 z^3 + A_4 z^4}{1 + B_1 z + B_2 z^2 + B_3 z^3 + B_4 z^4}, \quad (48)$$

$$\frac{T(z)}{T(0)} \cong \frac{1 + A_5 z + A_6 z^2 + A_7 z^3 + A_8 z^4}{1 + B_5 z + B_6 z^2 + B_7 z^3 + B_8 z^4}, \quad (49)$$

where  $T(z)$  is the temperature,  $\rho(0) = 1.225$  g/cm<sup>3</sup>,  $T(0) = 288.2$  K, and the coefficients are given in Table I. We obtained a model buoyancy frequency profile using Eq. (3). We obtained a model sound-speed profile using the formula,<sup>28</sup>

$$c^2 = \gamma RT, \quad (50)$$

where  $R = 287.05$  J·K<sup>-1</sup>·kg<sup>-1</sup> is the ratio of the gas constant and the mean molecular weight of air and  $\gamma = 1.4$  is the ratio of the specific heats of air at constant pressure and at constant volume. The model profiles appear in Fig. 6. The model density profile agrees closely with the standard profile. The model sound-speed profile agrees qualitatively with the standard profile, which is sufficient for our purposes. To approximate the effects of attenuation in the upper atmosphere and prevent artificial reflections from the computational boundary at  $z = 200$  km, we replace  $\omega$  with the complex frequency,<sup>18</sup>

$$\tilde{\omega} \equiv \omega + i\mu k_0^2 \rho^{-1}, \quad (51)$$

where  $\mu = 1.5 \times 10^{-5}$  N·s·m<sup>-2</sup> is the viscosity. Since  $\rho$  decreases exponentially with altitude, the attenuation is negligible in the lower atmosphere but large in the upper atmosphere.

Example E involves a 3-mHz source at  $z = 10$  km. We take  $c_0 = 160$  m/s,  $n = 10$ , and  $\theta = 10^\circ$ . Parabolic equation solutions appear in Fig. 7. There is a large difference between the gravity and acousto-gravity wave solutions. To illustrate

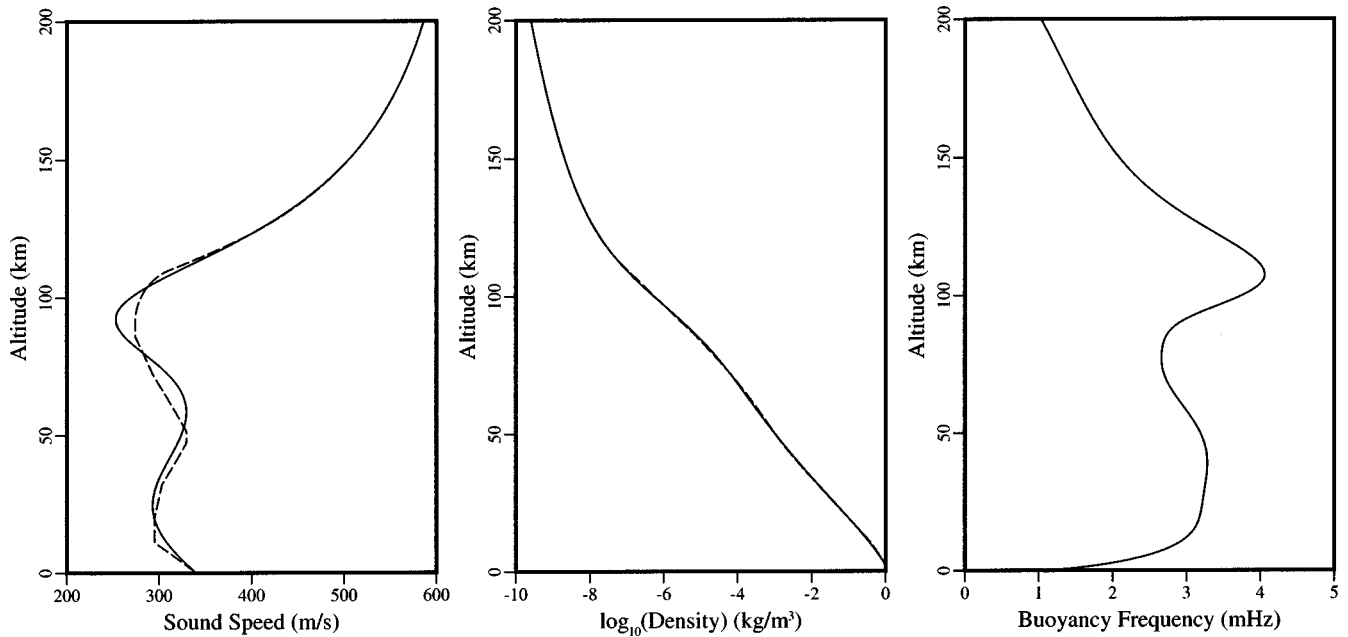


FIG. 6. The model atmospheric profiles (solid curves) for examples E and F, which were obtained by fitting four-term rational approximations to standard profiles (Ref. 27) (dashed curves). The model sound-speed profile agrees qualitatively with the standard profile, which contains cusps in the lower atmosphere. The model density profile is in good agreement with the standard profile. The bulges in the buoyancy frequency profile can duct gravity waves. The displayed quantity is  $N/2\pi$ .

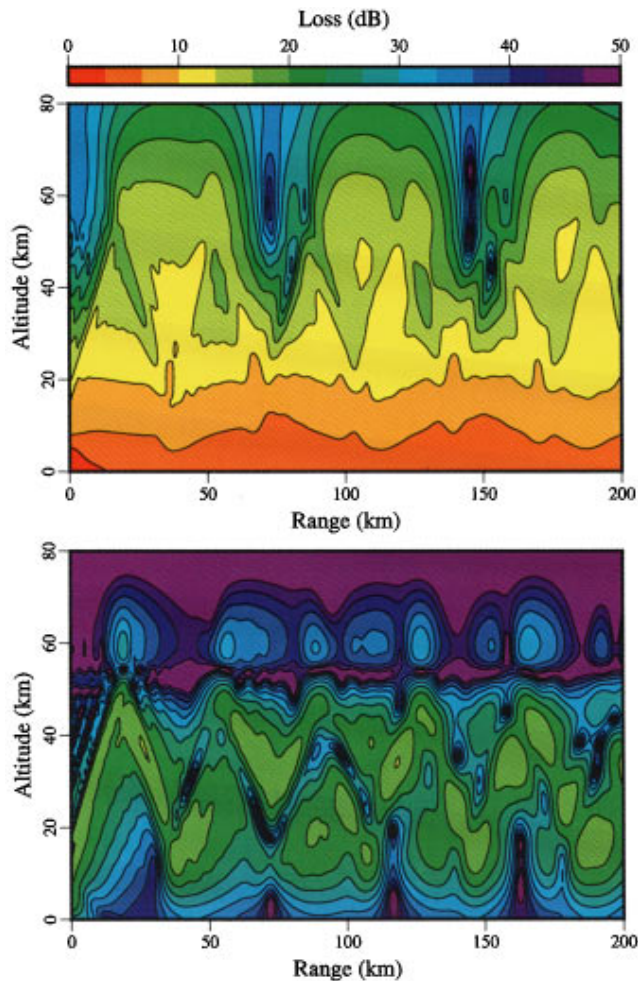


FIG. 7. Transmission loss for example E, which involves a 3-mHz source at  $z = 10$  km. There is a large difference between the acousto-gravity (top) and gravity (bottom) wave solutions. The acousto-gravity wave field contains a large amount of energy near the ground due to the contribution of the Lamb wave. The gravity wave field is weak outside the lower bulge in the buoyancy frequency profile, which appears in Fig. 6.

the nature of these differences, we consider the wave number spectrum.<sup>29</sup> Taking the Fourier transform of Eq. (20) in  $x$ , we obtain

$$(k_0^2(1+X) - h^2)\hat{p} = 2i\delta(z-z_0), \quad (52)$$

where  $h$  is the horizontal wave number and  $\hat{p}(h, z)$  is the spectrum. The acousto-gravity and gravity wave spectra appear in Fig. 8. The peaks in the spectra line up fairly well for large  $h$ . There are large differences for small  $h$ , including an additional peak in the acousto-gravity wave spectrum that corresponds to the Lamb wave.<sup>30</sup> Example F is similar to example E but involves variable topography and an initial condition consisting of the Lamb wave. The maximum altitude of the topographical features is 3 km. The parabolic

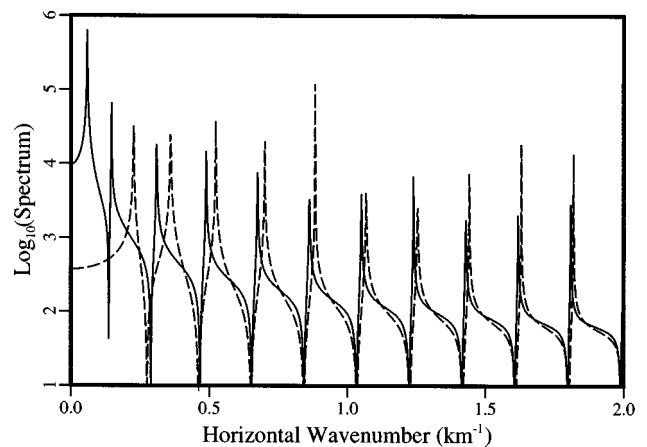


FIG. 8. Wave number spectra at  $z = 10$  km for example E. The peaks in the acousto-gravity wave spectrum (solid curve) are nearly aligned with the peaks in the gravity wave spectrum (dashed curve) for relatively large horizontal wave numbers. There are large differences in the spectra for small horizontal wave numbers, including an extra peak in the acousto-gravity wave spectrum that corresponds to the Lamb wave.

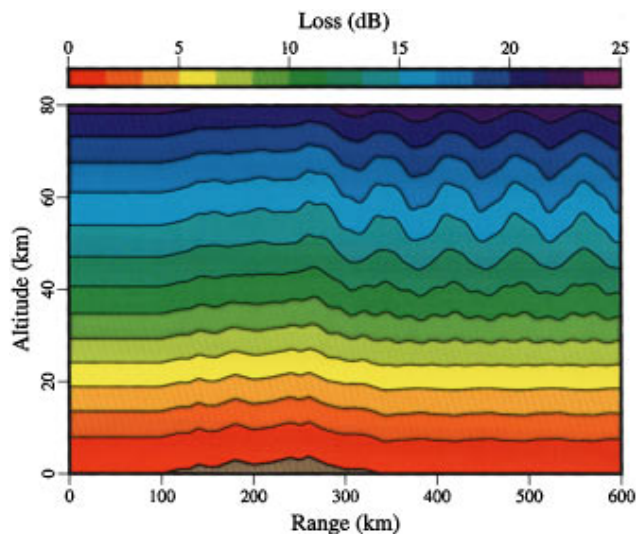


FIG. 9. Transmission loss for example F, which involves a 3-mHz Lamb wave and variable topography. To the left of the variable topography, the field consists of a single mode that decays exponentially with altitude. The variable topography causes mode coupling.

equation solution appears in Fig. 9. The variable topography causes mode coupling.

## V. CONCLUSION

Parabolic equation techniques have been developed for gravity and acousto-gravity wave propagation in range-dependent media. The self-starter can be used to generate initial conditions. The split-step Padé algorithm can be used to propagate the field through range-independent regions. Energy-conservation and single-scattering techniques can be used to handle vertical interfaces between range-independent regions. The accuracy of the single-scattering approximation needs to be tested for problems involving sloping boundaries. With these techniques, it is easy to avoid singularities at which  $N = \omega$ . The rational approximations can be designed to annihilate modes that correspond to very small wavelengths. The techniques have been tested for idealized problems and applied to realistic problems involving range-dependent oceanic and atmospheric waveguides. Possible extensions of this work include accounting for effects associated with ambient flow, viscosity, out-of-plane propagation, nonlinearity, and the earth's rotation and curvature.

## ACKNOWLEDGMENTS

This work was supported by the Office of Naval Research and a National Research Council-NRL Research Associateship. The authors thank John A. Colosi for providing the ocean data and useful suggestions.

<sup>1</sup>F. B. Jensen, W. A. Kuperman, M. B. Porter, and H. Schmidt, *Computational Ocean Acoustics* (American Institute of Physics, New York, 1994), pp. 343–412.

- <sup>2</sup>A. E. Gill, *Atmosphere-Ocean Dynamics* (Academic, New York, 1982), pp. 169–175.
- <sup>3</sup>P. L.-F. Liu and T.-K. Tsay, "On weak reflection of water waves," *J. Fluid Mech.* **131**, 59–71 (1983).
- <sup>4</sup>A. C. Radder, "On the parabolic equation method for water-wave propagation," *J. Fluid Mech.* **95**, 159–176 (1979).
- <sup>5</sup>M. B. Porter, F. B. Jensen, and C. M. Ferla, "The problem of energy conservation in one-way models," *J. Acoust. Soc. Am.* **89**, 1058–1067 (1991).
- <sup>6</sup>M. D. Collins and E. K. Westwood, "A higher-order energy-conserving parabolic equation for range-dependent ocean depth, sound speed and density," *J. Acoust. Soc. Am.* **89**, 1068–1075 (1991).
- <sup>7</sup>M. D. Collins, "An energy-conserving parabolic equation for elastic media," *J. Acoust. Soc. Am.* **94**, 975–982 (1993).
- <sup>8</sup>M. D. Collins and R. B. Evans, "A two-way parabolic equation for acoustic backscattering in the ocean," *J. Acoust. Soc. Am.* **91**, 1357–1368 (1992).
- <sup>9</sup>A. E. Gill, Ref. 2, pp. 85–87.
- <sup>10</sup>M. D. Collins, "A split-step Padé solution for the parabolic equation method," *J. Acoust. Soc. Am.* **93**, 1736–1742 (1993).
- <sup>11</sup>B. T. R. Wetton and G. H. Brooke, "One-way wave equations for seismoacoustic propagation in elastic waveguides," *J. Acoust. Soc. Am.* **87**, 624–632 (1990).
- <sup>12</sup>M. D. Collins, "Higher-order Padé approximations for accurate and stable elastic parabolic equations with application to interface wave propagation," *J. Acoust. Soc. Am.* **89**, 1050–1057 (1991).
- <sup>13</sup>R. J. Cederberg and M. D. Collins, "Application of an improved self-starter to geoacoustical inversion," *IEEE J. Ocean Eng.* **22**, 102–109 (1997).
- <sup>14</sup>F. A. Milinazzo, C. A. Zala, and G. H. Brooke, "Rational square-root approximations for parabolic equation algorithms," *J. Acoust. Soc. Am.* **101**, 760–766 (1997).
- <sup>15</sup>J. F. Lingeitch and M. D. Collins, "Wave propagation in range-dependent poro-acoustic waveguides," *J. Acoust. Soc. Am.* **104**, 783–790 (1998).
- <sup>16</sup>M. D. Collins, "A self-starter for the parabolic equation method," *J. Acoust. Soc. Am.* **92**, 2069–2074 (1992).
- <sup>17</sup>M. D. Collins, "The stabilized self starter," *J. Acoust. Soc. Am.* (submitted for publication).
- <sup>18</sup>A. E. Gill, Ref. 2, pp. 294–297.
- <sup>19</sup>A. E. Gill, Ref. 2, p. 130.
- <sup>20</sup>J. Lighthill, *Waves in Fluids* (Cambridge University Press, Cambridge, 1978), pp. 308–316.
- <sup>21</sup>K. Fuchs and G. Mueller, "Computation of synthetic seismograms with the reflectivity method and comparison with observations," *Geophys. J. R. Astron. Soc.* **23**, 417–433 (1971).
- <sup>22</sup>H. Schmidt, "SAFARI User's Guide," SAFLANTCEN SR-113 (SAFLANT Undersea Research Centre, La Spezia, Italy, 1988), p. 109.
- <sup>23</sup>J. A. Colosi, S. M. Flatté, and C. Bracher, "Internal-wave effects on 1000-km oceanic acoustic pulse propagation: Simulation and comparison with experiment," *J. Acoust. Soc. Am.* **96**, 452–468 (1994).
- <sup>24</sup>C. Garrett and W. H. Munk, "Space-time scales of internal waves: A progress report," *J. Geophys. Res.* **80**, 291–297 (1975).
- <sup>25</sup>J. A. Colosi and M. Brown, "Efficient numerical simulation of stochastic internal-wave-induced sound-speed perturbation fields," *J. Acoust. Soc. Am.* **103**, 2232–2235 (1998).
- <sup>26</sup>*Internal Waves and Small Scale Processes*, edited by C. Wunsch and B. Warren (MIT, Cambridge, 1981), pp. 264–291.
- <sup>27</sup>*U.S. Standard Atmosphere* (U.S. Government Printing Office, Washington, DC, 1976).
- <sup>28</sup>A. D. Pierce, *Acoustics: An Introduction to its Physical Principles and Applications* (Acoustical Society of America, New York, 1989), p. 28.
- <sup>29</sup>H. Schmidt and F. B. Jensen, "A full wave solution for propagation in multilayered viscoelastic media with application to Gaussian beam reflection at fluid-solid interfaces," *J. Acoust. Soc. Am.* **77**, 813–825 (1985).
- <sup>30</sup>A. E. Gill, Ref. 2, pp. 173–175.

# Acoustics of fluid-filled boreholes with pipe: Guided propagation and radiation

V. N. Rama Rao<sup>a)</sup> and J. K. Vandiver

*Department of Ocean Engineering, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139*

(Received 14 July 1997; accepted for publication 9 March 1999)

The study of the guided propagation and radiation from radially layered boreholes has applications in monitoring exploratory drilling. Axisymmetric propagation through pipes immersed inside fluid-filled bores in infinite elastic spaces was modeled using full 3-D elastodynamic equations. Three propagating modes are present for frequencies below 1 kHz. These are the lowest order modes and are characterized through their dispersion, attenuation, modeshapes, and particle motions. Mode I is confined mainly to the cross section of the pipe, while modes II and III are present in the fluid layers. Modes I and II are weakly dispersive, while mode III is strongly dispersive. Mode III is the equivalent of the Stoneley mode present in fluid-filled bores without pipe and is most influenced by the properties of the surrounding elastic space. Mode III has a cutoff frequency, below which it radiates, when the shear speed of the surrounding elastic space is sufficiently low. © 1999 Acoustical Society of America. [S0001-4966(99)03606-1]

PACS numbers: 43.20.Mv, 43.20.Bi, 43.20.Hq [DEC]

## INTRODUCTION

Monitoring the dynamics of drillpipes in fluid-filled boreholes and the drilling process is important in improving the efficiency of exploratory drilling. Many drilling monitoring systems make vibration and acoustic measurements in boreholes at the surface and attempt to infer downhole signals and drilling conditions. Hence, understanding wave propagation in boreholes with a drillpipe is critical for effectively processing and interpreting these data and for designing optimal data acquisition parameters.

This study develops a model for axisymmetric wave propagation in a fluid-filled borehole with a pipe, as shown in Fig. 1. In this model, there are four cylindrically symmetric layers with alternating fluid and elastic layers. The innermost layer is fluid and the outermost is an infinite elastic solid. The bore in the infinite elastic solid represents the borehole in the earth and the cylinder represents the drillpipe and the fluid inside the pipe and in the annulus between the pipe and the bore is drilling fluid. The earth will be referred as the "formation" while the two fluid layers will be referred as the "inner fluid" and "outer fluid." Formations are commonly referred to as being slow or soft when the formation shear speed is less than the compressional speed in the drilling fluid and fast or hard when it is not. The entire length of pipe in the borehole or the drillstring is typically composed of 30-ft pipe sections called drillpipes and drillcollars. These come in a variety of sizes (i.d. and o.d. combinations), with drillcollars having relatively larger wall thickness. Further, in both cases, each 30-ft pipe is homogeneous along its length except at the ends where there are tool joints, or short sections with expanded cross section for connectivity purposes. This study models a drillpipe and drillcollar as infinitely long pipes without tool joints as it captures the essen-

tial wave propagation features. The discontinuities due to the different sizes of drillpipes and drillcollars in a borehole would introduce scattering and reverberation to transient signals propagating up the borehole.<sup>1</sup> The periodic occurrence of tooljoints would additionally impose pass band and stop band structure in the frequency domain response.<sup>1,2</sup>

Early studies<sup>3-9</sup> focused on propagation in empty and fluid-filled boreholes in infinite elastic solids without pipe. It was established that at low frequencies, when wavelengths are much larger than the borehole radius, the Stoneley mode or Tube mode is the only propagating mode in a simple fluid-filled borehole. At low frequencies, this mode has plane wavefronts in the fluid while at high frequencies it is a surface mode existing at the borehole-formation interface with amplitudes that decay away from the interface. The Stoneley mode is referred to as such because this limiting behavior at high frequencies is that of the Stoneley surface wave. Additionally, at high frequencies, when the wavelengths are of the order of borehole radius or lesser, there also exist pseudo-Rayleigh modes or shear normal modes. These exhibit oscillatory amplitudes inside the borehole and a decaying amplitude in the formation. These modes have a low-frequency cutoff below which they are highly attenuated. Related studies on propagation in fluid-filled elastic pipes submerged in air and fluids<sup>10-15</sup> have reported similar results.

Later studies examined the effects of formation logging tools and casings that introduce additional radial layering within the borehole. Cheng and Toksöz<sup>16</sup> modeled the presence of a logging tool in the borehole. Unlike a drillpipe, the modeled tool was solid, and did not include fluid inside it. Their results showed that the lowest order propagating mode in this system was still the Stoneley mode. Additionally, the presence of a rigid tool reduced the cross-sectional area of the borehole and hence increased the cutoff frequencies of the pseudo-Rayleigh modes. Tubman and others<sup>17</sup> investigated another form of radial layering, a poorly bonded casing, to examine its effect on formation logging. They used a

<sup>a)</sup>Presently at the Earth Resources Laboratory, Department of Earth, Atmospheric and Planetary Sciences, Massachusetts Institute of Technology, Cambridge, MA 02142.



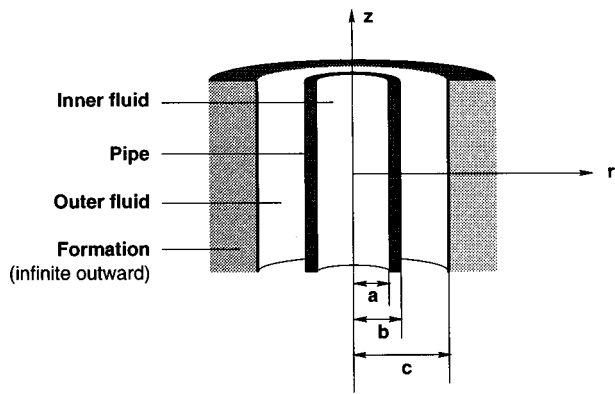


FIG. 1. Schematic of a fluid-filled borehole with pipe.

layer matrix technique to model an elastic cylinder in the borehole with fluid trapped between the cylinder and borehole wall and showed that the Stoneley mode and an additional fluid mode were present. A similar result was reported by Lee,<sup>1</sup> who developed a comprehensive model of a drill pipe in a fluid-filled borehole under a low-frequency approximation. He predicted a total of three modes, with two Stoneley-like modes and a mode localized in the pipe. This model was very accurate in hard formations and when the wavelengths were long compared to the borehole dimensions. Using Lee's infinite length model, Rao<sup>18</sup> used a transfer matrix approach to derive transfer functions and impulse responses of finite length boreholes with drillpipe.

While the low-frequency model was adequate for low frequencies (below 440 Hz in limestone) and hard formations, a more comprehensive model was necessary to examine dispersion and radiation in soft formations, and behavior at higher frequencies. In this study, the fluid-filled pipe in the fluid-filled borehole is modeled using the full 3-D elastodynamic equations. This allows for the treatment of thick cylinders, soft formations, and high frequencies. Propagating modes are identified and dispersion curves, modeshapes, attenuations, and particle motions are presented. In soft formations one of the modes exhibits dispersion and under certain conditions a cutoff behavior, characteristics that could not be either modeled or identified with the low-frequency model. Additionally, time domain pressure transients in the borehole due to a volume source are also given which exemplifies the multi-mode response of this layered cylindrical waveguide.

## I. FORMULATION

The schematic of the model is shown in Fig. 1. The fluids are assumed to be inviscid and the solids are assumed to be elastic and isotropic. The fluids and solids are assumed to be lossless. Axisymmetric solutions are considered. In a cylindrical coordinate system the displacement potentials in the elastic layers satisfy the wave equation,<sup>5</sup>

$$\frac{\partial^2 \Phi_j}{\partial r^2} + \frac{1}{r} \frac{\partial \Phi_j}{\partial r} + \frac{\partial^2 \Phi_j}{\partial z^2} = \frac{1}{\alpha_j^2} \frac{\partial^2 \Phi_j}{\partial t^2}, \quad (1)$$

$$\frac{\partial^2 \Psi_j}{\partial r^2} + \frac{1}{r} \frac{\partial \Psi_j}{\partial r} - \frac{\Psi_j}{r^2} + \frac{\partial^2 \Psi_j}{\partial z^2} = \frac{1}{\beta_j^2} \frac{\partial^2 \Psi_j}{\partial t^2}, \quad (2)$$

where  $\Phi_j$  and  $\Psi_j$  are the compressional and shear potentials in layer  $j$ , and  $\alpha_j$  and  $\beta_j$  are the compressional and shear speeds in that layer. Propagation in the fluid layers is governed by Eq. (1) alone. The subscripts  $j$  are 1 for the inner fluid, 2 for the pipe, 3 for the outer fluid, and 4 for the formation.

Axisymmetric wave propagation in the axial direction ( $\sim e^{-ikz + i\omega t}$ ) and the radiation condition in which energy only radiates outward give the following set of solutions:<sup>5,9</sup>

$$\begin{aligned} \bar{\Phi}_1(r, z, \omega) &= \int_{-\infty}^{\infty} A_1 I_0(lr) e^{-ikz} dk + \bar{S}, \\ \bar{\Phi}_2(r, z, \omega) &= \int_{-\infty}^{\infty} [A_2 I_0(m_\phi r) + B_2 K_0(m_\phi r)] e^{-ikz} dk, \\ \bar{\Psi}_2(r, z, \omega) &= \int_{-\infty}^{\infty} [C_2 I_1(m_\psi r) + D_2 K_1(m_\psi r)] e^{-ikz} dk, \\ \bar{\Phi}_3(r, z, \omega) &= \int_{-\infty}^{\infty} [A_3 I_0(nr) + B_3 K_0(nr)] e^{-ikz} dk, \\ \bar{\Phi}_4(r, z, \omega) &= \int_{-\infty}^{\infty} A_4 K_0(p_\phi r) e^{-ikz} dk, \\ \bar{\Psi}_4(r, z, \omega) &= \int_{-\infty}^{\infty} B_4 K_1(p_\psi r) e^{-ikz} dk, \end{aligned} \quad (3)$$

where  $I_\xi$ ,  $K_\xi$  are modified Bessel functions of order  $\xi$ , and  $k$  is the axial wave number.  $l$ ,  $m$ ,  $n$ , and  $p$  are radial wave numbers in the inner fluid, pipe, outer fluid, and the formation, respectively, and  $\phi$ ,  $\psi$  are subscripts of the compressional and shear components of the wave number, respectively:

$$\begin{aligned} l &= \left\{ k^2 - \frac{\omega^2}{\alpha_1^2} \right\}^{1/2}, & m_\phi &= \left\{ k^2 - \frac{\omega^2}{\alpha_2^2} \right\}^{1/2}, \\ m_\psi &= \left\{ k^2 - \frac{\omega^2}{\beta_2^2} \right\}^{1/2}, & n &= \left\{ k^2 - \frac{\omega^2}{\alpha_3^2} \right\}^{1/2}, \\ p_\phi &= \sqrt{k^2 - \frac{\omega^2}{\alpha_4^2}}, & \text{if } k &\geq \frac{\omega}{\alpha_4}, \\ &= i \sqrt{\frac{\omega^2}{\alpha_4^2} - k^2}, & \text{if } k &< \frac{\omega}{\alpha_4}, \\ p_\psi &= \sqrt{k^2 - \frac{\omega^2}{\beta_4^2}}, & \text{if } k &\geq \frac{\omega}{\beta_4}, \\ &= i \sqrt{\frac{\omega^2}{\beta_4^2} - k^2}, & \text{if } k &< \frac{\omega}{\beta_4}. \end{aligned} \quad (4)$$

The square root symbol signifies the root with positive real part. The square roots defining the radial wave numbers introduce branch cuts and the appropriate sign convention is given.

$A_1(k, \omega), A_2(k, \omega), B_2(k, \omega), \dots, B_4(k, \omega)$  are unknown coefficients and  $\bar{S}$  represents the potential of a source, which is assumed to be an impulsive point dilatational source in an infinite fluid medium,<sup>9</sup>

TABLE I. Borehole and formation specifications.

$j$	Layer	Compression speed $\alpha_j$ m/s	Shear speed $\beta_j$ m/s	Density $\rho_j$ kg/m <sup>3</sup>
1.	Inner fluid (water)	1558	...	1000
2.	Drillpipe (steel)			
	Drillcollar (steel)	5900	3400	7800
3.	Outer fluid (water)	1558	...	1000
4.	Formation			
	Hard (sandstone)	3279	1893	2000
	Soft (shale or clays)	1409	813	2000

$$\bar{S}(r, z, \omega) = \frac{-V_0}{4\pi^2} \int_{-\infty}^{\infty} K_0(lr) e^{-kz} dk, \quad (5)$$

where  $V_0$  is the volume displacement of the source. The inner fluid potential  $\bar{\Phi}_1$  is expressed as a sum of two potentials, one for the source in an infinite fluid medium and another which accounts for the effects of the boundary at  $r = a$ .

**Boundary conditions**

The three boundary conditions at each of the three layer interfaces specify the nine unknown coefficients which define the various potentials. The conditions are continuity of radial displacement, continuity of radial stress, and the absence of shear stress at the inner and outer radius of the pipe and the inner radius of the borehole. Once the displacements and stresses are expressed in terms of the potentials,<sup>5</sup> the three boundary conditions can be imposed at each of the three radial locations. The resulting equations in matrix form are as follows:

$$[\Xi(k, w)]\{Y(k, w)\} = 0, \quad (6)$$

where  $[\Xi(k, w)]$  is a  $9 \times 9$  matrix with 36 nonzero elements,  $\{Y(k, w)\}$  is the vector of unknown potential coefficients  $\{A_1, A_2, B_2, \dots, B_4\}^T$ , and is presented in Appendix A.

When a volume source is present on the axis of the borehole, its effects can be included at the boundaries to obtain

$$[\Xi(k, w)]\{Y(k, w)\} = \{\mathcal{F}(\bar{S}, k, w)\}, \quad (7)$$

where  $\{\mathcal{F}(\bar{S}, k, w)\}$  is a vector with terms involving the source (see Appendix A). This equation can be solved for the potential coefficients,  $\{Y(k, w)\}$ .

**II. DISPERSION**

The eigenvalues are obtained as roots of the characteristic equation

$$\text{Det}[\Xi(k, w)] = 0. \quad (8)$$

The locus of the roots  $k^*$  in the frequency wave number plane gives the dispersion characteristics of the various modes. This is obtained by solving for the roots of the above equation at different frequencies. Modes that have phase

TABLE II. Dimensions (borehole diameter=12 in.).

	i.d. in.	o.d. in.
Drillpipe	4.19	5.00
Drillcollar	3.00	8.00

speeds less than the slowest medium speed correspond to real wave number roots and those which have phase speeds that are greater correspond to complex roots. Muller's root finding algorithm<sup>19</sup> was used to locate the roots.

The modal phase speeds are examined in two ways—as a function of frequency for two different formations and as a function of formation speed for three discrete frequencies. This exemplifies their behavior in a variety of situations. The properties of the two formation types, a hard and soft formation, are given in Table I. The hard formation shear speed corresponds to sandstone and that of the soft formation to shales and clays. The compression speed is assumed to be  $\sqrt{3}$  times the shear speed (poisson's ratio=0.25) in both cases. Two common sizes of pipe, a 5-in. drillpipe and an 8-in. drillcollar, are considered. The borehole is 12 in. in diameter. Complete dimensions are provided in Table II. All calculations are done up to a frequency of 1000 Hz as this corresponds to the typical bandwidth of drilling vibration data. The model itself is formulated under no frequency limitations. Damping was absent in all layers.

**A. Frequency variation**

The dispersion characteristics of mode I for drillpipe and drillcollar in a soft and hard formation are plotted in Fig. 2. The phase speed in drillpipe is 5395 m/s in soft formation and 5397 m/s in hard formation, while that in drillcollar is 5381 m/s in soft formation and 5384 m/s in hard formation. Mode I is nondispersive up to 1 kHz and not sensitive to the properties of the surrounding fluid and formation. Mode I phase speed is close to the speed of axial waves through steel pipes in air ( $c_l = \sqrt{E/\rho} = 5378$  m/s).

Mode II is also nondispersive below 1 kHz and is not very sensitive to formation properties as shown in Fig. 3.

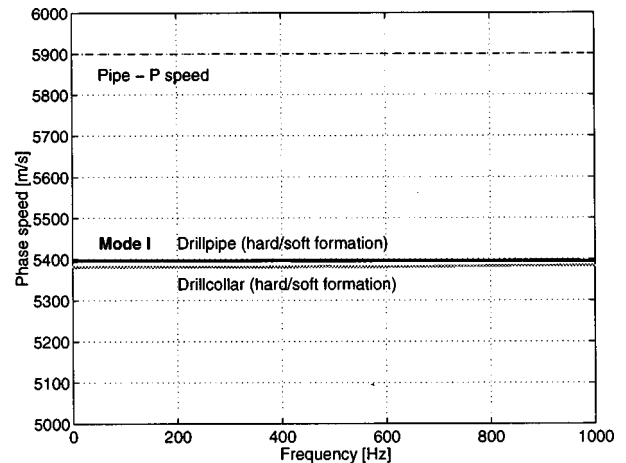


FIG. 2. Dispersion characteristics of mode I in drillpipe (black) and drillcollar (gray). The hard and soft formation curves coincide.

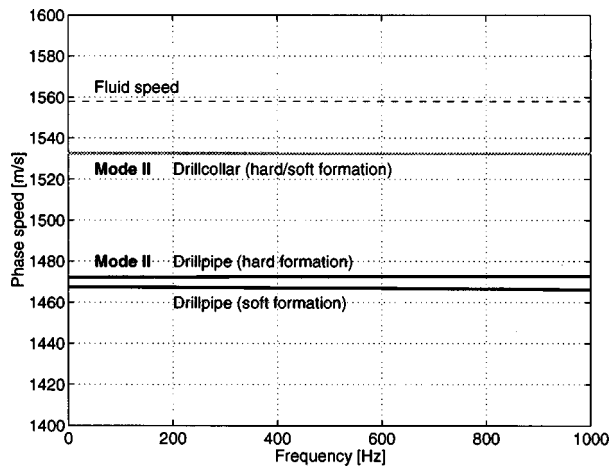


FIG. 3. Dispersion characteristics of mode II in drillpipe and drillcollar. The hard and soft formation curves for drillcollar coincide.

Mode II phase speed in drillpipe is 1466 m/s (soft formation) and 1472 m/s (hard formation) and in drillcollar is 1530 m/s (soft or hard formation). The phase speed of this mode is greater in drillcollar because it is a thicker and stiffer pipe. This mode has a phase speed close to that of plane waves in fluids confined in elastic pipes ( $c_f = 1471$  m/s in drillpipe and 1547 m/s in drillcollar,  $c_f$  defined in Appendix B). Thus this mode is controlled mainly by the properties of the pipe and very weakly by the fluid and formation layers outside the pipe.

Mode III is strongly influenced by the formation. It is the equivalent of the Stoneley mode found in fluid-filled boreholes without pipe. In the hard formation, Fig. 4, it is weakly dispersive below 1 kHz with phase speed increasing from 1296 to 1305 m/s for the drillpipe and 1220 to 1238 m/s for the drillcollar. In the soft formation, Fig. 5, it is strongly dispersive with the phase speed decreasing from 864 to 762 m/s for the drillpipe and 749 to 712 m/s for the drillcollar. In comparison, the low-frequency phase speed of the Stoneley mode in a borehole without drillpipe,  $c_S$ , is 925 m/s in soft formation and 1344 m/s in hard formation ( $c_S$  defined in Appendix B).

In drillpipe, mode III phase speed is greater than the

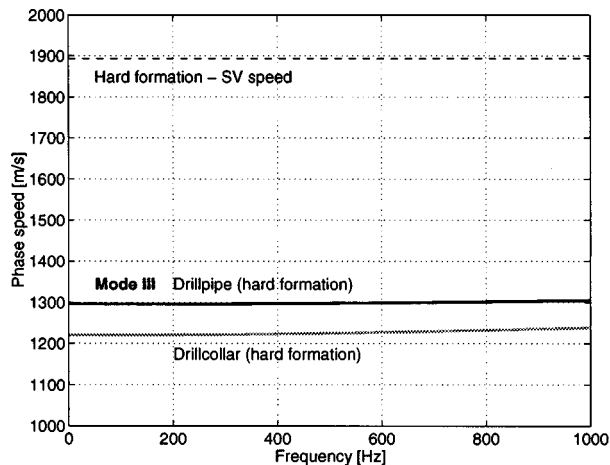


FIG. 4. Dispersion characteristics of mode III in drillpipe and drillcollar in a hard formation.

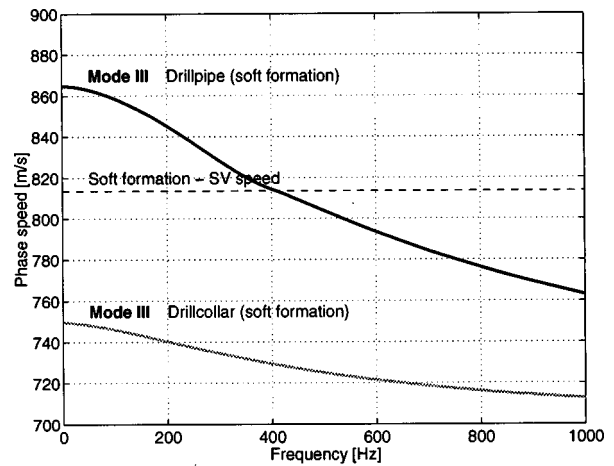


FIG. 5. Dispersion characteristics of mode III in drillpipe and drillcollar in a soft formation.

formation shear speed below 410 Hz with this frequency being termed the *cutoff frequency*. Below the cutoff frequency, this mode radiates energy into the formation, while above it energy in this mode is trapped within the borehole. It is worth noting that mode III in the drillcollar does not have a cutoff frequency for this formation, and hence does not radiate. It must also be noted that the term “cutoff frequency” used here does not refer to the frequency where the phase speed tends to infinity and group speed tends to zero. However, in both usages, frequency components below the “cutoff frequency” are attenuated.

## B. Formation variation

The migration of the three modal phase speeds was computed as the formation shear speed varied from 200 to 2000 m/s. The formation compression speed is assumed to be  $\sqrt{3}$  times the shear speed (Poissons ratio=0.25) and density is 2000 kg/m<sup>3</sup>.

For both drillpipe and drillcollar, change in the phase speeds of modes I and II with formation shear speed and frequency (0–1000 Hz) is less than 0.6% and is not shown. However, mode III phase speed variation is more significant and complex, as shown in Fig. 6, for the drillpipe. Three curves are plotted for 1, 100, and 1000 Hz. Curves for the

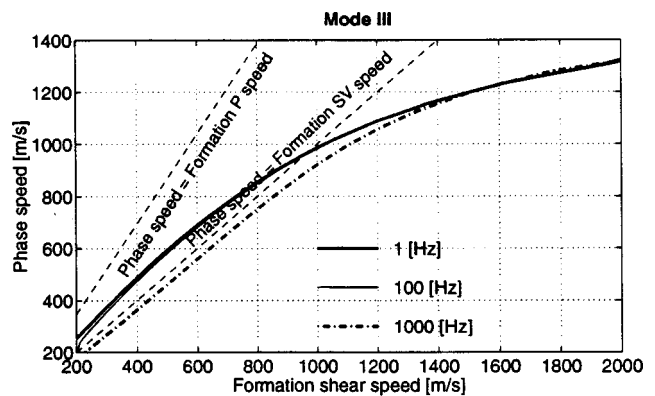


FIG. 6. In drillpipe: Variation of mode III phase speed with formation shear speed. Curves at 1 and 100 Hz coincide (formation compression speed = 1.732 × shear speed, density is constant at 2000 kg/m<sup>3</sup>).

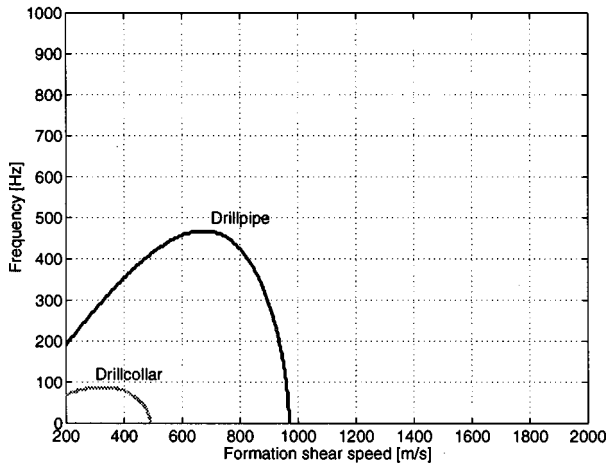


FIG. 7. Cutoff frequencies for drillpipe and drillcollar in different formations (formation compression speed=1.732×shear speed, density=2000 kg/m<sup>3</sup>).

drillcollar are similar, but shifted toward slower phase speeds and are not shown.

The 1-Hz and 100-Hz curves coincide for most formations. The 1-Hz and the 1-kHz curves crossover at a formation shear speed of 1600 m/s. This implies that mode III phase speed decreases with increasing frequency in formations below the crossover point and increases with increasing frequency above the crossover point. Additionally, mode III dispersion is minimal in formations close to the crossover point.

Further, at 1 Hz, mode III radiates into formations with shear speed lesser than 970 m/s. Thus a cutoff frequency exists for these formations, above which this mode does not radiate. This is exemplified in Fig. 7 where the cutoff frequencies for drillpipe and drillcollar in a range of formations is presented. The region enclosed by the curves are parameter ranges where this mode is leaky. The plot shows that cutoff frequencies exist only for formations with shear speeds slower than 970 m/s with drillpipe and 500 m/s with drillcollar. This is the case for the present choice of material and geometric parameters.

### III. ATTENUATION

The real part of the axial wave number gives the dispersion characteristics while the imaginary part provides the attenuation characteristics of the modes. The attenuation for the  $i$ th mode,  $\Gamma_i$ , is computed in dB/Km as,

$$\Gamma_i(k_i^*, \omega) = 20 \log_{10}(e^{\Im(k_i^*)1000}), \quad (9)$$

where  $\Im(k_i^*)$  is the imaginary component of the axial wave number of the  $i$ th mode ( $i=I,II,III$ ). The attenuations for the modes in drillpipe and drillcollar are shown in Fig. 8. No damping was included in the model. Hence all axial attenuation is due to radial leakage of energy into the formation. In a real borehole, damping in the various media and scattering due to inhomogeneities along the borehole would introduce additional attenuation. Modes that are subsonic relative to

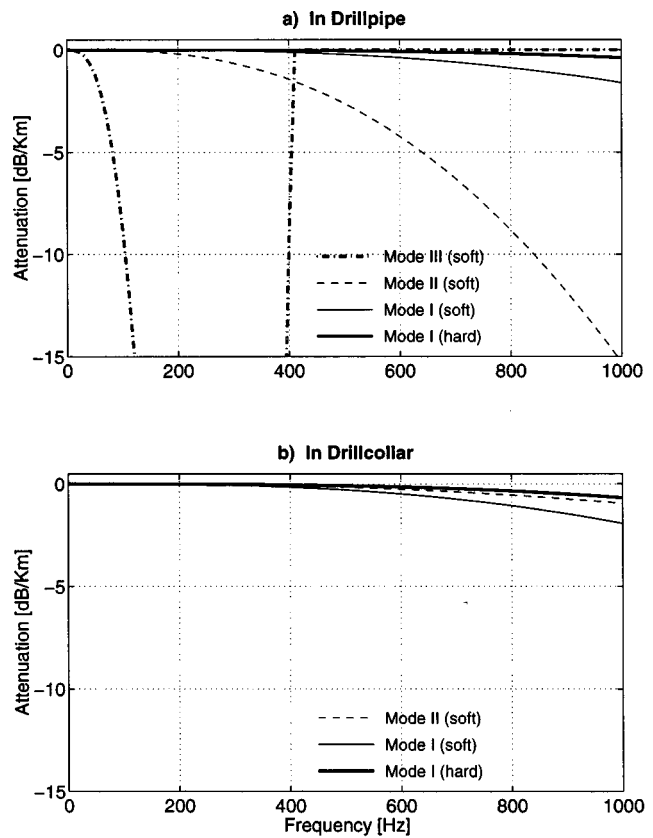


FIG. 8. Attenuation characteristics of the three modes in (a) drillpipe and (b) drillcollar. Mode III in drillpipe attains a maximum attenuation of  $-92$  dB/Km at 290 Hz and mode III in drillcollar is unattenuated and hence not shown.

the formation shear speed are unattenuated are not shown in the plots.

In both the drillpipe and drillcollar, and in hard and soft formations, mode I is leaky over the entire frequency range, but with small attenuation levels because it couples poorly to the formation, i.e., it is associated with small formation displacements and stresses. The attenuation increases with frequency and is greater in the soft formation than in the hard formation. Mode II in drillpipe has considerable attenuation in soft formations at higher frequencies. Downhole signal components propagating to the surface as mode II would be distorted by this frequency dependent attenuation. It is unattenuated in the hard formation. Mode II attenuation in drillcollar is much lesser due to the fact that the drillcollar, being a thicker pipe than drillpipe, causes a poorer coupling to the formation.

Mode III in drillcollar is unattenuated in hard and soft formations as it slower than the formation shear speed in both cases. In drillpipe it is unattenuated in the hard formation. In the soft formation, it has nonzero attenuation only below the cutoff frequency of 410 Hz. Radiation into the formation increases as the frequency increases from zero while it decreases to zero as frequency converges to cutoff frequency, causing a peak in the attenuation of  $-93$  dB/Km at 290 Hz. As the frequency decreases below the cutoff frequency, the radiation into the formation increases because the phase speed gets progressively greater than the formation shear speed. However, as the frequency tends toward zero,

the shear radiation into the formation drops to zero despite the phase speed being greater than the formation speed. This is because, in the limit of zero frequency or very large wavelengths, the formation deformations and stresses become purely compressional and hence shear radiation is not possible. However, mode III cannot radiate in compression as it is slower than the formation compression speed. Thus these two mechanisms cause a peak in attenuation. The attenuation characteristics of mode III are very important in understanding the propagation of signals in the outer fluid layer.

#### IV. MODESHAPES

Once the roots  $k^*$  of the characteristic equation are known, the coefficient vector,  $\{Y(k^*, \omega)\}$ , for each mode is computed from Eq. (6), correct to a multiplicative constant. Then stresses and displacements corresponding to each mode are then computed from the potentials in the various layers to give the modeshapes. For the sake of brevity, modeshapes in drillpipe alone are presented here. Modeshapes in drillcollar have similar qualitative behavior. The plots show the variation of the displacements and stresses along the radius, from the axis of the borehole to twice borehole radius. The phase difference between the axial and radial displacements controls the direction of the particle orbits and that between displacements and stresses controls energy flux or power flow per unit area.

##### Hard and soft formation

For the hard formation, axial displacement and stress modeshapes are shown in Fig. 9, while the radial displacement and stress modeshapes are shown in Fig. 10. For the soft formation, the axial displacement and stress modeshapes are shown in Fig. 11 and the radial displacement and stress modeshapes being qualitatively similar to those of the hard formation are not shown.

Mode I is characterized by axial displacements and stresses that are uniform over the cross section of the pipe. The pipe is the conduit for axial power transmission of this mode. The radial displacements of this mode expand the inner fluid and pipe cross sections and compress the outer fluid cross section. This results in pressures of opposing signs in the two fluid layers. Continuity of radial stress across interfaces leads to a node in the radial stress in pipe. The modeshape in both formations are similar.

Mode II is characterized by uniform axial displacements and stresses mainly in the inner fluid layer, leading to axial power flow along the inner fluid layer for this mode. The radial stresses show that the pressure in the inner fluid is mostly balanced by the radial stress in the pipe and to a much smaller extent by the outer fluid and formation. This accounts for the insensitivity of this mode to change in formation properties. The pressures in the inner and outer fluid layers are of the same sign for this mode. The modeshapes are real in the hard formation and complex in the soft formation as this mode radiates in the latter.

Mode III has axial displacements and stresses in the outer fluid layer and in the inner fluid layer. The axial power flows in both fluid layers in the hard formation while it is mainly in the outer fluid in the soft formation. A pressure

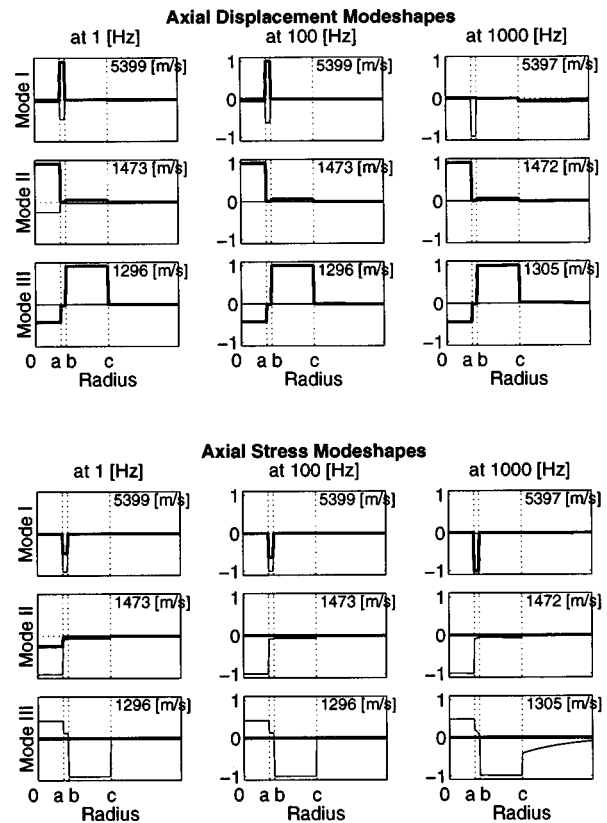


FIG. 9. In drillpipe, hard formation: Axial displacement and stress modeshapes at 1, 100, and 1000 Hz. Real part of modeshapes—thick line; imaginary part—thin line.

perturbation in the outer fluid deforms both the pipe and formation in proportion to their stiffnesses, as the two stiffnesses are in series. When their stiffnesses are comparable, or when the formation is hard, the pipe deforms and causes a pressure change in the inner fluid. However, when the formation is much weaker than the pipe, as in a soft formation, the outer fluid pressure results in little pipe deformation and hence marginal inner fluid pressure. A similar behavior can be observed in mode II also, but to a lesser extent. The outer and inner fluid pressures are of opposite signs for this mode.

#### V. PARTICLE MOTION

The axial and radial displacement modeshapes were normalized individually and hence their relative magnitudes were not obvious in the modeshapes. The particle orbits at a given frequency, normalized by the largest displacement, axial or radial, captures the dominant motions associated with each mode.

For most conditions, axial displacements are purely real and radial displacements are purely imaginary, as can be seen in the modeshapes. Hence, the particle orbits are elliptical. The direction of the orbit depends on the sign of the phase difference between the axial and radial displacements, with positive values corresponding to anti-clockwise orbits and negative values to clockwise orbits. The particle orbits in drillpipe, for the three modes in hard and soft formation, are shown in Figs. 12 and 13, respectively. They are calculated at the axis of the borehole, the center of the pipe, and the outer fluid layer, and in the formation at one and one half

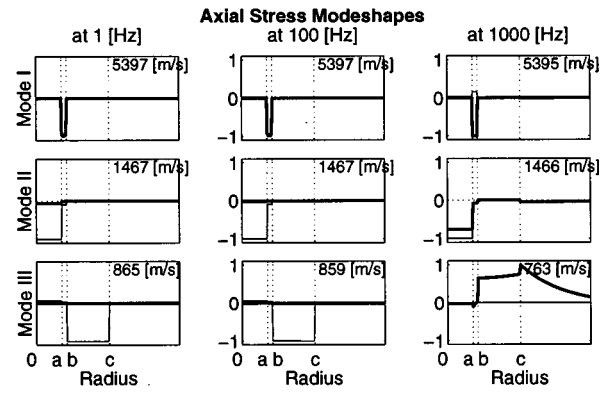
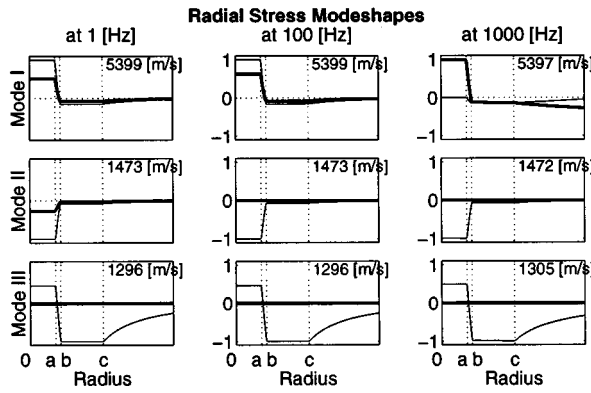
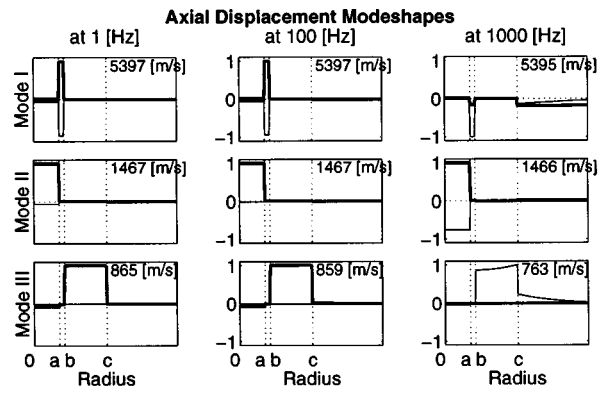
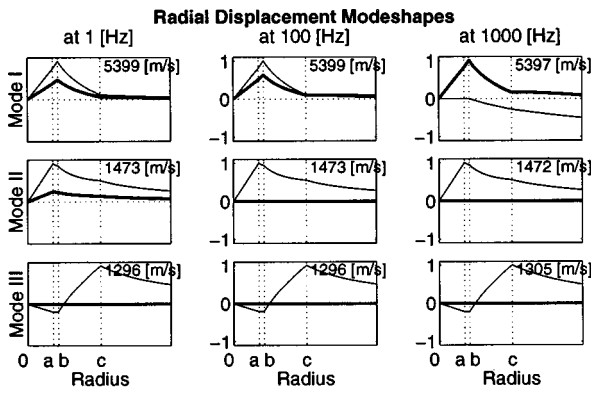


FIG. 10. In drillpipe, hard formation: Radial displacement and stress modeshapes at 1, 100, and 1000 Hz. Real part of modeshapes—thick line; imaginary part—thin line.

FIG. 11. In drillpipe, soft formation: Axial displacement and stress modeshapes at 1, 100, and 1000 Hz. Real part of modeshapes—thick line; imaginary part—thin line.

borehole radius. In both formations, modes I and II have dominantly axial orbits upto 1 kHz. Mode III acquires a radial component at higher frequencies and this effect is more pronounced in soft formations than in hard formations. Additionally the relative sizes of the outer fluid and formation orbits become comparable in the soft formation at high frequencies.

## VI. TRANSIENT RESPONSES

Once the potential coefficients are determined using Eq. (5), the stresses in the borehole layers can be computed from the potentials.<sup>5</sup> These stresses are in the axial wave-number frequency domain and can be transformed into axial space and time domain:

$$f(r, z, t) = \int_{-\infty}^{\infty} d\omega e^{i\omega t} G(\omega) \int_{-\infty}^{\infty} dk e^{-ikz} F(r, k, \omega), \quad (10)$$

where  $F(r, k, \omega)$  denotes the physical quantity to be transformed and  $G(\omega)$  is the source spectrum, assumed to be Gaussian:

$$G(\omega) = e^{-\omega^2/500}. \quad (11)$$

Thus the source time function is a Gaussian pulse, about 0.1 s wide. The two integrals were computed using fast Fourier transforms as discussed in Refs. 20 and 21. Time responses of pressure in the inner fluid (at  $r=a/2$ ), axial stress in the pipe [at  $r=(a+b)/2$ ], and pressure in the outer fluid [at  $r$

$= (b+c)/2$ ] were computed at an axial location 1000 m from a volume source of amplitude  $V_0 = 10 \text{ cm}^3$ . The computations were done for the drillpipe in the hard and soft formation.

Mode I (Pipe mode)				
	$r=0$	$r=(b+a)/2$	$r=(c+b)/2$	$r=3c/2$
at 1 [Hz]	(-)	(+)	(+)	(+)
at 1000 [Hz]	(-)	(+)	(+)	(+)
Mode II (Inner fluid mode)				
at 1 [Hz]	(+)	(+)	(+)	(+)
at 1000 [Hz]	(+)	(+)	(+)	(+)
Mode III (Outer fluid mode)				
at 1 [Hz]	(+)	(+)	(+)	(+)
at 1000 [Hz]	(+)	(+)	(+)	(+)

FIG. 12. In drillpipe, hard formation: Particle orbits of the three modes in the inner fluid, pipe, outer fluid and formation. Anti-clockwise orbit (+); clockwise orbit (-).

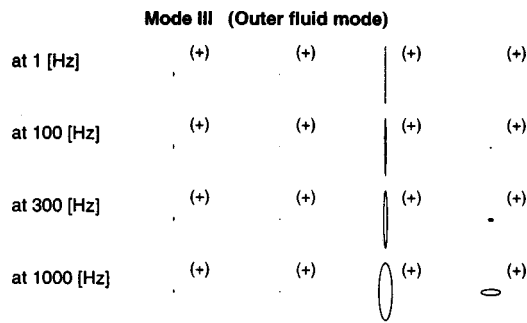


FIG. 13. In drillpipe, soft formation: Particle orbits of mode III in the inner fluid, pipe, outer fluid and formation. Orbits of modes I and II are similar to those in the hard formation and are not shown. Anti-clockwise orbit (+); clockwise orbit (-).

The results for the hard and soft formation are in Figs. 14 and 15, respectively. In response to a source pulse, one propagating pulse in the inner fluid layer and two in the pipe and outer fluid layer can be observed. The responses in the pipe and inner fluid layer for the soft formation were identical to that in the hard formation and are not shown.

The response pulses are nondispersive in the current frequency range (0–64 Hz), as expected from the dispersion curves seen earlier. Mode I is felt only in the pipe in both formations, while mode II is present in all three layers (inner fluid, pipe, and outer fluid). Mode II is strongest in the inner fluid where it is directly excited by the source. Further, mode II pressure is of the same phase in all three layers, as expected from modeshape results (refer to Subsec. IV). The outer fluid response in the hard formation is an order of magnitude larger than that in the soft formation. This is due to the strong coupling of inner and outer fluid pressures in the hard formation.

Receivers in the pipe and outer fluid layer would measure two separate pulses in response to a single source pulse. This is in a borehole with no axial discontinuities, like changes in pipe or borehole radii or formation properties. Typical boreholes have many axial discontinuities which would additionally introduce reverberation. Further, higher-

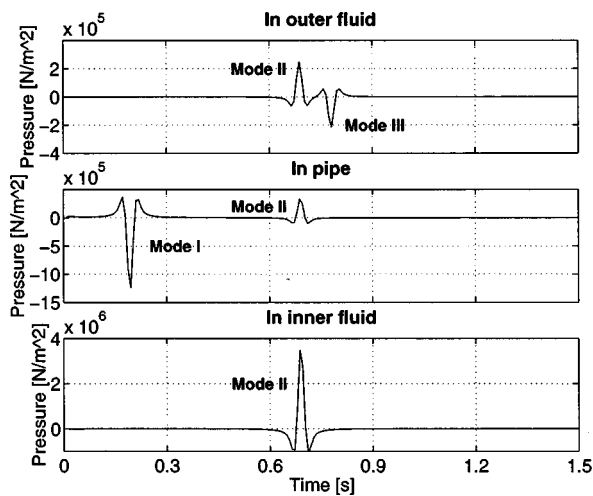


FIG. 14. Hard formation—Pressure response in the three layers in a borehole with drillpipe, 1000 m away from a volume source that is on the axis of the borehole.

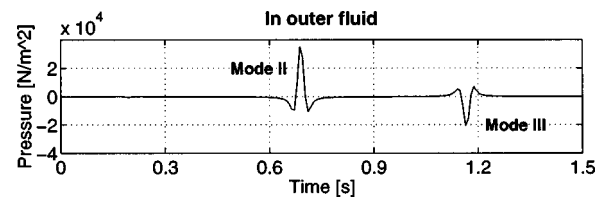


FIG. 15. Soft formation—Pressure response in the outer fluid layer in a borehole with drillpipe, 1000 m away from a volume source that is on the axis of the borehole (response in pipe and inner fluid layer is the same as that for the hard formation).

frequency data in soft formations would additionally be complicated by modal dispersion. Thus interpreting surface measurements made while drilling requires a clear understanding of the propagation characteristics of the borehole modes. The next step toward improved data interpretation would involve incorporating axial discontinuities in a model of the type discussed here.

## VII. SUMMARY

A model for axisymmetric wave propagation in fluid-filled pipes in boreholes was developed. Dispersion characteristics, attenuations, modeshapes, and particle motions as a function of frequency and formation properties were computed and analyzed for two common pipe sizes, a 5-in. drillpipe and an 8-in. drillcollar.

Three propagating modes exist for frequencies under 1000 Hz. Mode I, characterized by dominant axial motion in the cross section of the pipe, is nondispersive and insensitive to a wide range of formations. It propagates at a phase speed close to longitudinal waves in pipe. While it is supersonic compared to the formation, it radiates very weakly because of poor coupling to the formation.

Mode II, characterized by dominant axial motion in the inner fluid layer, is also nondispersive and insensitive to a wide range of formations. This mode radiates into soft formations. The frequency dependent axial attenuation associated with this radiation will distort signal components that propagate up as mode II. This has implications for drilling data communication systems that use the inner fluid layer for transmission and reception.

Mode III is dispersive with characteristics that depend on pipe, formation, and frequency. Mode III dispersion is stronger in drillpipe than in drillcollar. Under some conditions it has a low-frequency cutoff below which it radiates into the formation. The parameter ranges (frequency and formation properties) over which this mode is leaky are larger for drillpipe than drillcollar. Thus mode III radiation while drilling is most likely to occur from the regions of the borehole where the drillpipe is present. Interpretation of outer fluid data recorded at the surface requires a knowledge of the dispersion and attenuation characteristics of this mode all along the borehole.

The modal characteristics predicted by this model are useful for understanding borehole acoustics and can serve as a benchmark for approximate models.

## ACKNOWLEDGMENTS

We wish to acknowledge the support of the members of the Drilling Engineers Association—Project 29 Consortium. The members are AGIP, Amoco, Baker Hughes Inteq, BP, British Gas, Elf Aquitaine, Mobil, Statoil, Total, and U.S. Bureau of Mines. We also acknowledge the valuable comments of the reviewers.

## APPENDIX A: BOUNDARY CONDITIONS

The terms of the matrix  $\Xi$ , in Eq. (7) are:

$$\begin{aligned}\Xi_{11} &= I_1(la), \\ \Xi_{12} &= \frac{k^2 - m_\psi^2}{k^2 + m_\psi^2} m_\phi I_1(m_\phi a), \\ \Xi_{13} &= -\frac{k^2 - m_\psi^2}{k^2 + m_\psi^2} m_\phi K_1(m_\phi a), \\ \Xi_{21} &= \frac{\lambda_1}{\mu_2} (l^2 - k^2) I_0(la) + \frac{2l}{a} I_1(la), \\ \Xi_{22} &= -(k^2 + m_\psi^2) I_0(m_\phi a), \\ \Xi_{23} &= -(k^2 + m_\psi^2) K_0(m_\phi a), \\ \Xi_{24} &= -2\iota k m_\psi I_0(m_\psi a), \quad \Xi_{25} = -2\iota k m_\psi K_0(m_\psi a), \\ \Xi_{32} &= -2\iota k m_\phi I_1(m_\phi a), \quad \Xi_{33} = 2\iota k m_\phi K_1(m_\phi a), \\ \Xi_{34} &= (k^2 + m_\psi^2) I_1(m_\psi a), \quad \Xi_{35} = (k^2 + m_\psi^2) K_1(m_\psi a), \\ \Xi_{44} &= \frac{m_\psi^2 - k^2}{2\iota k} I_1(m_\psi b), \quad \Xi_{45} = \frac{m_\psi^2 - k^2}{2\iota k} K_1(m_\psi b), \\ \Xi_{46} &= -n I_1(nb), \quad \Xi_{47} = n K_1(nb), \\ \Xi_{52} &= -(k^2 + m_\psi^2) I_0(m_\phi b), \\ \Xi_{53} &= -(k^2 + m_\psi^2) K_0(m_\phi b), \\ \Xi_{54} &= 2\iota k m_\psi I_0(m_\psi b), \quad \Xi_{55} = -2\iota k m_\psi K_0(m_\psi b), \\ \Xi_{56} &= \frac{\lambda_3}{\mu_2} (k^2 - n^2) I_0(nb) - \frac{2n}{b} I_1(nb), \\ \Xi_{57} &= \frac{\lambda_3}{\mu_2} (k^2 - n^2) K_0(nb) + \frac{2n}{b} K_1(nb), \\ \Xi_{62} &= -2\iota k m_\phi I_1(m_\phi b), \quad \Xi_{63} = 2\iota k m_\phi K_1(m_\phi b), \\ \Xi_{64} &= (k^2 + m_\psi^2) I_1(m_\psi b), \quad \Xi_{65} = (k^2 + m_\psi^2) K_1(m_\psi b), \\ \Xi_{76} &= n I_1(nc), \quad \Xi_{77} = -n K_1(nc), \\ \Xi_{78} &= p_\phi K_1(p_\phi c), \quad \Xi_{79} = -\iota k K_1(p_\psi c), \\ \Xi_{86} &= \frac{\lambda_3}{\mu_4} (n^2 - k^2) I_0(nc) + \frac{2n}{c} I_1(nc), \\ \Xi_{87} &= \frac{\lambda_3}{\mu_4} (n^2 - k^2) K_0(nc) - \frac{2n}{c} K_1(nc), \\ \Xi_{88} &= -(k^2 + p_\psi^2) K_0(p_\phi c), \quad \Xi_{89} = 2\iota k p_\psi K_0(p_\psi c),\end{aligned}$$

$$\Xi_{98} = 2\iota k p_\phi K_1(p_\phi c), \quad \Xi_{99} = (k^2 + p_\psi^2) K_1(p_\psi c);$$

and those of  $\{\mathcal{F}(\bar{S}, k, w)\}$  are:

$$\begin{aligned}\mathcal{F}_1 &= \frac{-V_0}{4\pi^2} l K_1(la), \\ \mathcal{F}_2 &= \frac{-V_0}{4\pi^2} \left\{ \frac{\lambda_1}{\mu_2} (k^2 - l^2) K_0(la) + \frac{2l}{a} K_1(la) \right\}.\end{aligned}$$

## APPENDIX B: RELATED PHASE SPEEDS

The phase speed of plane waves in fluids confined within elastic pipes in air is given by,<sup>22</sup>

$$c_f = \left\{ \rho \left( \frac{1}{B} + \frac{1}{M} \right) \right\}^{-1/2}, \quad (\text{B1})$$

where

$$M = \frac{E_p(a^2 - b^2)}{2[(1 + \nu)(a^2 + b^2) - 2\nu b^2]}, \quad (\text{B2})$$

and  $B$  is the bulk modulus of the fluid,  $\rho$  is its density,  $b$  and  $a$  are the inner and outer radii of the pipe,  $E_p$  is the youngs modulus of pipe, and  $\nu$  is its Poissons ratio. This is an approximate expression derived under a quasi-static approximation and is valid at low frequencies.

The phase speed of the Stoneley mode in fluid-filled boreholes without pipe is also given by Eq. (12),<sup>5</sup>

$$\text{with } M = \rho_F \beta^2, \quad (\text{B3})$$

and  $\rho_F$  is the density of the formation and  $\beta$  is the formation shear speed. This expression is valid at low frequencies, when the wavelength in the formation is much larger than the borehole radius. The exact dispersion relation is presented in Refs. 3 and 16.

<sup>1</sup>H. Y. Lee, "Drillstring Axial Vibration and Wave Propagation in Boreholes," Ph.D. thesis, Massachusetts Institute of Technology, Cambridge, MA (1991).

<sup>2</sup>D. S. Drumheller, "Acoustical properties of drillstrings," J. Acoust. Soc. Am. **85**, 1048–1064 (1989).

<sup>3</sup>M. A. Biot, "Propagation of elastic waves in a cylindrical bore containing a fluid," J. Appl. Phys. **23**, 997–1009 (1952).

<sup>4</sup>P. A. Heelan, "Radiation from a cylindrical source of finite length," Geophysics **18**, 685–696 (1953).

<sup>5</sup>J. E. White, *Underground Sound: Application of Seismic Waves* (Elsevier Science, New York, 1983).

<sup>6</sup>J. E. White and R. L. Sengbush, "Shear waves from explosive sources," Geophysics **28**, 1109–1119 (1963).

<sup>7</sup>E. W. Peterson, "Acoustic wave propagation along a fluid-filled cylinder," J. Appl. Phys. **45**, 3340–3350 (1974).

<sup>8</sup>A. Abo-Zena, "Radiation from a finite cylindrical explosive source," Geophysics **42**, 1384–1393 (1977).

<sup>9</sup>M. W. Lee and A. H. Balch, "Theoretical seismic radiation from a fluid-filled borehole," Geophysics **47**, 1308–1314 (1982).

<sup>10</sup>L. D. Lafleur and F. D. Shields, "Low-frequency propagation modes in a liquid-filled elastic tube waveguide," J. Acoust. Soc. Am. **97**, 1435–1445 (1995).

<sup>11</sup>W. L. Jacobi, "Propagation of sound waves along liquid cylinders," J. Acoust. Soc. Am. **21**, 120–127 (1949).

<sup>12</sup>V. Easwaran and M. L. Munjal, "A note on the effect of wall compliance on lowest-order mode propagation in fluid-filled/submerged impedance tubes," J. Acoust. Soc. Am. **97**, 3494–3501 (1995).

<sup>13</sup>J. E. Greenspon and E. G. Singer, "Propagation in fluids inside thick viscoelastic cylinders," J. Acoust. Soc. Am. **97**, 3502–3509 (1995).

<sup>14</sup>B. K. Sinha, T. J. Plona, S. Kostek, and S. K. Chang, "Axisymmetric



- propagation in fluid-loaded cylindrical shells. I—Theory,” *J. Acoust. Soc. Am.* **92**, 1132–1143 (1992).
- <sup>15</sup>T. J. Plona, B. K. Sinha, S. Kostek, and S. K. Chang, “Axisymmetric propagation in fluid-loaded cylindrical shells. II—Theory versus experiment,” *J. Acoust. Soc. Am.* **92**, 1144–1155 (1992).
- <sup>16</sup>C. H. Cheng and M. N. Toksöz, “Elastic wave propagation in a fluid-filled borehole and synthetic acoustic logs,” *Geophysics* **46**, 1042–1053 (1981).
- <sup>17</sup>K. Tubman, C. H. Cheng, S. P. Cole, and M. N. Toksöz, “Synthetic full waveform acoustic logs in radially layered boreholes. II—Poorly bonded casing,” *Geophysics* **51**, 902–913 (1986).
- <sup>18</sup>V. N. Rama Rao, “Acoustic transmission through fluid-filled pipes in boreholes,” Master’s thesis, Massachusetts Institute of Technology, Cambridge, MA (1991).
- <sup>19</sup>P. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes in Fortran* (Cambridge University Press, Cambridge, 1992).
- <sup>20</sup>F. B. Jensen, W. B. Kuperman, M. B. Porter, and H. Schmidt, *Computational Ocean Acoustics* (AIP Press, Woodbury, NY, 1994).
- <sup>21</sup>M. Bouchon, “Calculation of complete seismograms for an explosive source in a layered medium,” *Geophysics* **45**, 197–203 (1980).
- <sup>22</sup>P. M. Morse and H. Feshbach, *Methods of Theoretical Physics—II* (McGraw-Hill, New York, 1953).

# The arccos and Lommel formulations—Approximate closed-form diffraction corrections

Charles J. Daly<sup>a)</sup> and N. A. H. K Rao<sup>b)</sup>

Chester F. Carlson Center for Imaging Science, Rochester Institute of Technology, Rochester, New York 14623

(Received 30 December 1997; accepted for publication 9 March 1999)

A closed-form frequency-domain formalism for spatially integrated diffraction corrections is proposed. Spatially integrated diffraction corrections are necessary when trying to characterize material with ultrasonic probing. In the case of piston transducers and point receivers, the Lommel diffraction formulation is used when the excitation is monochromatic, and the arccos diffraction formulation is used when the excitation is impulsive. The Lommel and arccos formulations are usually treated separately; here, they are connected. Specifically, the arccos diffraction formulation and Lommel diffraction formulation are shown to form an approximate Fourier transform pair. Since the Lommel formulation is amenable to closed-form spatial integration, Lommel functions are used to derive diffraction corrections for unfocused piston transducers operating in receive-only (one-way) mode or transmit/receive (two-way) mode. Results obtained from the proposed closed-form frequency-domain formalism are qualitatively compared with results based on the closed-form time-domain or impulse-response formalism. It will be shown that the proposed frequency-domain formalism has theoretical and practical value. Finally, specific computational considerations are discussed as necessary. © 1999 Acoustical Society of America. [S0001-4966(99)03506-7]

PACS numbers: 43.20.Rz, 43.20.Fn, 43.20.Bi [DEC]

## INTRODUCTION

Diffraction corrections are necessary when trying to characterize material, such as germanium or human tissue, with ultrasound.<sup>1</sup> They are also necessary when trying to predict or calibrate transducer responses.<sup>2,3</sup> Traditionally, the Lommel diffraction formulation<sup>1,4</sup> has been used when the excitation is monochromatic, and the arccos diffraction formulation<sup>5,6</sup> has been used when the excitation is a delta function. In this theoretical paper, the two formulations are connected, and Lommel functions are used to derive closed-form spatially integrated diffraction corrections for unfocused piston transducers operating in pulsed mode.

Piston transducers have been researched for over 50 years. Indeed, closed-form spatially integrated diffraction corrections have already been derived using both the arccos diffraction formulation<sup>2,3,7</sup> and the Lommel diffraction formulation.<sup>8,9</sup> This paper complements and extends the existing literature in three important ways. First, the arccos and Lommel diffraction formulations are compared, and the Fourier equivalence of the two is rigorously demonstrated. The term *Fourier equivalence* implies the two formulations form an approximate Fourier transform pair. Second, the theoretical development serves to review and unify the arccos and Lommel diffraction formulations for unfocused piston transducers. Papers on focused piston transducers are listed in Refs. 8, 10, and 11.

The third point requires detailed discussion. This paper proposes a closed-form frequency-domain formalism for spa-

tially integrated one-way and two-way diffraction correction. The frequency-domain formalism is based on the Lommel diffraction formulation and, as such, serves as an alternative to the well-established time-domain or impulse response formalism based on the arccos diffraction formulation.<sup>2,6</sup> Thus, we are primarily interested in frequency-domain results, particularly for spatially integrated diffraction effects.

To establish the formalism, we assume delta function excitation of an infinitely baffled, unfocused piston transducer that has an infinite bandwidth or Dirac response.<sup>2</sup> To validate the proposed frequency-domain formalism, we inverse Fourier transform the frequency-domain results and obtain estimated impulse responses. These estimated responses are compared to responses obtained from the time-domain formalism. Our primary purpose in doing so is to establish the theoretical validity of the proposed formalism.

An immediate criticism, then, is that the proposed frequency-domain formalism will suffer from Gibb's phenomenon<sup>12</sup> because an infinite bandwidth can never be adequately sampled. The criticism is valid; however, we counter that real transducers are bandlimited. Thus, the frequency-domain formalism will be of interest and of value if it can be easily and accurately computed across some bandwidth of interest. We will show that it can.

Although we point out a few possible advantages of the frequency-domain formalism, it is not our intention to stimulate or engage in unproductive debate. Indeed, we acknowledge at the outset that the time-domain formalism is the gold standard in ultrasonic diffraction theory. This acknowledgment, however, should not preclude theoretical research into alternative formalisms.

The paper is organized as follows. Section I explains

<sup>a)</sup>Electronic mail: cjd6905@rit.edu

<sup>b)</sup>Electronic mail: narpci@rit.edu

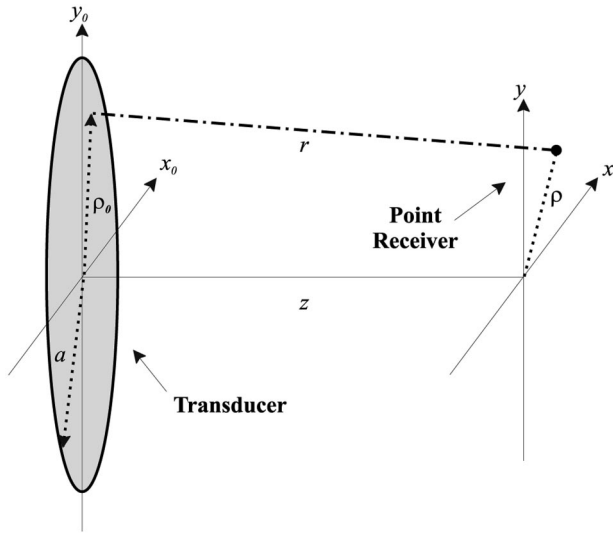


FIG. 1. Piston transducer and fictitious point receiver/scatterer.

more assumptions and defines terms and expressions. Section II demonstrates the Fourier equivalence of the arccos and Lommel diffraction formulations as an approximate Fourier transform pair. In Sec. III, closed-form spatially integrated diffraction corrections are derived. Specific computational considerations are discussed where appropriate. Detailed error analyses are not presented, and our results are compared with the existing literature in a qualitative fashion only.

## I. PRELIMINARIES

Results from scalar diffraction theory will be applied to the diffraction problem depicted in Fig. 1. A pressure-sensitive piston transducer of radius  $a$  is shown. In the theoretical development, all transducers have a Dirac response and are assumed to be unfocused. The Kirchoff boundary conditions along with an infinite rigid baffle and spatially uniform excitation are assumed. For mathematical tractability, dispersion and multiple scattering are considered negligible. Attenuation is ignored, and only frequency-independent scattering is considered.

The transmitting transducer is symmetrically positioned at the origin of the  $z=0$  plane. The position and dimensions of the pressure-sensitive receiver will be explicitly stated in the development. The velocity of the transducer disk, often denoted  $u_0$ , due to either monochromatic or impulsive excitation is assumed to be unity and does not appear in the development. The theory will be developed for unfocused transducers but should be easily extended to the focused case via the Lommel diffraction formulation for focused radiators developed by Papoulis (Ref. 13, pp. 351–353). Other assumptions will be explained as the need arises.

The term *one-way* implies that an ultrasonic transducer emits energy and another transducer coaxially located with the transmitter some distance  $z$  away acts as a receiver. The usual goal of one-way ultrasonic probing is to extract information about the medium between the transmitter and receiver. This information is encoded on the output voltage of the receiver. The output voltage is assumed to be propor-

tional, in some manner, to the total pressure on the face of the receiver,<sup>14,2,15</sup> and the total pressure is found by spatially integrating the incoming pressure field over the receiving aperture. Note the receiving aperture is referred to as a *measurement circle* by Williams.<sup>16</sup> When properly formulated, this spatial integration gives a quantitative estimate of attenuation at different frequencies due to one-way diffraction.

The term *two-way* implies reflection imaging in which diffraction occurs during transmission and reflection. The notion of one-way and two-way is subject to quantitative interpretation. Specifically, some authors<sup>1,2,8,17,18</sup> state that equations derived for the one-way case can be used for the two-way case by simply doubling the depth  $z$ . This claim is based on an optical or mirror-image interpretation of reflection imaging and involves perfect reflection from an infinite plane. On the other hand, some authors<sup>19–21</sup> state that two-way diffraction can be characterized by an autoconvolution of the one-way impulse response. This claim is based on a linear systems or autoconvolution interpretation of reflection imaging. Only the mirror-image interpretation of reflection imaging is considered here. Discussion of spatially integrated autoconvolution diffraction corrections can be found in Refs. 22 and 23.

A number of mathematical expressions involving Bessel functions will be encountered in the derivations,<sup>1,13,24</sup> and they are defined here. Following Wolf,<sup>24</sup>  $u$  and  $v$  are real variables,  $n$  is a non-negative integer, and  $J_n$  denotes a Bessel function of the first kind of order  $n$ .  $U_n$  and  $V_n$  are Lommel functions defined by

$$U_n(u, v) = \sum_{s=0}^{\infty} (-1)^s \left(\frac{u}{v}\right)^{n+2s} J_{n+2s}(v), \quad (1)$$

$$V_n(u, v) = \sum_{s=0}^{\infty} (-1)^s \left(\frac{v}{u}\right)^{n+2s} J_{n+2s}(v). \quad (2)$$

Because they are infinite summations, the Lommel functions can be computed only approximately, and these approximations can be programmed either recursively<sup>4,23</sup> or explicitly in a do-loop. Do-loops were used in this work.  $W_n$  and  $Y_n$  denote the related functions

$$W_n(u, v) = \sum_{s=0}^{\infty} (-1)^s (s+1) \left(\frac{v}{u}\right)^{n+2s} J_{n+2s}(v), \quad (3)$$

$$Y_n(u, v) = \sum_{s=0}^{\infty} (-1)^s (n+2s) \left(\frac{v}{u}\right)^{n+2s} J_{n+2s}(v). \quad (4)$$

$W_n$  are referred to here as Wolf functions, while  $Y_n$  are referred to as Hopkins functions.

Finally, since  $U_n$  converges too slowly for calculation when  $u/v > 1$ , the following formulas from Gray and Mathews [Ref. 4, p. 185, Eq. (20)] will prove useful:

$$U_{2n+1}(u, v) + V_{-2n+1}(u, v) = (-1)^n \sin\left(\frac{1}{2} \left[ u + \frac{v^2}{u} \right]\right), \quad (5a)$$

$$-U_{2n}(u, v) + V_{-2n}(u, v) = (-1)^n \cos\left(\frac{1}{2} \left[ u + \frac{v^2}{u} \right]\right). \quad (5b)$$

Special case formulas for  $u/v = 1$  can be found in Gray and Mathews.

## II. AN APPROXIMATE FOURIER TRANSFORM PAIR

In this section, derivations of the arccos and Lommel diffraction formulations are outlined, the two formulations are compared, and their equivalence as an approximate Fourier transform pair is rigorously demonstrated. The reader is referred to Oberhettinger,<sup>5</sup> Papoulis (Ref. 13, pp. 329–331), Stepanishen,<sup>6</sup> and Harris<sup>24</sup> for complete details on the derivations. The Lommel diffraction formulation is discussed first.

Assuming monochromatic excitation of the transducer in Fig. 1, the velocity potential sensed by a fictitious point receiver  $P$  located at some off-axis distance  $\rho = \sqrt{x^2 + y^2}$  can be written

$$H_1(\rho, z, \omega) = \frac{1}{2\pi} \int_{\sigma_o} f(\sigma_o) \frac{e^{-jkr}}{r} d\sigma_o, \quad (6)$$

where  $\sigma_o$  is the area of the transmitter. Throughout this paper, the subscript  $o$  denotes the  $z = 0$  plane, and the subscript 1 denotes one-way propagation. The velocity distribution across the face of the transducer is  $f(\sigma_o)$  and is unity due to the assumption of spatially uniform excitation.

Equation (6) represents the Rayleigh-Sommerfeld diffraction integral for an infinitely baffled transducer.<sup>6,25</sup> It is important to note that Eq. (6) is based on Huygen's principle and represents continuous integration of the free-space Green's function for a point source over a continuum of point sources which contains, mathematically speaking, an infinite number of point sources. As an aside, the velocity potential  $H_1(\rho, z, \omega)$  is designated  $\psi_p$  in Ref. 1.

As usual, the time-dependence of  $H_1(\rho, z, \omega)$  on  $e^{j\omega t}$  is implied. The spatial wave number,  $k$ , is related to temporal frequency,  $\omega$ , via  $k = \omega/c$ . Thus, the dependence of  $H_1(\rho, z, \omega)$  on  $\omega$  is implicit in two ways.

The Fresnel approximation,<sup>13</sup> in conjunction with circular symmetry, allows Eq. (6), to be estimated,

$$\begin{aligned} H_1(\rho, z, \omega) &\approx \hat{H}_1(\rho, z, \omega) \\ &= \frac{1}{z} e^{-jk(z + (\rho^2/2z))} \int_0^a e^{-jk(\rho_o^2/2z)} \\ &\quad \times J_0\left(\frac{k\rho}{z}\rho_o\right) \rho_o d\rho_o, \end{aligned} \quad (7)$$

where  $\rho_o = \sqrt{x_o^2 + y_o^2}$  is the off-axis distance at the  $z = 0$  plane and  $\rho = \sqrt{x^2 + y^2}$  is the off-axis distance at the  $z$  plane (Ref. 13, p. 330). The hat notation signifies estimation throughout this paper.

A prominent and familiar feature of Fresnel diffraction is its interpretation as a convolution involving a quadratic phase term.<sup>13,26</sup> This feature is obscured in Eq. (7). However, if the singularity function,

$$p_a(\rho_o) = \begin{cases} 1, & \rho_o \leq a; \\ 0, & \rho_o > a, \end{cases} \quad (8)$$

is introduced in the integrand of Eq. (7) and the upper limit of integration changed to  $\infty$ , then Eq. (7) becomes

$$\begin{aligned} \hat{H}_1(\rho, z, \omega) &= \frac{1}{z} e^{-jk(z + \rho^2/2z)} \int_0^\infty p_a(\rho_o) e^{-jk\rho_o^2/2z} \\ &\quad \times J_0\left(\frac{k\rho}{z}\rho_o\right) \rho_o d\rho_o \end{aligned} \quad (9)$$

which may be interpreted as the Hankel transform of the product of the singularity function  $p_a(\rho_o)$  and a quadratic phase term. The convolution theorem for Hankel transforms allows Eq. (9) to be rewritten

$$\hat{H}_1(\rho, z, \omega) = \frac{1}{k} e^{-jk(z + \rho^2/2z)} \left[ \frac{a}{\rho} J_1\left(\frac{ka\rho}{z}\right) * \frac{1}{j} e^{jk\rho^2/2z} \right], \quad (10)$$

where the convolution is with respect to  $k\rho/z$ . The familiar interpretation of Fresnel diffraction is made explicit in Eq. (10).

Equations (7) and (10) can be calculated numerically, but a closed-form expression would simplify the calculation. Equation (7) can be cast in closed form via Lommel functions. The closed-form result is

$$\begin{aligned} \hat{H}_1(\rho, z, \omega) &= \frac{1}{k} \exp\left[-j\left(kz + \frac{v^2}{2u} + \frac{u}{2}\right)\right] \\ &\quad \times [U_1(u, v) + jU_2(u, v)], \end{aligned} \quad (11)$$

where the substitutions  $u = ka^2/z$  and  $v = ka\rho/z$  result in more compact notation. Equation (11) is the Lommel diffraction formulation; it is easily programmed because of its closed form and is amenable to either direct or recursive calculation.

Seki *et al.*<sup>1</sup> used a variant of Eq. (11) to calculate pressure as a function of depth and off-axis distance; pressure can be obtained from Eq. (11) by multiplying by  $j\omega\rho$ , where  $\rho$  is medium density. Results obtained from Eq. (11) are plotted in Fig. 2. The plots agree well with those found in Seki's 1956 paper and serve to validate the use of Eq. (11) in the present work.

Three computational issues deserve mention here. First, Eq. (5) must be used when the ratio  $u/v > 1$ . Specifically,  $U_n$  must be expressed in terms of  $V_n$  because  $U_n$  converges too slowly for calculation when  $u/v > 1$ . Second, sufficient terms must be included in finite do-loop approximations to  $U_n$ . In our work,  $J_n$  up to and including  $n = 52$  have been used to calculate  $U_2$ . This high an order can be used if underflow is not too objectionable. Third, on-axis values of Eq. (11), that is when  $\rho = 0$ , can be calculated via appropriate handling of the Lommel functions (Ref. 27, p. 540) when  $v = 0$ , or they can be calculated from a separate formula which is easily derived by explicitly integrating Eq. (7) with  $\rho = 0$ . The latter method was used in our computations.

In contrast to the Lommel diffraction formulation which can only be derived analytically, the arccos diffraction formulation can be derived either analytically<sup>5</sup> or geometrically.<sup>6</sup> A modified version of Oberhettinger's analytic derivation is outlined here because Oberhettinger's derivation points to the Fourier equivalence of the arccos and

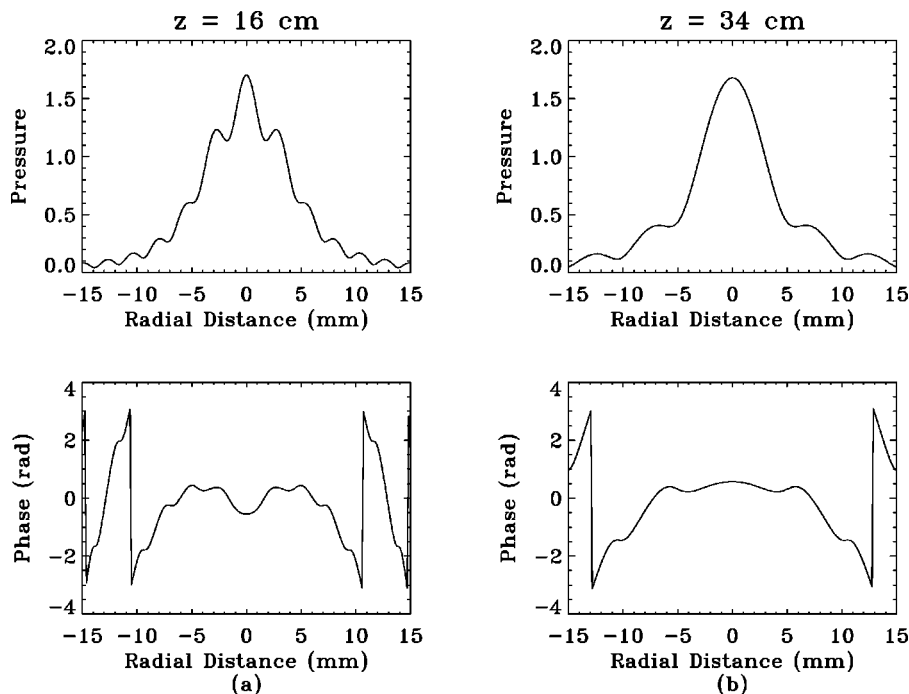


FIG. 2. Pressures after Seki *et al.*<sup>1</sup> via Eq. (11).

Lommel diffraction formulations; the modification involves application of the Fourier transform instead of the Laplace transform.

Assuming the transducer in Fig. 1 is excited by an impulse, the velocity potential sensed by a fictitious point receiver located at  $P$  can be written

$$h_1(\rho, z, t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} H_1(\rho, z, \omega) e^{j\omega t} d\omega$$

$$= \mathcal{F}^{-1}\{H_1(\rho, z, \omega)\}, \quad (12)$$

where  $\mathcal{F}^{-1}$  denotes the inverse Fourier transform. Thus, Eq. (12) is the inverse Fourier transform of  $H_1(\rho, z, \omega)$  in Eq. (6). Note,  $h_1(\rho, z, t)$  is designated  $\Phi_D(t)$  in Oberhettinger's paper.<sup>5</sup>

With a transformation from rectangular to polar coordinates ( $x = \rho \cos \phi$  and  $y = \rho \sin \phi$ ),  $H_1(\rho, z, \omega)$  in Eq. (6) becomes

$$H_1(\rho, z, \omega) = \int_{\phi_o=0}^{2\pi} \int_{\rho_o=0}^a \{[\rho^2 + \rho_o^2 - 2\rho\rho_o \times \cos(\phi - \phi_o) + z^2]^{-1/2} \exp[-jk(\rho^2 + \rho_o^2 - 2\rho\rho_o \cos(\phi - \phi_o) + z^2)]\} \rho_o d\rho_o d\phi_o. \quad (13)$$

After several steps, Oberhettinger obtains exact expressions for  $h_1(\rho, z, t)$ . For  $\rho < a$ ,

$$h_1(\rho, z, t) = \begin{cases} 0, & ct < z; \\ c, & z < ct < R'; \\ \frac{c}{\pi} \arccos \left[ \frac{(ct)^2 - z^2 + \rho^2 - a^2}{2\rho((ct)^2 - z^2)^{1/2}} \right], & R' < ct < R; \\ 0, & ct > R, \end{cases} \quad (14)$$

and, for  $\rho > a$ ,

$$h_1(\rho, z, t) = \begin{cases} 0, & ct < R'; \\ \frac{c}{\pi} \arccos \left[ \frac{(ct)^2 - z^2 + \rho^2 - a^2}{2\rho((ct)^2 - z^2)^{1/2}} \right], & R' < ct < R; \\ 0, & ct > R. \end{cases} \quad (15)$$

Taken together, Eqs. (14) and (15) represent the arccos diffraction formulation where  $R' = \sqrt{z^2 + (a - \rho)^2}$  and  $R = \sqrt{z^2 + (a + \rho)^2}$ .

Note that Eq. (13), which is the Rayleigh-Sommerfeld integral for diffraction from a piston transducer, led to the arccos diffraction formulation. On the other hand, Eq. (7), which is the Fresnel approximation to Eq. (13), led to the Lommel diffraction formulation. Thus, we conclude that the arccos and Lommel diffraction formulations are closely related, and this relationship is explored more fully later in this section.

For now, we do well to describe the well-known behavior of the arccos diffraction formulation.<sup>6</sup> Figure 3, which will be discussed in detail later, can be used as a visual aid. For a fixed depth  $z$ , the on-axis velocity-potential impulse response  $h_1(\rho, z, t)$  is a rectangular pulse starting at  $t = z/c$ ; its amplitude is  $c$ . As  $\rho$  increases, the start time of the pulse remains  $t = z/c$  but the trailing edge of the pulse moves closer to  $t = z/c$ . Simultaneously, the fall time of the trailing edge increases, and the trajectory of the fall is governed by the arccos term in Eq. (14). In short, the pulslike nature of the impulse response gradually decays with increasing  $\rho$ . For  $\rho > a$ , the impulse response no longer resembles a rectangular pulse, and its maximum value is something less than  $c$ . In addition, its start time is delayed in proportion to  $\rho$ .

For a fixed off-axis distance  $\rho$ ,  $h_1(\rho, z, t)$  has the same general shape at any depth  $z$  but is compressed in time as  $z$

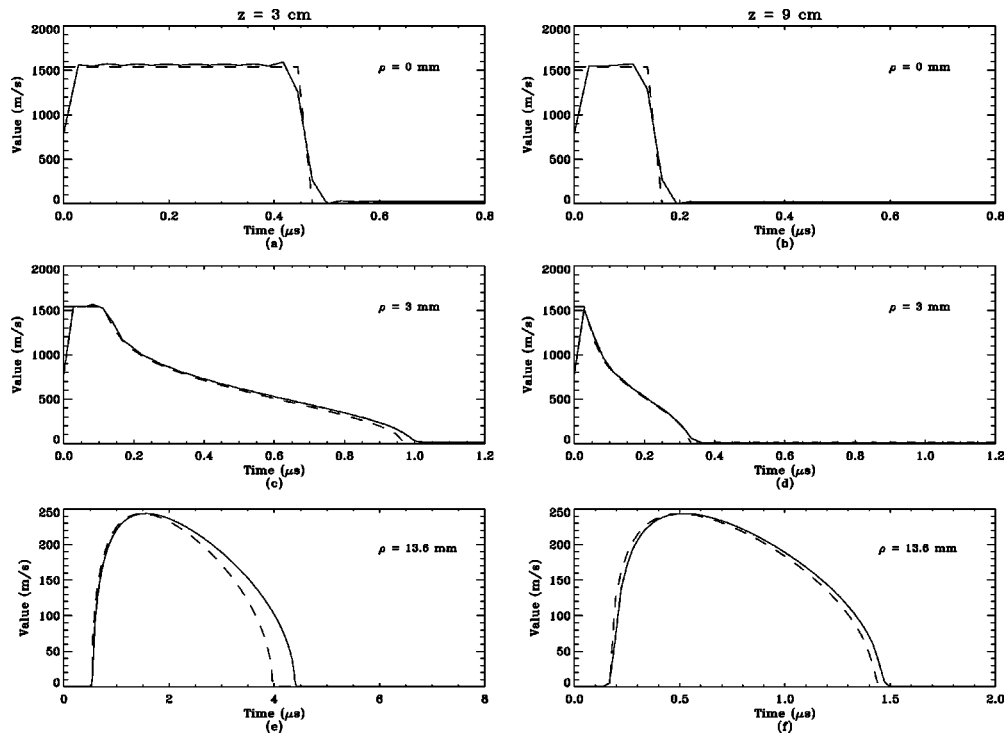


FIG. 3. One-way point-receiver impulse responses for the Lommel (solid) and arccos (dashed) diffraction formulations.

increases. The relationship can be quantified by expanding  $R'$  and  $R$  via binomial expansion, subtracting the smaller from the larger, and dividing for different values of  $z$ . The result is that for large  $z$ ,  $h_1(\rho, z, t) = h_1(\rho, z, z_r t/z)$ , where  $z_r$  is some appropriately chosen reference plane.<sup>2</sup> Researchers in wavelet theory might find this an interesting physical problem since time scaling arises in a natural fashion.

As was just mentioned, the arccos and Lommel diffraction formulations are closely related. The Lommel diffraction formulation is a monochromatic frequency-domain solution based on the Fresnel approximation to the Rayleigh-Sommerfeld integral of scalar diffraction theory. Hence, the derivation of Lommel diffraction formulation permits monochromatic diffraction from a circular aperture to be interpreted as convolution involving a depth-dependent quadratic phase factor as in Eq. (10). On the other hand, the arccos diffraction formulation is a polychromatic time-domain solution based on the Rayleigh-Sommerfeld integral. The arccos formulation permits an interpretation of impulsive diffraction from a piston transducer in terms of a depth-dependent time-scaling operation.

The Fourier equivalence of the Lommel and arccos diffraction formulations will now be demonstrated. Consider again the general form of the arccos diffraction formulation, Eq. (12). In this equation,  $h_1(\rho, z, t)$  is the inverse Fourier transform of some unspecified function,  $H_1(\rho, z, \omega)$ . Of course,  $H_1(\rho, z, \omega)$  could be obtained by calculating  $h_1(\rho, z, t)$  and Fourier transforming the result. Doing so, however, does not advance the goal of finding closed-form spatially integrated diffraction corrections.

Recall the Lommel diffraction formulation,  $\hat{H}_1(\rho, z, \omega)$  in Eq. (11), is a closed-form estimate of  $H_1(\rho, z, \omega)$  in Eq. (13). Further, the Lommel diffraction formulation is written

in terms of  $k = \omega/c$ . Theoretically, discrete Fourier coefficients for the arccos diffraction formulation can be estimated using  $\hat{H}_1(\rho, z, \omega)$ . These coefficients can then be inverse Fourier transformed to obtain

$$\hat{h}_1(\rho, z, t) = \mathcal{F}^{-1}\{\hat{H}_1(\rho, z, \omega)\}, \quad (16)$$

which is, as already mentioned, an estimate of the impulse response predicted by the arccos diffraction formulation.

Thus, the Lommel diffraction formulation  $\hat{H}_1(\rho, z, \omega)$  is a closed-form approximation of the Fourier transform of the arccos diffraction formulation  $h_1(\rho, z, t)$ , and Eq. (11) can be used to sample the arccos diffraction formulation in the frequency domain. In short, the Lommel and arccos diffraction formulations form an approximate Fourier transform pair,

$$\mathcal{F}\{h_1(\rho, z, t)\} \approx \hat{H}_1(\rho, z, \omega), \quad (17)$$

where  $\mathcal{F}$  is the Fourier transform.

The Lommel diffraction formulation was used in conjunction with Eq. (16) to compute  $\hat{h}_1(\rho, z, t)$ , an estimate of  $h_1(\rho, z, t)$ , for three off-axis positions at two depths,  $z = 3$  and  $z = 9$  cm. The speed of sound was set at  $c = 1540$  m/s, and the diameter of the piston was set at  $2a = 13$  mm. We reiterate that the transducer was assumed to have an infinitely broadband or Dirac response, and the excitation was assumed to be an impulse. The sampling frequency was set at  $f_s = 36$  MHz; thus, the Nyquist frequency was 18 MHz. Note the sampling rate is consistent with 2X oversampling of a real 2.25-MHz piston transducer with a cutoff frequency of 4.5 MHz. More will be said about real transducers later.

The results are shown in Fig. 3. The off-axis positions are annotated in the figure. The impulse responses for a given  $\rho$  are plotted on the same time scale, referenced to  $t$

$=z/c$ , to emphasize the depth-dependent time scaling mentioned earlier. In all figures where the two diffraction formulations are compared, Lommel-derived results are plotted with solid lines, while arccos-derived results are plotted with dashed lines. In this work, the arccos diffraction formulation is the gold standard against which the Lommel diffraction formulation is compared.

Before analyzing the results, we discuss Gibb's phenomenon<sup>12</sup> and its impact on this work. Under the stated assumptions and for practical geometries, impulse responses computed using the arccos diffraction formulation have compact support in the time domain; consequently, their Fourier transforms have infinite bandwidth in the frequency domain. In practice, the Lommel diffraction formulation can be sampled only over some finite bandwidth (truncation in the frequency domain); consequently, impulse responses based on the Lommel diffraction formulation will suffer from Gibb's phenomenon.

As a result, we expect that Lommel-based results will fail to capture temporal discontinuities and will simultaneously exhibit ringing in the neighborhood of any temporal discontinuities. The degree of failure and extent of ringing are functions of the sampling rate; higher sampling rates will capture temporal discontinuities more faithfully but simultaneously introduce more ringing. In short, impulse responses based on the Lommel diffraction formulation and Eq. (16) can never show exact agreement with those based on the arccos diffraction formulation in Eqs. (14) and (15).

The plots in Fig. 3(a) and (b) show on-axis impulse responses. As was explained earlier, the on-axis impulse response for a piston transducer is a rectangular pulse of amplitude  $c$  that gets compressed in time with increasing depth  $z$ . The on-axis impulse responses computed with the Lommel diffraction formulation capture this behavior. As expected, they do not capture the discontinuities at the beginning and end of each pulse. Note that for  $\rho=0$ , the Fresnel approximation is very good. Thus, it can be argued that the disagreement at the discontinuities is due primarily to Gibb's phenomenon.

Figure 3(c) and (d) show impulse responses for  $\rho=3$  mm. Here again, the disagreement, mostly near discontinuities, is due to Gibb's phenomenon and not due to the Fresnel approximation. The Lommel-based results are consistent with the results predicted by the arccos diffraction formulation.

The plots in Fig. 3(e) and (f) show impulse responses for  $\rho=13.6$  mm. Since  $\rho>a$ , each impulse response will have a maximum amplitude less than  $c$  and will start at some time later than  $t=z/c$ . This behavior is confirmed in the plots. Note that the Lommel diffraction formulation overestimates the time duration of both impulse responses. Since  $\rho$  is large, the disagreement here is primarily due to the fact that the Lommel diffraction formulation is based on the Fresnel approximation.

Overall, the results show satisfactory agreement and confirm the validity of the Fourier equivalence of the arccos and Lommel diffraction formulations as an approximate Fourier transform pair. Clearly, the magnitude and phase responses computed using the Lommel diffraction formulation

capture the salient features of the arccos diffraction formulation. Thus, no discussion or graphs of frequency-domain results are included at this point. We will discuss frequency-domain results in great detail in the next section.

The three computational issues already mentioned apply here, and five new ones require discussion. Because these issues will resurface, these issues will be referred to as the *five general computational issues*. First, the arccos impulse is real; hence, Fourier coefficients need be calculated for positive frequencies only. Negative-frequency coefficients are simply the complex conjugate of the positive-frequency coefficients.

This computational benefit is negated by the fact that the Lommel diffraction formulation is ill-defined at  $\omega=0$ . Thus, a DC frequency coefficient cannot be calculated directly. It can, however, be indirectly calculated by exploiting the positivity of the arccos diffraction formulation. In this work, discrete Fourier coefficients were calculated via Eq. (11) and inverse Fourier transformed with the fast Fourier transform (FFT). The resulting samples were forced to be greater than or equal to zero. In short,  $\hat{h}_1(\rho, z, t)$  was forced to be positive. These two issues represent a trade-off inherent in any Lommel-based solution.

The  $k$  in the denominator of Eq. (11) is the third issue. Since the coefficients calculated from the Lommel formulation are ultimately sent to an FFT algorithm, continuous or discrete frequencies may be used with  $k$ . Discrete frequencies were used in our implementation. If the results are to be scaled to a maximum value of unity, the choice is immaterial.

Fourth, as explained earlier, estimated impulse responses will suffer from ringing due to Gibb's phenomenon due to truncation in the frequency domain. If desired, this artifact can be reduced with frequency-domain windowing; a window  $w(f) = \text{sinc}(0.25\pi f/f_s)$ , where  $\text{sinc}(x) = \sin(\pi x)/(\pi x)$ , was used to produce the results shown in Fig. 3. The window is admittedly *ad hoc*, but it produced satisfactory results.

Finally, Eq. (11) gives no indication of how many frequency samples are required in estimating the arccos impulse response. For a given off-axis position  $\rho$  and sampling frequency  $f_s = 1/\Delta t$ , the minimum number of samples required can be computed via  $(R-z)/(c\Delta t)$  or  $(R-R')/(c\Delta t)$ , whichever is appropriate. Further, when comparing the two formulations, accurate book-keeping in terms of sampling frequency, zero-padding, amplitude scaling, and phase is essential because results are being computed in conjugate domains.

Clearly, the Lommel diffraction formulation is more difficult to compute than the arccos diffraction formulation. Nonetheless, this section has formally connected the two formulations, and this connection is of historical and theoretical interest. Indeed, it appears to be of practical interest. For example, Chen *et al.* used the Lommel diffraction formulation in their 1994 paper on acoustic coupling to and from a flat plate.<sup>8</sup> Furthermore, the Lommel diffraction formulation may find application in numerical computation of complicated diffraction expressions<sup>21,28</sup> involving piston transducers and the free-space Green's function for a point source.

### III. SPATIALLY INTEGRATED DIFFRACTION CORRECTION

Over 20 years ago, Huntington, *et al.*,<sup>27</sup> Williams,<sup>16,18</sup> Seki *et al.*,<sup>1</sup> Bass,<sup>14</sup> Rhyne,<sup>3</sup> and Rogers and Van Buren<sup>9</sup> researched closed-form spatially integrated diffraction corrections. Khimunin<sup>28</sup> and Benson and Kiyohara<sup>31</sup> reported numerical results. More recently, Cassereau *et al.*<sup>2</sup> and Chen *et al.*<sup>8</sup> have continued the work of the early researchers.

The researchers just mentioned derived or calculated one-way diffraction corrections and ultimately invoked the optical or mirror-image interpretation of ultrasound to extend the one-way corrections to the two-way case; we invoke the same interpretation here. The works of Rhyne, Cassereau *et al.*, and Rogers and Van Buren are particularly germane to this paper, and more will be said about them in this section.

Given that one-way diffraction has been researched so extensively, we must explain why we are revisiting the topic. First, we wish to establish that the Fourier equivalence of the arccos and Lommel diffraction formulations is valid for spatially integrated diffraction. Second, we will gain new insight into diffraction from a piston transducer and derive at least one new equation of academic, if not practical, interest. Finally, we will gain confidence that the Lommel diffraction formulation can be applied in the numerical computation of complicated diffraction expressions involving piston transducers.

Closed-form results will be derived for two cases: (i) a piston receiver with radius  $b \leq a$  and (ii) a piston receiver with infinite radius. In both cases, the transmitter, a piston transducer with radius  $a$ , and the receiver are coaxially located and separated by a distance  $z$ . Spatially integrating the Lommel diffraction formulation [Eq. (11)] for the first case yields

$$\langle \hat{H}_1(z, \omega) \rangle_b = 2\pi \int_0^b \hat{H}_1(\rho, z, \omega) \rho d\rho, \quad (18)$$

where the symbol  $\langle \rangle$  subscripted with  $b$  denotes spatial integration over a disk of radius  $b$ . Note the angular integration from 0 to  $2\pi$  has been completed.

The integral in Eq. (18) can be solved for  $b \leq a$  with Wolf functions. Specifically, expand the integrand in terms of the  $V_n$  Lommel functions [Eq. (5)] and integrate using Lemma 9 in Ref. 24. With  $v_b = kab/z$  and  $u = ka^2/z$ , the result is

$$\begin{aligned} \langle \hat{H}_1(z, \omega) \rangle_b = & -\frac{2\pi z}{k^2} e^{-jkz} \left\{ j \frac{v_b^2}{2u} + e^{-j((u/2) + (v_b^2/2u))} \right. \\ & \times [2W_2(u, v_b) - jW_1(u, v_b) \\ & \left. + jW_3(u, v_b) \right\}, \quad (19) \end{aligned}$$

which can be simplified via the Hopkins functions to

$$\begin{aligned} \langle \hat{H}_1(z, \omega) \rangle_b = & -\frac{2\pi z}{k^2} e^{-jkz} \left\{ j \frac{v_b^2}{2u} + e^{-j((u/2) + (v_b^2/2u))} \right. \\ & \left. \times [Y_2(u, v_b) - jY_1(u, v_b)] \right\}. \quad (20) \end{aligned}$$

Equation (20) is relatively new in the literature on diffraction from an unfocused piston transducer.<sup>22,29</sup> It is, however, a difficult expression to compute and, as a result, may be of academic interest only. Nonetheless, the Hopkins functions are highly convergent, and the computational burden of calculating them may be eased by recursion relations.<sup>33</sup>

Equation (20) is valid for  $b \leq a$ ; however, only the case  $b = a$  is investigated. Using relations developed by Wolf,<sup>23,30</sup> it can easily be shown that Eq. (20) reduces to

$$\begin{aligned} \langle \hat{H}_1(z, \omega) \rangle_a = & -\frac{2\pi z}{k^2} e^{-j[kz + (u/2)]} \left\{ j \frac{u}{2} e^{j(u/2)} + e^{-j(u/2)} \right. \\ & \left. \times \left[ \frac{u}{2} J_1(u) - j \frac{u}{2} J_0(u) \right] \right\}. \quad (21) \end{aligned}$$

Unlike Eq. (20), Eq. (21) is relatively easy to compute. We remark here that Eq. (21) is, with the exception of a multiplicative constant, the same as the result derived by Rogers and Van Buren.<sup>9</sup> Thus, comparing Eq. (21) to their result will yield no new historical insight.

Instead, the validity of Eq. (21) was checked against the work done by Bass and Williams. The parameters used were the same as in Bass's 1958 article:  $c = 1200$  m/s,  $a = 1$  cm, and  $f = 0.956$  MHz. The data were obtained from (a) Bass's 1958 equation [Ref. 14, Eq. (14)], (b) Williams' 1970 equation [Ref. 18, Eq. (6)], and (c) Eq. (21). Note Williams (Ref. 18, p. 286) corrected two typos in Bass's 1958 equation.

The squared-magnitude results in decibels (dB) are presented in Figs. 4 and 5. The graphs show the spatially integrated diffraction effects, due to monochromatic excitation, plotted as a function of  $S = z\lambda/a^2$ ; two graphs of Eq. (21) are included for the purposes of comparison. The oscillatory behavior of Bass's result at low  $S$  is due to the small number of terms used in his equation. The overall results, however, show excellent agreement, and the plots confirm the well-established fact that attenuation due to diffraction increases with depth  $z$ . Thus, we have gained new academic perspective on classic research with the help of Eq. (21), which is a special case of Eq. (20).

We now extend the Fourier equivalence of the arccos and Lommel diffraction formulations to spatially integrated diffraction effects. Spatially integrating the velocity-potential transfer function  $H_1(\rho, z, \omega)$  and subsequently inverse Fourier transforming the result yields

$$\langle h_1(z, t) \rangle_b = \mathcal{F}^{-1} \{ \langle H_1(z, \omega) \rangle_b \}. \quad (22)$$

Now, we invoke the Fourier equivalence of the arccos and Lommel diffraction formulations and write

$$\langle \hat{h}_1(z, t) \rangle_b = \mathcal{F}^{-1} \{ \langle \hat{H}_1(z, \omega) \rangle_b \}. \quad (23)$$

Hence, Eq. (20) can be used to estimate the spatially integrated impulse response associated with the arccos diffraction formulation.

Theoretically, substituting  $k = \omega/c$  in Eq. (23) should allow one to estimate the Fourier coefficients of the spatially integrated arccos impulse response. These coefficients can then be inverse Fourier transformed to obtain an estimate of the spatially integrated arccos impulse response. This reason-



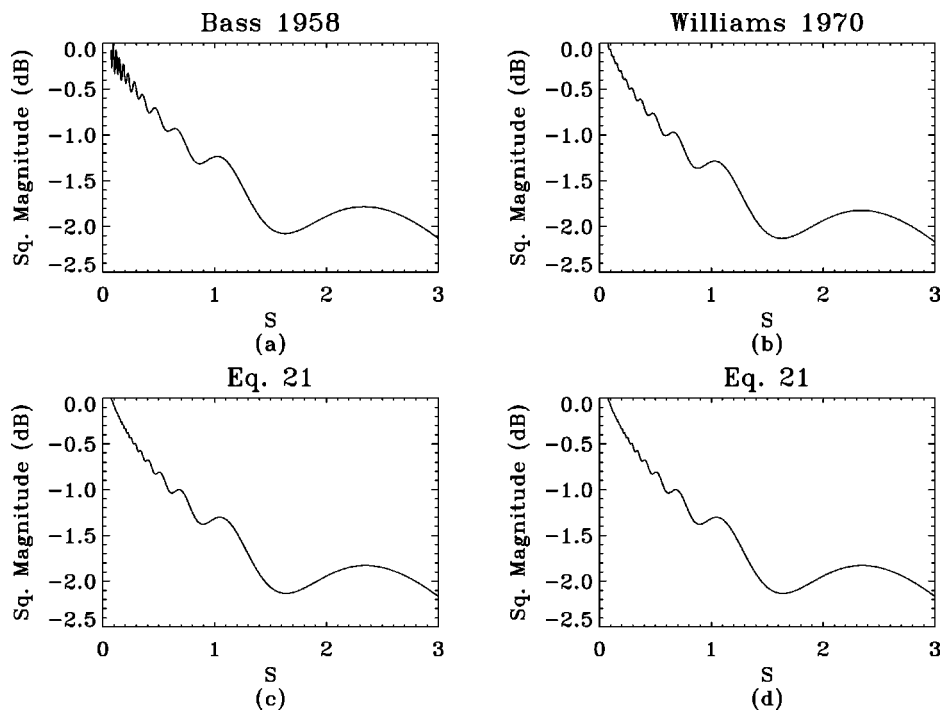


FIG. 4. Attenuation caused by diffraction as a function of  $S$  in both the near and far fields.

ing is simply an extension of the Fourier equivalence of the Lommel and arccos diffraction formulations developed for a point receiver in the previous section.

Some discussion is required before computing and comparing spatially integrated impulse responses. First, Rhyne<sup>3</sup> derived closed-form expressions for the spatially integrated arccos diffraction formulation for the case  $b = a$ ; it serves as the gold standard in this work. As an important aside, Cassereau *et al.*<sup>2</sup> and, later, Daly and Rao,<sup>7</sup> generalized Rhyne's work; their closed-form expressions are completely general and include Rhyne's result as a special case.

Second, impulse responses computed using Rhyne's expressions have compact support in the time domain; conse-

quently, their Fourier transforms have infinite bandwidth in the frequency domain. Like the Lommel diffraction formulation, Eq. (21) must be sampled over some finite bandwidth; consequently, impulse responses based on this equation will suffer from Gibb's phenomenon. Thus, the comments made earlier about impulse responses based on the Lommel diffraction formulation apply here. Those comments should be kept in mind when the comparison is made.

Figure 6(a) and (b) show spatially integrated one-way impulse responses estimated via Eq. (21) (solid lines) and spatially integrated one-way impulse responses calculated by using Rhyne's expression (dashed lines). The impulse responses were calculated for  $b = a$  at two depths:  $z = 3$  and

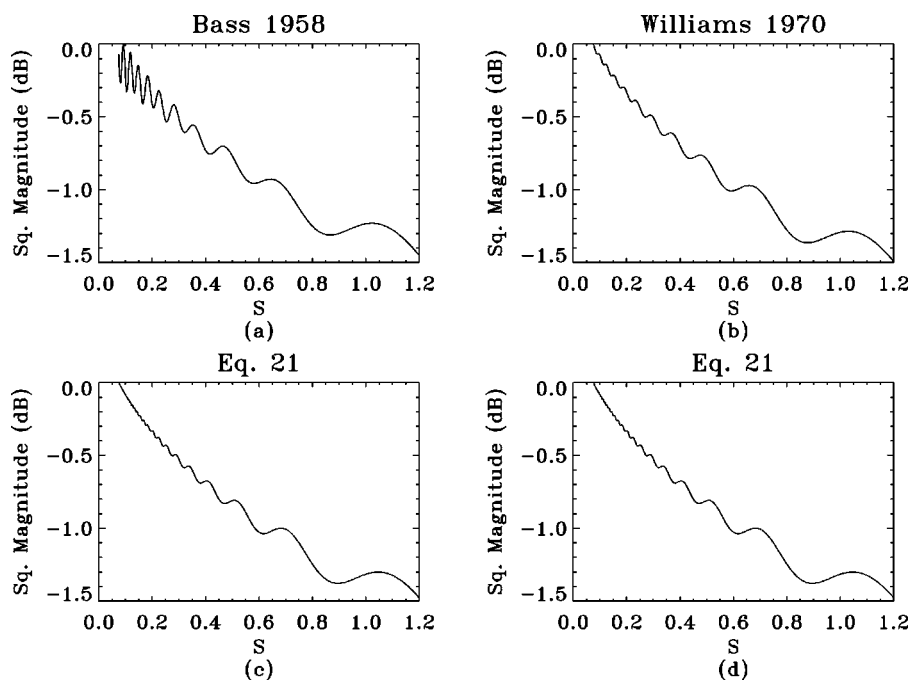


FIG. 5. Attenuation caused by diffraction as a function of  $S$  in the near field.

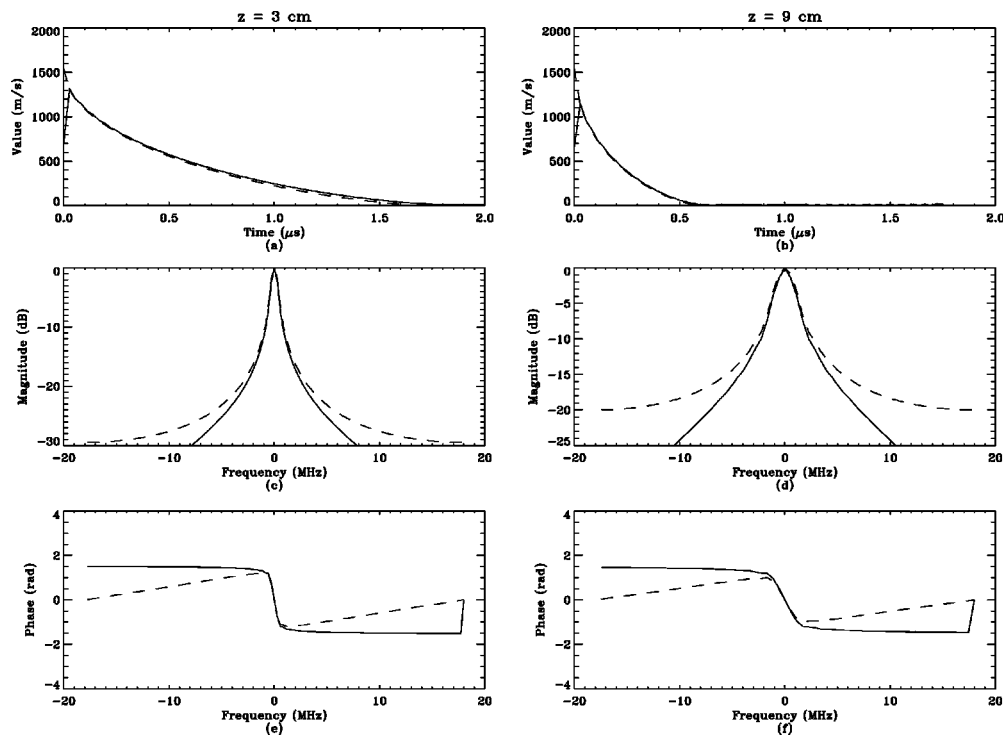


FIG. 6. Spatially integrated one-way impulse responses: Eq. (21) (solid) and Rhyne's expression (Ref. 3) (dashed).

$z=9$  cm. The speed of sound was set at  $c=1540$  m/s, and the diameter of the piston was set at  $2a=13$  mm. The transducer was assumed to have an infinitely broadband response, and the excitation was assumed to be an impulse. The sampling frequency was set at  $f_s=36$  MHz; thus, the Nyquist frequency was 18 MHz.

With the exception of discontinuities, the impulse responses based on Eq. (21) are consistent with the results computed using Rhyne's expression and results computed by Kuc and Regula;<sup>34</sup> see Ref. 2 for an excellent discussion on the origin and effects of temporal discontinuities in spatially integrated impulse responses.

The five general computational issues mentioned in the previous section apply here. The window  $w(f)$  discussed earlier was used in computing the Lommel-based impulse responses. Thus, ringing due to Gibb's phenomenon is reduced in the plots at the expense of a small amount of low-pass filtering. The impulse responses show satisfactory agreement. It is also important to reiterate how easy Eq. (21) is to compute. Furthermore, the utility of Eq. (21) is obvious if characterization of diffraction effects in the frequency domain is the main objective.

Figure 6(b) and (c) and Fig. 6(d) and (e) show the squared magnitude responses (dB) and the phase responses associated with the impulse responses in Fig. 6(a) and (b), respectively. Taking an optimistic point of view, we can say the magnitude responses show satisfactory agreement, particularly at the lower frequencies. Indeed, better agreement can be had at higher frequencies if the sampling frequency is increased, but the cost is more samples.

The phase responses do not agree as favorably. This is not surprising when one considers the physical origins of the results being compared. Specifically, the arccos-derived re-

sults are based on the Rayleigh-Sommerfeld diffraction integral, while the Lommel-derived results are based on the Fresnel diffraction integral. Hence, the two diffraction integrals differ primarily in terms of their phase.<sup>26</sup> This, in conjunction with Gibb's phenomenon, helps explain the phase differences exhibited in the plots.

It is crucial to note Eq. (21) was derived under the assumption of an ideal piston transducer with a Dirac response. Thus, Eq. (21) is completely general in terms of frequency. Real transducers, however, are bandlimited.

For example, consider a real 2.25 MHz unfocused piston transducer with diameter  $2a=13$  mm. A typical bandwidth for such a transducer is 2 to 4 MHz centered at 2.25 MHz. Clearly, the results shown in Fig. 6 apply to the real transducer just described. Indeed, they apply quite well, particularly in terms of magnitude, with just 2X oversampling. Thus, if a diffraction correction were desired for this transducer, Eq. (21) could be used to calculate an inverse filter directly in the frequency domain. Furthermore, higher sampling rates could be used, and the results applied to real transducers operating at a higher frequency than 2.25 MHz.

Of course, the time-domain formalism developed by Rhyne<sup>3</sup> and extended by Cassereau *et al.*<sup>2</sup> could be used, and the impulse-response formalism is, in fact, more general. On the other hand, Eqs. (20) and (21) are amenable to calculation directly in the frequency domain across any bandwidth of interest, and Eq. (21) is very easy to compute. Clearly though, the time-domain formalism is easier to compute and more general than the Lommel-based equations. Nonetheless, we have demonstrated the validity of the proposed frequency-domain formalism for spatially integrated one-way diffraction.

More insight can be gained by considering the second

case: a piston receiver of infinite extent located some distance  $z$  from the transmitter. Consideration of this case leads to a theoretical result which further illustrates the utility of the proposed frequency-domain formalism. As an aside, this case can also be investigated in the time domain using the spatially integrated impulse formalism; the reader is referred to Ref. 2 for details.

Spatially integrating the Lommel diffraction formulation of Eq. (11) for this case yields

$$\langle \hat{H}_1(z, \omega) \rangle_\infty = 2\pi \int_0^\infty \hat{H}_1(\rho, z, \omega) \rho \, d\rho, \quad (24)$$

where the upper limit of infinity is not problematic because of the quickly converging Lommel functions in the integrand. With the help of Watson (Ref. 27, p. 541), Wheelon [Ref. 35, pp. 76–77, Eqs. (1.608) and (1.610)], and Euler's formula, it can be shown that

$$\langle \hat{H}_1(z, \omega) \rangle_\infty = -j \frac{\pi a^2}{k} e^{-jkz}. \quad (25)$$

This closed-form result is the same as the result reported by Williams in Ref. 18, Eq. (40) for monochromatic diffraction with a theoretically infinite receiver.

If Eq. (25) is used to calculate pressure, it yields the pressure,  $\rho \pi a^2 e^{-jkz}$ , "produced by a section of area  $\pi a^2$  cut out of a plane wave that has the same particle velocity, [in our case unity], as does the piston source (Ref. 18, p. 289)." In essence, the magnitude of the pressure detected by the infinite receiver is the same at all  $z$ -planes. No pressure/energy is lost because (i) the receiver is infinite and (ii) no loss mechanism has been introduced into the theory. Diffraction, in this case, introduces only a depth-dependent phase shift via the  $e^{-jkz}$  term.

Still more insight can be gained by examining Eqs. (21) and (25) more closely. Equation (21) can be rewritten,

$$\langle \hat{H}_1(z, \omega) \rangle_a = -j \frac{\pi a^2}{k} e^{-jkz} - \frac{\pi a^2}{k} e^{-jk[z + (a^2/z)]} \times \left[ J_1\left(\frac{ka^2}{z}\right) - jJ_0\left(\frac{ka^2}{z}\right) \right]. \quad (26)$$

Note that the first terms of Eqs. (26) and (25) are identical. Thus, the second term in Eq. (26) represents combined diffraction/finite-receiver effects which serve to modify the infinite receiver solution of Eq. (25).<sup>14,16</sup> Equation (20) can be manipulated and interpreted in a similar fashion. At any rate, the  $k = \omega/c$  in the denominator of the infinite receiver solution indicates that one-way diffraction is dominated by a  $1/f$  low-pass filtering effect. This final insight further demonstrates the theoretical and practical value of the proposed frequency-domain formalism.

#### IV. CONCLUSION

Section II demonstrated the Fourier equivalence of the arccos and Lommel diffraction formulations as an approximate Fourier transform pair. In Sec. III, we used this equivalence to propose a new closed-form frequency-domain formalism describing spatially integrated diffraction effects

from an unfocused piston transducer. The proposed frequency-domain formalism is based on the Lommel diffraction formulation and, in general, is in good agreement with results predicted by the time-domain or impulse-response formalism.

From a geometrical point of view, the time-domain formalism is superior to the proposed frequency-domain formalism because the former is based on the Rayleigh-Sommerfeld diffraction integral; thus, it is theoretically valid in the entire half-space in front of the transducer. On the other hand, the proposed frequency-domain formalism is strictly valid only in the Fresnel region. Nonetheless, the frequency-domain formalism has allowed us to compute a closed-form expression describing spatially integrated diffraction effects.

From a Fourier point of view, the Lommel-based formalism allows direct calculation of diffraction effects in the frequency domain (excluding DC) at whatever frequencies desired, and only positive coefficients need be calculated. Specifically, if a diffraction filter is to be used for inverse filtering or calibration of a real bandlimited transducer response, the filter can be readily calculated over the bandwidth of interest. If an impulse response is desired, the Lommel-based equations are problematic because they require a cumbersome estimation of DC and exhibit Gibb's phenomenon upon inverse Fourier transformation.

Computationally, the Lommel-based formalism is particularly easy to calculate for special case geometries ( $b = a$ ) because no infinite summations are required. In general though, the time-domain formalism is easier to compute than the proposed frequency-domain formalism.

Ultimately, the time-domain formalism is superior to the proposed frequency-domain formalism in terms of geometrical validity and computational ease and efficiency. Nonetheless, the proposed frequency-domain formalism is of historical, theoretical, and practical interest. Indeed, we gleaned numerous insights from the proposed frequency-domain formalism and showed that it has potential application in diffraction correction (inverse filtering) of real bandlimited transducers. More details on the proposed frequency-domain formalism can be found in Refs. 23 and 32.

#### ACKNOWLEDGMENTS

The authors thank Emil Wolf, Wilson Professor of Optical Physics at the University of Rochester, New York, for his assistance. He discussed certain aspects of this work with Captain Daly, provided references, and pointed the way to others. The elegant results he derived in the 1950's made this paper possible. The authors also thank the reviewers for their insightful comments and helpful suggestions.

<sup>1</sup>H. Seki, A. Granato, and R. Truell, "Diffraction effects in the ultrasonic field of a piston source and their importance in the accurate measurement of attenuation," *J. Acoust. Soc. Am.* **28**, 230–238 (1956).

<sup>2</sup>D. Cassereau, D. Guyomar, and M. Fink, "Time deconvolution of diffraction effects—Application of calibration and prediction of transducer waveforms," *J. Acoust. Soc. Am.* **84**, 1073–1085 (1988).

<sup>3</sup>T. Rhyne, "Radiation coupling of a disk to a plane and back or a disk to disk: An exact solution," *J. Acoust. Soc. Am.* **61**, 318–324 (1977).

<sup>4</sup>A. Gray and G. Mathews, *A Treatise on Bessel Functions and Their Ap-*

- lications to Physics* (Dover, New York, 1966). (First published in 1922.)
- <sup>5</sup>F. Oberhettinger, "On transient solutions of the "baffled piston" problem," *J. Res. Natl. Bur. Stand., Sect. B* **65**, 1–5 (1961).
  - <sup>6</sup>P. Stepanishen, "Transient radiation from pistons in an infinite planar baffle," *J. Acoust. Soc. Am.* **49**, 1627–1638 (1971).
  - <sup>7</sup>C. Daly and N. Rao, "Spatially averaged impulse response for an unfocused piston transducer," *J. Acoust. Soc. Am.* **105**, 1563–1566 (1999).
  - <sup>8</sup>X. Chen, K. Schwarz, and K. Parker, "Acoustic coupling from a focused transducer to a flat plate and back to the transducer," *J. Acoust. Soc. Am.* **95**, 3049–3054 (1994).
  - <sup>9</sup>P. Rogers and A. Van Buren, "An exact expression for the Lommel diffraction correction integral," *J. Acoust. Soc. Am.* **55**, 728–728 (1974).
  - <sup>10</sup>M. Fink and J. Cardoso, "Diffraction effects in pulse-echo measurement," *IEEE Trans. Sonics Ultrason.* **SU-31**, 313–329 (1984).
  - <sup>11</sup>A. Penttinen and M. Luukkala, "The impulse response and pressure nearfield of a curved ultrasonic radiator," *J. Phys. D* **9**, 1547–1557 (1976).
  - <sup>12</sup>R. Ziemer, W. Tranter, and D. Fanin, *Signals and Systems* (Macmillan, New York, 1983).
  - <sup>13</sup>A. Papoulis, *Systems and Transforms with Applications in Optics* (McGraw-Hill, New York, 1968), Chaps. 5 and 9.
  - <sup>14</sup>R. Bass, "Diffraction effects in the ultrasonic field of a piston source," *J. Acoust. Soc. Am.* **30**, 602–605 (1958).
  - <sup>15</sup>G. Harris, "Transient field of a baffled planar piston transducer having an arbitrary vibration amplitude distribution," *J. Acoust. Soc. Am.* **70**, 186–204 (1981).
  - <sup>16</sup>A. Williams, Jr., "The piston source at high frequencies," *J. Acoust. Soc. Am.* **23**, 1–6 (1951).
  - <sup>17</sup>G. Kino, *Acoustic Waves: Devices, Imaging, & Analog Signal Processing* (Prentice-Hall, Englewood Cliffs, NJ, 1987).
  - <sup>18</sup>A. Williams, Jr., "Integrated signal on circular piston receiver centered in a piston beam," *J. Acoust. Soc. Am.* **48**, 285–289 (1970).
  - <sup>19</sup>M. Fink, "Theoretical study of pulsed echographic focusing procedures," in *Acoustical Imaging*, edited by P. Alais and A. Metherell (Plenum, New York, 1982), Vol. 10.
  - <sup>20</sup>J. Hunt, M. Arditì, and F. Foster, "Ultrasound transducers for pulse-echo medical imaging," *IEEE Trans. Biomed. Eng.* **BME-30**, 453–481 (1983).
  - <sup>21</sup>J. Weight and A. Hayman, "Observations of the propagation of very short ultrasonic pulses and their reflection by small targets," *J. Acoust. Soc. Am.* **63**, 396–404 (1978).
  - <sup>22</sup>J. Cardoso and M. Fink, "Echographic diffraction filters and the diffraction function for random media through an instantaneous time-frequency approach," *J. Acoust. Soc. Am.* **90**, 1074–1084 (1991).
  - <sup>23</sup>C. Daly, "The Arccos and Lommel Diffraction Formulations," Ph.D. thesis, Rochester Institute of Technology, 1998.
  - <sup>24</sup>E. Wolf, "Light distribution near focus in an error-free diffraction image," *Proc. R. Soc. London, Ser. A* **204**, 533–548 (1951).
  - <sup>25</sup>G. Harris, "Review of transient field theory for a baffled planar piston," *J. Acoust. Soc. Am.* **70**, 10–19 (1981).
  - <sup>26</sup>J. Gaskill, *Linear Systems, Fourier Transforms, and Optics* (Wiley, New York, 1978).
  - <sup>27</sup>G. Watson, *A Treatise on The Theory of Bessel Functions* (Cambridge University Press, New York, 1980). (First published in 1922.)
  - <sup>28</sup>E. Madsen, M. Insana, and J. Zagzebski, "Method of data reduction for accurate determination of acoustic backscatter coefficient," *J. Acoust. Soc. Am.* **76**, 913–923 (1984).
  - <sup>29</sup>H. Huntington, A. Emslie, and V. Hughes, "Ultrasonic delay lines. I," *J. Franklin Inst.* **245**, 16–23 (1948).
  - <sup>30</sup>A. Khimunin, "Numerical calculation of the diffraction corrections for the precise measurement of ultrasound absorption," *Acustica* **27**, 173–181 (1972).
  - <sup>31</sup>G. Benson and O. Kiyohara, "Tabulation of some integral functions describing diffraction effects in the ultrasonic field of a circular piston source," *J. Acoust. Soc. Am.* **55**, 184–184 (1974).
  - <sup>32</sup>C. Daly and N. Rao, "Time- and frequency-domain descriptions of spatially averaged one-way diffraction for an unfocused piston transducer," *Ultrasonics* **37**, 209–221 (1999).
  - <sup>33</sup>E. Wolf, "The  $X_n$  and  $Y_n$  functions of Hopkins, occurring in the theory of diffraction," *J. Opt. Soc. Am.* **43**, 218 (1953).
  - <sup>34</sup>R. Kuc and D. Regula, "Diffraction effects in reflected ultrasound spectral estimates," *IEEE Trans. Biomed. Eng.* **BME-31**, 537–545 (1984).
  - <sup>35</sup>A. Wheelon, *Tables of Summable Series and Integrals Involving Bessel Functions* (Holden-Day, San Francisco, 1968).

# Directivity of a uniform-strength, continuous circular-arc source phased to the spatial position of its diameter

William Thompson, Jr.

*Applied Research Laboratory and Department of Engineering Science and Mechanics,  
The Pennsylvania State University, University Park, Pennsylvania 16802*

(Received 7 August 1998; accepted for publication 26 February 1999)

An expression is derived for the in-plane directivity function of a uniform-strength, circular-arc source (or receiver) wherein the output of each element of the arc is phase shifted by the wavelength distance from that element to the spatial position of the diameter of the circle which is parallel to the chord joining the end points of the arc. Numerical results are presented for a selected set of arc sizes and values of the radius-to-wavelength ratio. The results resemble the directional characteristics of a true straight line source of the chord length. The special cases of a phased semicircular arc and a phased complete circular arc are discussed. The directivity function of a uniform-strength, unphased, continuous circular-arc source, a previously established result, is also presented for comparison. © 1999 Acoustical Society of America. [S0001-4966(99)01006-1]

PACS numbers: 43.20.Rz, 43.30.Jx [DEC]

## INTRODUCTION

Conformal transducer arrays are commonly used in a variety of acoustic detection and localization scenarios. Cylindrical or spherical surfaces are often of interest, and a line of transducers mounted conformally on either surface could have the shape of an arc of a circle. Consequently, there is interest in the directivity patterns of arrays of this shape.

If one was computing the directivity patterns of arrays of equiamplitude point sources (or receivers) uniformly spaced along a straight line, one would naturally compare the patterns to those of a continuous, uniform strength, line array of the same length to provide a check on the validity of the results. Similarly, to provide a check on the validity of some computations of the directivity patterns of very large arrays of equiamplitude point sources uniformly spaced along a circular arc, and where the output of each point source is phase shifted by the wavelength distance from that source to the diameter of the circle which is parallel to the chord joining the end points of the arc (e.g., the  $y$ -axis of Fig. 1), it was desired to compare these patterns with those of a similarly phased, continuous, uniform-strength, circular-arc source of the same size. An expression for the in-plane directivity function of the latter source is herein derived. The directivity pattern in the orthogonal  $x$ - $z$  plane of this single arc source is not of interest here.

## I. ANALYSIS

Figure 1 shows a continuous, un baffled, circular-arc source, symmetric about the  $x$ -axis, with central angle  $2\alpha$ , radius  $R$ , and an observation point  $P$  in the plane of the source defined by coordinates  $\vec{r}$  and  $\phi$ . A differential element of the source is located at angular position  $\eta$ . Variable  $s$  is arc length along the array measured from some fixed reference point; in this case, the intersection of the  $x$ -axis with the array. The radiated pressure  $p$  at the observation point  $P$  is<sup>1</sup>

$$p(P) = B \int_{-R\alpha}^{R\alpha} \frac{e^{-jkr'}}{r'} v(s) ds. \quad (1)$$

The parameter  $B = j\rho_0 c k / 4\pi$  where  $k$  is the usual acoustic wave number ( $\omega/c$ ). The assumed harmonic time dependence factor,  $\exp(j\omega t)$ , has been suppressed throughout. The function  $v(s)$  describes the velocity distribution along the line. Two cases are considered, uniform strength with no phasing so that  $v(s) = v_0$ , a real constant, and uniform strength, where each element of the line is phased as if it were positioned along the  $y$ -axis, i.e., along the diameter of the circle which is parallel to the chord of the arc through its end points, so that  $v(s) = v_0 \exp(-jkx)$ .

Employing the usual far-field approximations, viz.,  $1/r' \approx 1/r$  for all elements and  $r' \approx r - R \cos(\phi - \eta)$  in the phase factor, and noting that  $s = R\eta$ , the radiated far-field pressure becomes

$$p(P) = \frac{BR}{r} e^{-jkr} \int_{-\alpha}^{\alpha} \exp[jkR \cos(\eta - \phi)] v(R, \eta) d\eta. \quad (2)$$

Since interest is primarily in the directivity function, the  $\phi$ -independent parameters outside the integral, as well as the magnitude  $v_0$  of the velocity distribution, are henceforth ignored.

### A. Uniform strength, no phasing

For the case  $v(R, \eta) = v_0$ , the resulting integral, designated  $I_1$ , becomes

$$I_1 = \int_{-\alpha}^{\alpha} \exp[jkR \cos(\eta - \phi)] d\eta. \quad (3)$$

This is evaluated in Appendix A by making a series expansion of the integrand and then integrating term by term. The result (the unnormalized directivity function) is

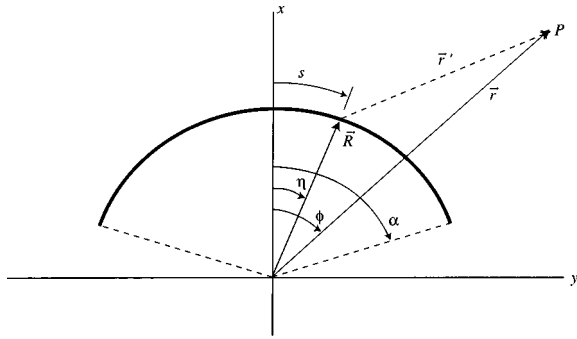


FIG. 1. Geometry of the problem.

$$I_1(\alpha, \phi) = 2 \left\{ \alpha J_0(kR) + \sum_{n=1}^{\infty} \frac{(-1)^n}{n} J_{2n}(kR) \sin(2n\alpha) \cos(2n\phi) + 2j \sum_{n=0}^{\infty} \frac{(-1)^n}{2n+1} J_{2n+1}(kR) \times \sin[(2n+1)\alpha] \cos[(2n+1)\phi] \right\} \quad (4a)$$

$$= 2 \sum_{n=0}^{\infty} (j)^n \frac{\epsilon_n}{n} J_n(kR) \sin(n\alpha) \cos(n\phi), \quad (4b)$$

where  $\epsilon_n = 1$  for  $n=0, 2$  for  $n > 0$ . This result is not new. It is alluded to, but not presented, in Ref. 2.

Two special cases are worth noting. For a uniform-strength semicircular source ( $\alpha = \pi/2$ ), the unnormalized directivity function is

$$I_1\left(\frac{\pi}{2}, \phi\right) = \pi J_0(kR) + 4j \sum_{n=0}^{\infty} \frac{J_{2n+1}(kR)}{2n+1} \cos[(2n+1)\phi], \quad (5)$$

while for a complete circular ring source ( $\alpha = \pi$ ), the response function at any far-field point in the plane of the ring is proportional to the term  $2\pi J_0(kR)$ , i.e., it is independent of the angle  $\phi$ , as is obvious. This agrees with published results.<sup>3</sup>

### B. Uniform strength, phased to y-axis

For the case  $v(R, \eta) = v_0 \exp(-jkR \cos \eta)$ , the integral in Eq. (2), denoted  $I_2$ , becomes, with the use of some trigonometric identities,

$$I_2 = \int_{-\alpha}^{\alpha} \exp[j2kR \sin(\phi/2) \sin(\eta - \phi/2)] d\eta. \quad (6)$$

This integral can be evaluated by again making a series expansion of the integrand (Appendix B). The result, the unnormalized directivity function of a uniform, continuous circular-arc source, phased to the position of the diameter, is

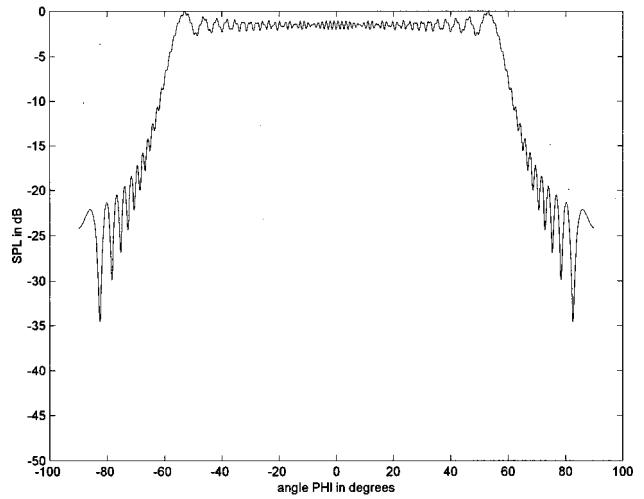


FIG. 2. Pattern of unphased, uniform strength, circular arc,  $\alpha = 60$  deg,  $kR = 300$ .

$$I_2(\alpha, \phi) = 2 \left\{ \alpha J_0[2kR \sin(\phi/2)] + \sum_{n=1}^{\infty} \frac{1}{n} J_{2n}[2kR \sin(\phi/2)] \times \sin(2n\alpha) \cos(n\phi) - 2j \sum_{n=0}^{\infty} \frac{J_{2n+1}[2kR \sin(\phi/2)]}{2n+1} \times \sin[(2n+1)\alpha] \sin[(2n+1)\phi/2] \right\} \quad (7a)$$

$$= \sum_{n=0}^{\infty} \frac{\epsilon_n}{n} J_n[2kR \sin(\phi/2)] \sin(n\alpha) \times [(-1)^n \exp(jn\phi/2) + \exp(-jn\phi/2)]. \quad (7b)$$

Again, for the special case of a semicircular arc ( $\alpha = \pi/2$ ), the result is

$$I_2\left(\frac{\pi}{2}, \phi\right) = \pi J_0[2kR \sin(\phi/2)] - 4j \sum_{n=0}^{\infty} (-1)^n \frac{J_{2n+1}[2kR \sin(\phi/2)]}{2n+1} \times \sin[(2n+1)\phi/2], \quad (8)$$

while for a complete circular source phased to the position of a diameter of that circle,

$$I_2(\pi, \phi) = 2\pi J_0[2kR \sin(\phi/2)]. \quad (9)$$

## II. RESULTS AND CONCLUSIONS

A necessary check on the analysis, and subsequent computer code, is provided by first computing directivity patterns for an unphased arc source using Eq. (4). It is known (Ref. 3) that for a uniform, unphased, circular-arc source, the magnitude of the radiation is the same on either the convex or

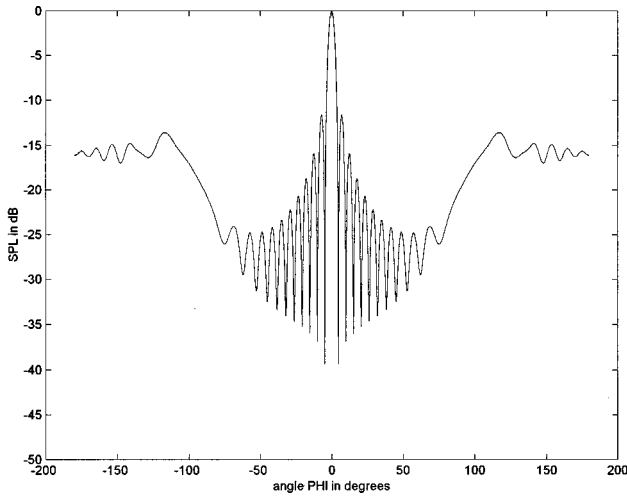


FIG. 3. Pattern of phased, uniform strength, circular arc,  $\alpha=45$  deg,  $kR=50$ .

concave side of the arc; this result is readily realized from Eq. (3), where the response at angle  $\phi - \pi$  is seen to be simply the complex conjugate of the response at angle  $\phi$ . Hence, this pattern is symmetric about the  $\pm 90$  deg directions as well as about the 0 and 180 deg directions. Furthermore, when the radius of the arc measures many wavelengths, the radiation on either side is essentially confined to, and quite uniform over, a single broad beam whose width is approximately equal to the central angle of the arc, provided that central angle is less than 180 deg. The normalized directivity pattern in Fig. 2, for the case  $\alpha=60$  deg and  $kR=300$ , clearly has these attributes, as was the case for all other  $\alpha$  and large  $kR$  values investigated. The pattern is shown only for the convex side of the arc. Note that as long as the  $kR$  value is sufficiently large, the beamwidth of the pattern of the unphased arc source is approximately the arc central angle.

Directivity patterns were then computed for the phased arc source from Eq. (7). In this case, the radiation to the concave and convex sides is different and the patterns are not symmetric about the  $\pm 90$  deg directions. Figures 3–9 are a

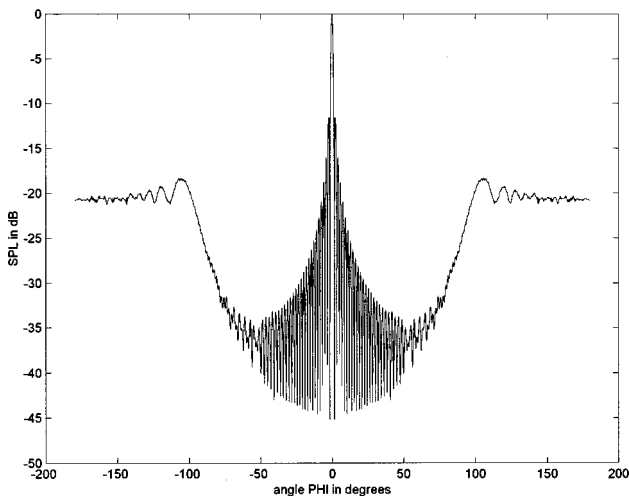


FIG. 4. Pattern of phased, uniform strength, circular arc,  $\alpha=45$  deg,  $kR=150$ .

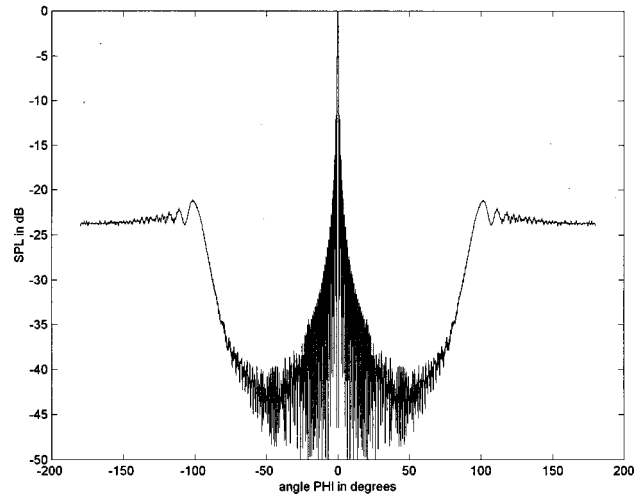


FIG. 5. Pattern of phased, uniform strength, circular arc,  $\alpha=45$  deg,  $kR=300$ .

selected set of the results. Figures 3–5 illustrate the effect of varying frequency ( $kR=50,150,300$ ) for a fixed arc size ( $\alpha=45$  deg). One sees the expected trends, *viz.*, a monotonic decrease in beamwidth with increasing frequency while the height of the first sidelobe on either side of the main beam remains constant at about  $-12.5$  dB. The envelope of the set of sidelobes has an interesting shape displaying well-defined minima near 50 deg on either side of the mainlobe. Hence, the pattern, when plotted in this linear format, has a characteristic letter ‘W’ appearance. Figures 6, 4, and 7–9 illustrate the effect of varying the size of the arc ( $\alpha=22.5,45,60,90,135$  deg) for a fixed frequency ( $kR=150$ ). While the beamwidth monotonically decreases with increasing size, the ‘W’ appearance to the pattern is only evident for  $\alpha$  less than 90 deg. For  $\alpha=90$  deg, or greater, the sidelobe envelope decreases rather monotonically with angle  $\phi$ . Also, the height of the first sidelobe on this last set of patterns increases with the angle  $\alpha$ .

It is of interest to contrast the pattern of a complete, uniform-strength, circular-arc (i.e., ring) source phased to the position of its diameter, Eq. (9), with that of a uniform-

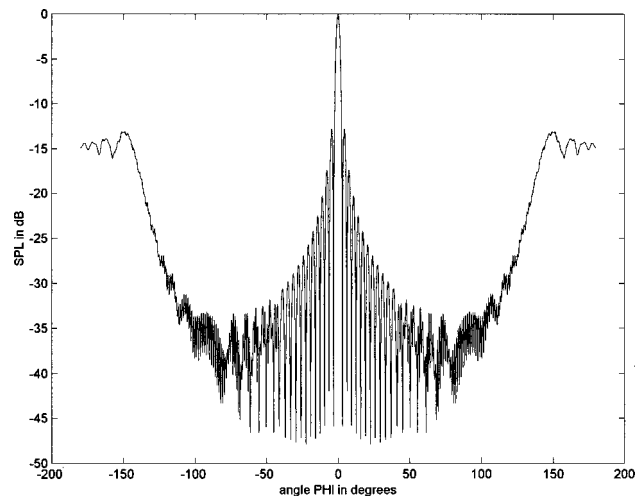


FIG. 6. Pattern of phased, uniform strength, circular arc,  $\alpha=22.5$  deg,  $kR=150$ .

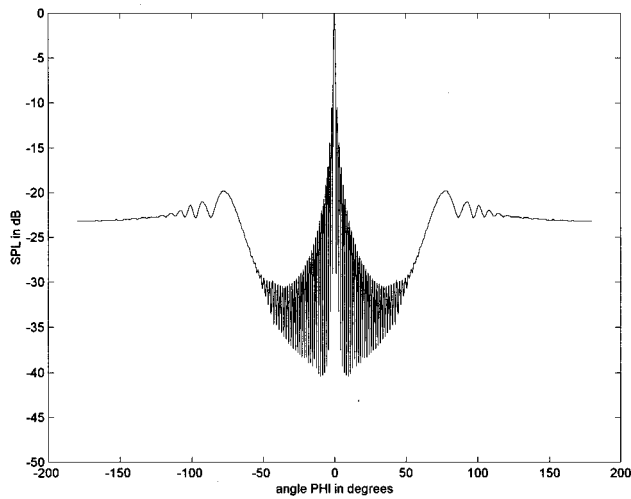


FIG. 7. Pattern of phased, uniform strength, circular arc,  $\alpha=60$  deg,  $kR=150$ .

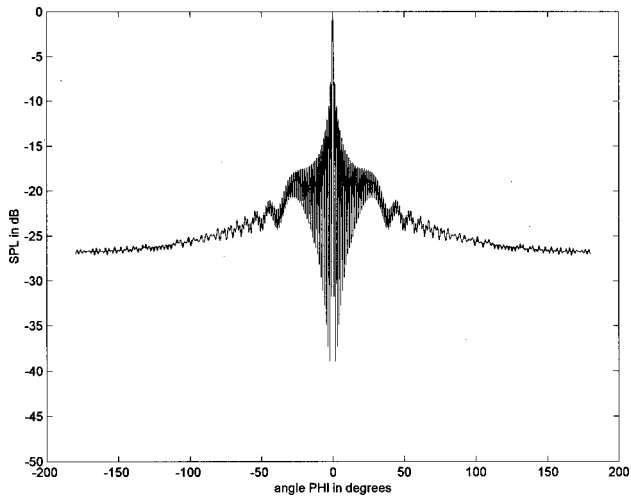


FIG. 8. Pattern of phased, uniform strength, circular arc,  $\alpha=90$  deg,  $kR=150$ .

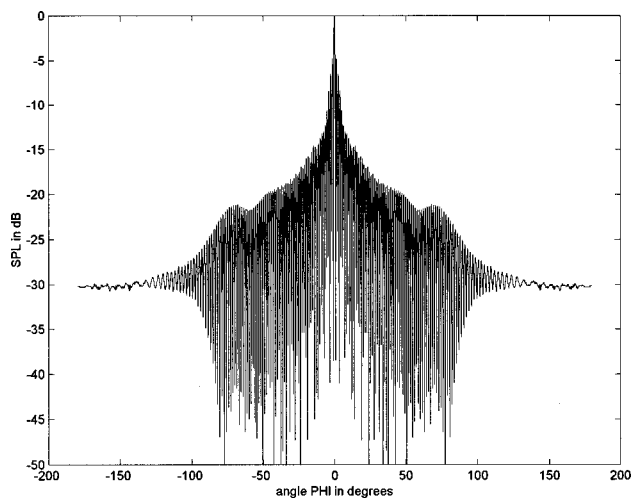


FIG. 9. Pattern of phased, uniform strength, circular arc,  $\alpha=135$  deg,  $kR=150$ .

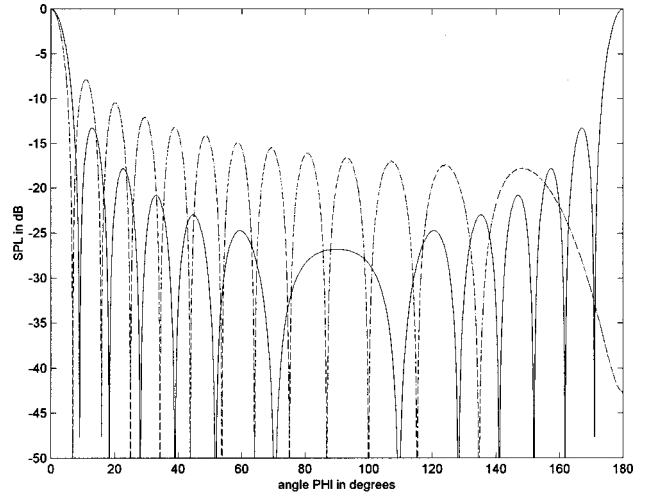


FIG. 10. Comparison of the pattern of a complete ( $\alpha=180$  deg), uniform strength, phased, circular-arc array (dashed curve) with that of a uniform-strength, straight-line array, length equal to circle diameter (solid curve),  $kR=20$ .

strength, straight line array whose length equals that diameter. One might think they should be similar results. However, further consideration of the physical shape of the phased ring source suggests that, because of its shape, it would tend to have more strength concentrated towards the ends of the diameter to which it has been phased. Hence, the effective strength is not uniform along that diameter. This inverse shading (greater strength at the ends compared to the middle) should result in an increase in the sidelobe levels with an attendant narrowing of the main beamwidth. That this effect does indeed happen is illustrated in Fig. 10 for the case  $kR=20$ . This relatively small value of  $kR$  was chosen to clearly delineate the sidelobe structure of the two directivity patterns. Also, only half of the pattern for positive values of  $\phi$  is shown. The phased ring source clearly displays higher sidelobes but a narrower beamwidth. This difference was evident for all frequencies examined. The other obvious result, which was noted earlier, is that the pattern of the phased ring is not symmetric about the  $\phi=90$  deg direction, while that of the straight line array is.

#### ACKNOWLEDGMENT

This material is based upon work supported by ASTO, through the Naval Sea Systems Command under Contract No. N00039-D-0042, Delivery Order No. 53.

#### APPENDIX A

Expanding the integrand of Eq. (3), integral  $I_1$  becomes

$$I_1 = \int_{-\alpha}^{\alpha} \cos[z \cos(\eta - \phi)] d\eta + j \int_{-\alpha}^{\alpha} \sin[z \cos(\eta - \phi)] d\eta, \quad (A1)$$

where  $z=kR$ . Employing the following standard Fourier-Bessel series expansions (Ref. 4, nos. 9.1.44 and 9.1.45)



$$\cos(z \cos \eta) = J_0(z) + 2 \sum_{n=1}^{\infty} (-1)^n J_{2n}(z) \cos(2n \eta), \quad (\text{A2})$$

and

$$\sin(z \cos \eta) = 2 \sum_{n=0}^{\infty} (-1)^n J_{2n+1}(z) \cos[(2n+1) \eta] \quad (\text{A3})$$

to replace each of the integrands in Eq. (A1), one obtains

$$\int_{-\alpha}^{\alpha} \cos[z \cos(\eta - \phi)] d\eta = 2 \left[ \alpha J_0(z) + \sum_{n=1}^{\infty} \frac{(-1)^n}{n} J_{2n}(z) \sin(2n\alpha) \cos(2n\phi) \right], \quad (\text{A4})$$

while

$$\int_{-\alpha}^{\alpha} \sin[z \cos(\eta - \phi)] d\eta = 4 \sum_{n=0}^{\infty} (-1)^n \frac{J_{2n+1}(z)}{2n+1} \times \{ \sin[(2n+1)\alpha] \cos[(2n+1)\phi] \}. \quad (\text{A5})$$

Equations (A4) and (A5), inserted into Eq. (A1), produce the result quoted in Eq. (4a).

## APPENDIX B

Expanding the integrand of Eq. (6), integral  $I_2$  becomes

$$I_2 = \int_{-\alpha}^{\alpha} \cos[z \sin(\eta - \phi/2)] d\eta + j \int_{-\alpha}^{\alpha} \sin[z \sin(\eta - \phi/2)] d\eta, \quad (\text{B1})$$

where  $z = 2kR \sin(\phi/2)$ . Employing the following standard Fourier–Bessel series expansions (Ref. 4, nos. 9.1.42 and 9.1.43)

$$\cos(z \sin \eta) = J_0(z) + 2 \sum_{n=1}^{\infty} J_{2n}(z) \cos(2n \eta) \quad (\text{B2})$$

and

$$\sin(z \sin \eta) = 2 \sum_{n=0}^{\infty} J_{2n+1}(z) \sin[(2n+1) \eta] \quad (\text{B3})$$

to replace each of the integrands above, one obtains

$$\int_{-\alpha}^{\alpha} \cos[z \sin(\eta - \phi/2)] d\eta = 2 \left[ \alpha J_0(z) + \sum_{n=1}^{\infty} \frac{J_{2n}(z)}{n} \sin(2n\alpha) \cos(n\phi) \right], \quad (\text{B4})$$

while

$$\int_{-\alpha}^{\alpha} \sin[z \sin(\eta - \phi/2)] d\eta = -4 \sum_{n=0}^{\infty} \frac{J_{2n+1}(z)}{2n+1} \sin[(2n+1)\alpha] \sin[(2n+1)\phi/2]. \quad (\text{B5})$$

Equations (B4) and (B5), inserted into Eq. (B1), produce the result quoted in Eq. (7a).

<sup>1</sup>S. Temkin, *Elements of Acoustics* (Wiley, New York, 1981), p. 286.

<sup>2</sup>I. Wolf and L. Malter, "Directional radiation of sound," *J. Acoust. Soc. Am.* **2**, 201–241 (1930).

<sup>3</sup>H. F. Olson, *Elements of Acoustical Engineering* (Van Nostrand, New York, 1947), 2nd ed., pp. 34–37.

<sup>4</sup>*Handbook of Mathematical Functions*, edited by M. Abramowitz and I. A. Stegun (Dover, New York, 1965), Chap. 9.

# A new theoretical approach to the analysis of nonlinear sound beams using the oblate spheroidal coordinate system

Tomoo Kamakura

*The University of Electro-Communications, 1-5-1, Chofugaoka, Chofu-shi, Tokyo 182-8585, Japan*

Tsuneo Ishiwata

*Oki Electric Industry Co., Ltd., 688, Ozuwa, Numazu-shi, Shizuoka 410-0873, Japan*

Kazuhisa Matsuda

*Koganei Technical High School, 6-8-9, Honmachi, Koganei-shi, Tokyo 184-0004, Japan*

(Received 16 April 1998; accepted for publication 14 February 1999)

In this article, the authors propose a new model equation which describes well progressive sound beams from a planar piston source with a circular aperture. The theory is based on the oblate spheroidal coordinate system and is easily extended to the analysis of nonlinear propagation of finite amplitude sound beams. The resultant beam equation facilitates numerical calculations in the far field. An experiment is performed in water using a 5-MHz planar ultrasonic transducer with a 9-mm radius aperture, and the results are compared with the theoretical prediction for sound pressure amplitudes of the fundamental and second harmonics. They are in excellent agreement with each other along and across the beam axis. © 1999 Acoustical Society of America. [S0001-4966(99)01106-6]

PACS numbers: 43.25.Cb [MAB]

## INTRODUCTION

Sound field analysis for directional beams reveals that there exist near-field and far-field regions which are roughly separated at the Rayleigh distance  $ka^2/2$  ( $k$ , wave number;  $a$ , source aperture radius), measured from the source along the beam axis. A simplified theoretical model for describing such beams in the near field, as well as in the far field, was obtained from the parabolic equation, which is a valid approximation of the scalar wave equation when the source aperture is much greater than the wavelength.<sup>1</sup> The parabolic approximation is essentially equivalent to the Fresnel approximation and is, therefore, a good way to predict wave behavior in the field beyond the range of  $a(ka)^{1/3}$  and in the paraxial region. Later, on the same analytical lines as the parabolic approximation, the Khokhlov-Zabolotskaya-Kuznetsov (KZK) equation, which combines efficiently three effects in beams (i.e., sound dissipation, diffraction, and nonlinearity), has been proposed. Today, the KZK equation is widely used for field analysis of the nonlinear interaction of sound beams.

The present article is concerned with another model beam equation based on an appropriate transformation of the coordinate systems. In the first section, the model equation which successfully accounts for beam spreading due to diffraction in the far field is derived from the scalar wave equation. The theory is developed in the oblate spheroidal coordinate system of the 11 separable coordinates in three dimensions, whose system is like the rectangular coordinate system near the origin and is like the spherical coordinate far from the origin.<sup>2</sup> In the next section, the beam equation is extended to a general form which includes the sound dissipation and nonlinearity in a thermoviscous fluid. In the last section, an experiment is performed in water using a 5-MHz planar transducer with a 9-mm radius aperture in order to verify the effectiveness of the present model equation. Sound

pressure amplitudes of the fundamental and second harmonics are numerically predicted and are compared with the experimental data along and across the beam axis.

## I. SPHEROIDAL BEAM EQUATION

The oblate spheroidal coordinate system  $(\sigma, \eta, \varphi)$ , which is shown in Fig. 1, may be defined in terms of the rectangular coordinate system  $(x, y, z)$  by

$$\begin{aligned}x &= b\sqrt{(1+\sigma^2)(1-\eta^2)}\cos\varphi, \\y &= b\sqrt{(1+\sigma^2)(1-\eta^2)}\sin\varphi, \\z &= b\sigma\eta\end{aligned}\quad (1)$$

with

$$-\infty < \sigma < \infty, \quad 0 \leq \eta \leq 1, \quad 0 \leq \varphi < 2\pi, \quad (2)$$

where  $b$  is the half-interfocal length.<sup>2</sup> Let us imagine that a planar sound source with circular aperture which is located at the origin  $z=0$  is radiating axisymmetric beams in the positive  $z$ -direction. In the near field the beams are presumably plane waves, and are spherically spreading due to diffraction in the far field. As can be seen in Fig. 1, the waves might be expected to propagate along the lines with  $\eta=\text{constant}$ .

In the spheroidal coordinate system, the scalar wave equation

$$\nabla^2 p - \frac{1}{c_0^2} \frac{\partial^2 p}{\partial t^2} = 0 \quad (3)$$

is transformed into

$$\begin{aligned}\frac{\partial}{\partial \sigma} \left[ (1+\sigma^2) \frac{\partial p}{\partial \sigma} \right] + \frac{\partial}{\partial \eta} \left[ (1-\eta^2) \frac{\partial p}{\partial \eta} \right] \\ - \frac{b^2(\sigma^2 + \eta^2)}{c_0^2} \frac{\partial^2 p}{\partial t^2} = 0,\end{aligned}\quad (4)$$

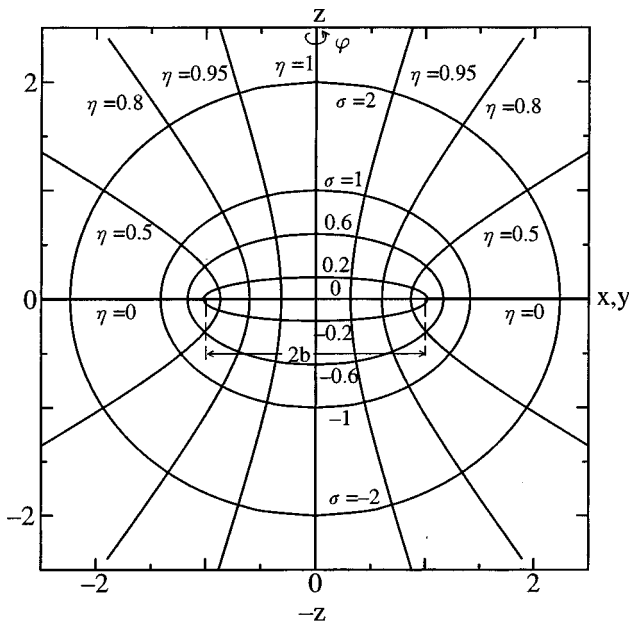


FIG. 1. Spheroidal coordinate system.

where  $p$  is the sound pressure,  $c_0$  is the sound speed,  $\nabla^2 = \partial^2/\partial x^2 + \partial^2/\partial y^2 + \partial^2/\partial z^2$ , and  $t$  is time. In order to observe progressive waves in a coordinate moving with  $c_0$ , the retarded time

$$t' = t - \frac{b\sigma\eta}{c_0} \quad (5)$$

is introduced. Substitution of Eq. (5) into Eq. (4) leads to

$$\frac{\partial}{\partial \sigma} \left[ (1 + \sigma^2) \frac{\partial \bar{p}}{\partial \sigma} \right] + \frac{\partial}{\partial \eta} \left[ (1 - \eta^2) \frac{\partial \bar{p}}{\partial \eta} \right] - 2kb\eta(1 + \sigma^2) \frac{\partial^2 \bar{p}}{\partial \sigma \partial \tau} - 2kb\sigma(1 - \eta^2) \frac{\partial^2 \bar{p}}{\partial \eta \partial \tau} = 0, \quad (6)$$

where  $\bar{p} = p/p_0$ ,  $p_0$  is the pressure amplitude at the source,  $\tau = \omega t'$ ,  $\omega$  is the angular frequency of sound, and  $k = \omega/c_0$  is the wave number. For radiation of high frequency waves, the first term on the left-hand side of Eq. (6) is discarded, and the following beam equation is obtained:

$$\eta \frac{\partial^2 \bar{p}}{\partial \tau \partial \sigma} - \frac{1}{2kb} \frac{1}{1 + \sigma^2} \frac{\partial}{\partial \eta} \left[ (1 - \eta^2) \frac{\partial \bar{p}}{\partial \eta} \right] + \frac{\sigma}{1 + \sigma^2} (1 - \eta^2) \frac{\partial^2 \bar{p}}{\partial \tau \partial \eta} = 0. \quad (7)$$

Instead of the variable  $\eta$ ,  $\theta = \cos^{-1} \eta$  is appropriate for a description of source directivity, because the surface of  $\eta = \text{constant}$ , passing through the origin, denotes an inclination at the angle  $\theta$  to the  $z$ -axis in the far field. Then, Eq. (7) takes the form

$$\cos \theta \frac{\partial^2 \bar{p}}{\partial \tau \partial \sigma} - \frac{\epsilon}{1 + \sigma^2} \left( \frac{\partial^2 \bar{p}}{\partial \theta^2} + \cot \theta \frac{\partial \bar{p}}{\partial \theta} \right) - \frac{\sigma}{1 + \sigma^2} \sin \theta \frac{\partial^2 \bar{p}}{\partial \tau \partial \theta} = 0, \quad (8)$$

where  $\epsilon = 1/2kb$  is much smaller than unity, as will be clarified just below. In this spheroidal beam equation, the axial variable  $\sigma$  and the length parameter  $b$  still remain undetermined. The next problem is to determine  $\sigma$  and  $b$  to meet actual source conditions.

Using Eq. (1) and the relation of  $\eta = \cos \theta$ , the axial distance from the source and radial distance from the axis are given by

$$z = b\sigma \cos \theta, \quad (9)$$

$$r = \sqrt{x^2 + y^2} = b\sqrt{1 + \sigma^2} \sin \theta. \quad (10)$$

In the same manner as the normalized KZK equation developed by Aanonsen *et al.*,  $\sigma$  is defined as the ratio of  $z$  to the Rayleigh length  $R_d = ka^2/2$ .<sup>3</sup> Consequently  $b$  is reduced from Eq. (9) to

$$b = \frac{z}{\sigma \cos \theta} = R_d \quad (\sigma = 1, \theta = 0). \quad (11)$$

On the source surface, the aperture imposes a condition on the angle from Eq. (10), i.e.,

$$\sin \theta_0 = \frac{a}{b} = \frac{a}{R_d} = \frac{2}{ka} \quad (\sigma = 0). \quad (12)$$

Substituting Eq. (12) into the expression of  $\epsilon$ , we obtain

$$\epsilon = \frac{1}{2kb} = \frac{1}{(ka)^2}. \quad (13)$$

Since the aperture is generally much larger than the wavelength in actual situations encountered during our experiments,  $ka \gg 1$  and  $\epsilon \ll 1$  is always satisfied.

## II. EXECUTION OF NUMERICAL COMPUTATION

We now suppose a planar sound source with uniform amplitude distribution to be radiating a sinusoidal ultrasound, for which the prescribed boundary conditions at  $\sigma = 0$  are

$$\bar{p} = \begin{cases} \sin \tau & (0 \leq \theta \leq \theta_0) \\ 0 & (\theta > \theta_0). \end{cases} \quad (14)$$

A steady-state solution of Eq. (8) is sought in the form

$$\bar{p} = g(\sigma, \theta) \sin \tau + h(\sigma, \theta) \cos \tau \quad (15)$$

and the resultant partial differential equations for spectral components  $g$  and  $h$  are expressed by

$$\begin{aligned} \cos \theta \frac{\partial g}{\partial \sigma} - \frac{\epsilon}{1 + \sigma^2} \left( \frac{\partial^2 h}{\partial \theta^2} + \cot \theta \frac{\partial h}{\partial \theta} \right) - \frac{\sigma}{1 + \sigma^2} \sin \theta \frac{\partial g}{\partial \theta} &= 0, \\ \cos \theta \frac{\partial h}{\partial \sigma} + \frac{\epsilon}{1 + \sigma^2} \left( \frac{\partial^2 g}{\partial \theta^2} + \cot \theta \frac{\partial g}{\partial \theta} \right) - \frac{\sigma}{1 + \sigma^2} \sin \theta \frac{\partial h}{\partial \theta} &= 0. \end{aligned} \quad (16)$$

The boundary conditions are written in terms of  $g$  and  $h$  from Eq. (14):

$$g = \begin{cases} 1 & (0 \leq \theta \leq \theta_0) \\ 0 & (\theta > \theta_0), \end{cases} \quad h = 0. \quad (17)$$

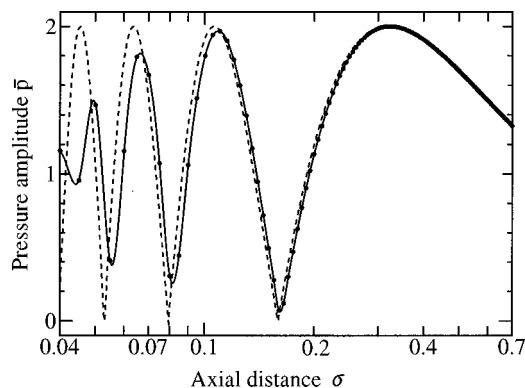


FIG. 2. On-axis pressure amplitudes.  $\omega/2\pi=2$  MHz,  $a=1$  cm, and  $c_0=1500$  m/s. Solid curve and dots denote the numerical solutions based on Eq. (8) and Eq. (18), respectively. Dashed curve is the analytical solution of Eq. (18).

Using a finite difference method, Eqs. (16) can be solved numerically subject to the boundary conditions of Eqs. (17). The main procedure for the solution follows, in principle, the lines of the original version of Aanonsen *et al.*<sup>3</sup> The pressure field  $\bar{p}(\sigma, \theta)$  is discretized in space by small step sizes of  $\Delta\sigma$  and  $\Delta\theta$  between grids, and is integrated with respect to  $\theta$  to advance the field through each incremental step from  $\sigma$  to  $\sigma+\Delta\sigma$ . When a general iteration scheme such as the Gauss–Seidel method is used for the solution of large linear systems,  $\Delta\sigma$  is allowed to be increased by the factor  $1+\sigma^2$ . Therefore the incremental step size is smaller in the range of  $\sigma < 1$  than the size proposed by Aanonsen *et al.* in the transformed beam equation, in which it is increased by the factor  $(1+\sigma)^2$ .<sup>4</sup> However, in the far field  $\sigma \gg 1$  both step sizes are closely increased proportional to  $\sigma^2$ .

Even if an integration range in the angle region is uniformly divided into a small step size  $\Delta\theta$ , the step size in the actual radial region  $\Delta r$  is not uniform owing to the nonlinear relation  $r=b\sqrt{1+\sigma^2}\sin\theta$ . This relation results in finer discretization in a large angle region. However, the size may be almost uniform so long as the integration range is limited in a paraxial region; i.e.,  $\theta_0$  is much smaller than  $90^\circ$ .

Numerical computation has been executed to validate the present model equation. Assuming ultrasounds travel down in water, input parameters to a computer are assigned as follows:  $\omega/2\pi=2$  MHz,  $a=10$  mm,  $c_0=1500$  m/s. These parameters determine the Rayleigh length 42 cm. Axial pressure amplitudes are shown in Fig. 2 in solid curves. For comparison, computed results are plotted in dots using the ordinary parabolic beam equation which is written

$$\frac{\partial^2 \bar{p}}{\partial \tau \partial \sigma} - \frac{1}{4} \left( \frac{\partial^2 \bar{p}}{\partial \xi^2} + \frac{1}{\xi} \frac{\partial \bar{p}}{\partial \xi} \right) = 0, \quad (18)$$

where  $\xi=r/a$ .<sup>1</sup> The dashed curve in the figure is the analytical solution of Eq. (18);  $2|\sin(1/2\sigma)|$ . Discretization conditions for both beam equations are almost the same; the integration ranges in the radial direction are extended up to  $7\theta_0$  for Eq. (8) and up to 7 for Eq. (15), and are divided into 400 equally spaced grids. The initial step sizes for  $\sigma$  are equal to  $3 \times 10^{-3}$ . In the far field, however, the size in Eq. (8) gets coarser as  $\sigma$  increases in comparison with the size in Eq. (18). Figure 2 shows very good agreement in magnitudes

between the solid curve and dots in the field structure, in particular, in the near-field oscillations. This result supports the validity of the present model equation for describing high frequency sound beams from a planar source within at least the same accuracy as the ordinary parabolic equation.

### III. EXTENSION TO THE NONLINEAR BEAM

When sound amplitude becomes finite, the waveform distorts with propagation due to the inherent nonlinearity of the medium. In order to predict theoretically the nonlinear behavior of sound beams, the following wave equation is obtained from the general equations of fluid dynamics:

$$\nabla^2 p - \frac{1}{c_0^2} \frac{\partial^2 p}{\partial t^2} + \frac{\delta}{c_0^4} \frac{\partial^3 p}{\partial t^3} = - \frac{\beta}{\rho_0 c_0^4} \frac{\partial^2 p^2}{\partial t^2}, \quad (19)$$

where  $\delta$  is the sound diffusivity and  $\beta$  is the coefficient of nonlinearity.<sup>1</sup> Based on a similar approximation as in the preceding section, Eq. (19) is reduced to the nonlinear beam equation

$$\cos\theta \frac{\partial^2 \bar{p}}{\partial \sigma \partial \tau} - \frac{\epsilon}{1+\sigma^2} \left( \frac{\partial^2 \bar{p}}{\partial \theta^2} + \cot\theta \frac{\partial \bar{p}}{\partial \theta} \right) - \frac{\sigma}{1+\sigma^2} \sin\theta \frac{\partial^2 \bar{p}}{\partial \theta \partial \tau} - \alpha b \frac{\sigma^2 + \cos^2\theta}{1+\sigma^2} \frac{\partial^3 \bar{p}}{\partial \tau^3} = \frac{\beta \rho_0 b \omega}{2\rho_0 c_0^3} \frac{\sigma^2 + \cos^2\theta}{1+\sigma^2} \frac{\partial^2 \bar{p}^2}{\partial \tau^2}, \quad (20)$$

where  $\rho_0$  is the medium density and  $\alpha$  is the thermoviscous attenuation coefficient which is related to  $\delta$  as  $\alpha = \delta\omega^2/2c_0^3$ .

Using a commercially available 5-MHz planar transducer with a 9-mm radius aperture, sound pressure was measured in fresh water. A miniature PVDF hydrophone with an active diameter 0.6 mm, which is calibrated within an accuracy of  $\pm 1$  dB from 2 to 10 MHz, detects locally the nonlinearly generated high frequency components in beams. Details of the experimental setup were presented in a previous article.<sup>5</sup>

Figure 3 shows on-axis sound pressure amplitudes for the fundamental and second harmonic components, in which these spectral components correspond to 5-MHz and 10-MHz sound waves. Circles in the figure are all the experi-

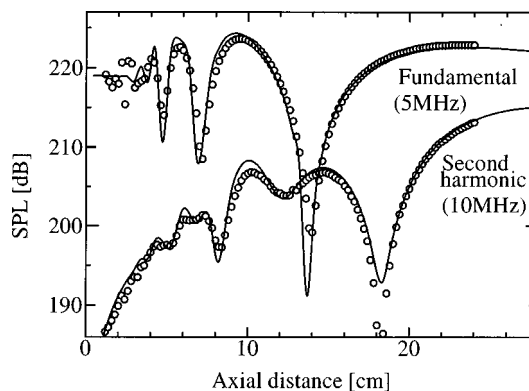


FIG. 3. On-axis pressure amplitudes of the fundamental and second harmonics in water.  $\omega/2\pi=5$  MHz,  $a=9$  mm, and  $p_0=126$  kPa. Solid curves are the numerical solutions based on Eq. (20), and symbols are the experimental data.

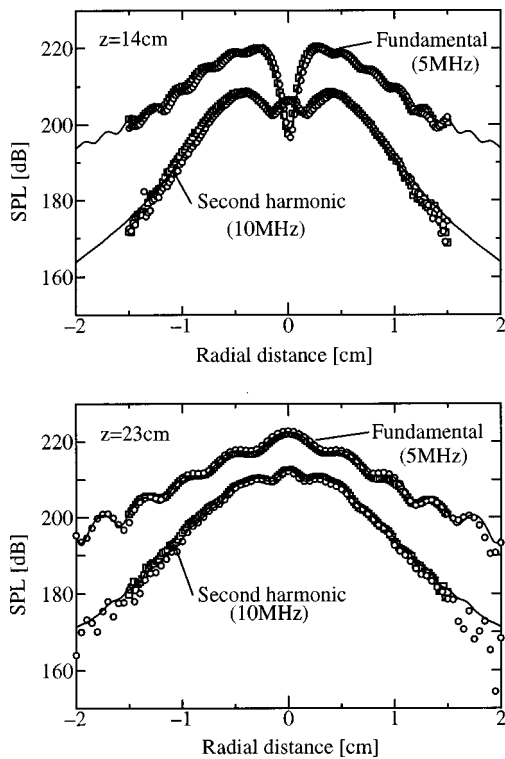


FIG. 4. Beam patterns of the fundamental and second harmonic components at  $z = 14$  cm (top) and  $z = 23$  cm (bottom). Source conditions are the same as in Fig. 3. Solid curves are the numerical solutions based on Eq. (20) and symbols are the measured data.

mental data, and solid curves are the numerically computed results based on the spheroidal beam equation. In terms of a Fourier expansion

$$\bar{p} = \sum_{n=1}^{\infty} (g_n \sin n\tau + h_n \cos n\tau), \quad (21)$$

a solution of Eq. (20) is obtained via the finite difference method. The first seven spectral components are retained, and the initial sound pressure  $p_0$  is set to 126 kPa. Numerical conditions such as step sizes are the same as in Fig. 2. Agreement of theory and experiment is excellent from near the source to the neighborhood of the last peaks of the pressure. Other comparisons of theory and experiment are given in

Fig. 4 for the beam patterns on the radial planes perpendicular to the axis. The distances from the source to the planes are 14 cm, where the last dip of the fundamental wave is found, and 23 cm, where the amplitude of the fundamental almost attains a peak. In the figures' symbols, circles and squares stand for the data on the  $x$ -axis and the  $y$ -axis, respectively. The experimental fact that measured data on the different axes are almost the same magnitudes verifies the beam to be axisymmetric. The theoretical prediction, which is designated by solid curves, agrees fairly well with experiments in the paraxial region. Thus we can confirm that the model equation given by Eq. (20) predicts quantitatively the harmonic generation of nonlinear sound beams within the framework of the present experimental conditions.

#### IV. CONCLUSION

The spheroidal beam equation which predicts well sound pressure fields near and far from the source has been discussed. The theory was justified within the same accuracy as the ordinary parabolic theory, and was easily extended to the analysis of nonlinear harmonic generation in finite amplitude sound beams. The analysis is limited to axisymmetric fields. However, the model equation which accommodates the spherically spreading sound waves is more amenable to numerical calculations up to the far field just like the transformed beam equation. An application of the present theory to the analysis of nonlinear sound beams emitted from a focusing source with wide aperture angle is now under way.

<sup>1</sup>M. F. Hamilton and C. L. Morfey, "Model equations," in *Nonlinear Acoustics*, edited by M. F. Hamilton and D. T. Blackstock (Academic, CA, 1998), Chap. 3.

<sup>2</sup>C. Flammer, *Spheroidal Wave Functions* (Stanford University Press, Stanford, CA, 1957).

<sup>3</sup>S. I. Aanonsen, T. Barkve, J. Naze Tjøtta, and S. Tjøtta, "Distortion and harmonic generation in the nearfield of a finite amplitude sound beam," *J. Acoust. Soc. Am.* **75**, 749–768 (1984).

<sup>4</sup>S. I. Aanonsen, M. F. Hamilton, J. Naze Tjøtta, and S. Tjøtta, "Nonlinear effects in sound beams," in *Proceedings of the 10th International Symposium on Nonlinear Acoustics*, edited by Akira Nakamura (Teikosha Press, Kadoma, Japan, 1984), pp. 45–48.

<sup>5</sup>T. Kamakura, T. Sudo, K. Matsuda, and Y. Kumamoto, "Time evolution of acoustic streaming from a planar ultrasound source," *J. Acoust. Soc. Am.* **100**, 132–138 (1996).

# Nonlinear pulse propagation in arbitrarily dispersive media: Tube waves in permeable formations

David Linton Johnson

*Schlumberger-Doll Research, Old Quarry Road, Ridgefield, Connecticut 06877-4108*

(Received 21 September 1998; accepted for publication 5 March 1999)

An approximate quasistatic equation, analogous to the Burgers equation, is derived to account for the combined effects on tube wave propagation of (a) dispersion/attenuation in permeable formations, and (b) quadratic nonlinearity of the fluid and of the formation. Numerical results for weak nonlinearity and narrow-band pulses indicate that pulse self-demodulation does occur, but over relatively large distances because of the relatively low-frequency band relevant for tube wave propagation in characteristic borehole geometries ( $f < 10$  kHz). The self-demodulated pulse shape can be very significantly distorted from that predicted by the conventional Burgers equation, depending upon the choice of relevant parameters such as the permeability, the carrier frequency, and the mudcake membrane stiffness. Numerically exact analytical formulas for the self-demodulated pulse shape, as well as for the energy in the second harmonic band, are derived for cases in which the pulse duration is long and the nonlinearity is relatively weak. These formulas are valid for any arbitrary dispersion/attenuation mechanism, and not just tube waves in permeable formations, as long as the propagation wave vector may be specified uniquely as a function of frequency. © 1999 Acoustical Society of America. [S0001-4966(99)03006-4]

PACS numbers: 43.25.Dc, 43.25.Ed [MFH]

## INTRODUCTION

A tube wave (often misnamed as a Stoneley wave) is a normal mode of propagation confined to the vicinity of a fluid-filled cylindrical borehole in an elastic solid. As a practical consideration, it is generally the dominant signal which appears in a typical borehole-logging measurement and thus it is important in a variety of contexts in the search for hydrocarbon sources. In a previous article (Johnson *et al.*, 1994, hereafter referred to as JKN), the nonlinear characteristics of a tube wave were analyzed in terms of the nonlinear third-order elastic constants of the solid and the nonlinearity of the fluid. The approximations invoked were: (1) the quasistatic limit, and (2) absence of attenuation/dispersion. The first of these merely limits the frequency range to a few kHz or below, for typical oil-field boreholes. This is not much of a restriction. The second is more restrictive inasmuch as it is known that tube waves are generally very attenuative, especially if the formation is permeable to fluid flow. This means that although one could calculate, in JKN, the initial growth of a second harmonic tube wave, or the pressure dependence of the speed of the tube wave one could not realistically calculate the large distance evolution of a tube wave. The formation of a shock discontinuity, or even the evolution of the second harmonic, would be severely affected by attenuation, and an effect such as self-demodulation of a pulse would be impossible to calculate at all. It is known that the dominant mechanism for attenuation (and the concomitant dispersion) that occurs in permeable formations is due to the fact that a pressure wave in the borehole causes the fluid to flow through the pores of the surrounding rock. This effect is well described by the Biot theory (see Sec. I A, below), and depends upon formation permeability and fluid viscosity. The purpose of the present article is to derive an approximate quasistatic equation of motion incorporating both nonlinearity as well as attenuation/dispersion due to this Biot mecha-

nism, and to investigate the consequences of these combined mechanisms on the propagation of tube wave pulses. The formalism, many of the analytical results, and the numerical technique are, however, applicable to any dispersion/attenuation mechanism. For the purposes of the present article, only amplitude-modulated pulses are considered.

It is convenient to do acoustic experiments using well-defined pulses as opposed to a continuous sine wave. In the present paper, we focus on the evolution of narrow-band (long duration) amplitude-modulated pulses for which the nonlinearity is relatively weak. Specifically, in the presence of nonlinearity a pulse will evolve second-harmonic components as well as low-frequency components; the latter is related to the shape of the envelope function. In the presence of attenuation, the second-harmonic components increase with propagation distance as long as the carrier component is still strong; then they will die away due to the attenuation mechanism. Similarly, the low-frequency components grow but they eventually outlive the components near the carrier frequency because they suffer much less attenuation. This effect, in which the nonlinearly generated low-frequency components outlive the original components near the carrier frequency, is pulse self-demodulation, an effect which is well understood in unbounded fluids in three dimensions (Averkiou *et al.*, 1993, and references therein). There are two major differences between the two cases: (1) In three-dimensional fluids, pulse self-demodulation typically can occur only in the far field; therefore, diffraction effects are essential to the understanding. In the present paper, the effect is strictly one-dimensional and this represents a considerable simplification. (2) In fluids, there is negligible dispersion and an attenuation which is proportional to frequency squared, and thus can be expressed simply in the time domain. In tube waves in permeable formations, both attenuation and dispersion are large and complicated functions of frequency. This

represents a significant complication, relative to unbounded fluids. These two differences also apply to second-harmonic generation or, indeed, to any aspect of nonlinear wave propagation.

In rocks, it is well-known that the third-order elastic constants are orders of magnitude larger than those in non-porous media (Winkler and Liu, 1996, and references therein). This fact argues in favor of enhanced nonlinear effects. On the other hand, typical frequencies are much lower than laboratory-based ultrasonic investigations; this fact argues in favor of diminished nonlinear effects. On the third hand, attenuation is large, which would argue that self-demodulation can be made to occur in a short distance, simply by increasing the amplitude. A purpose of the present article is to investigate, numerically, all of these combined effects. Although the article is theoretical, numerical examples are given for real-life borehole situations that occur in the oil-field industry so as to get some sense as to the size of the effects that can be expected. This sets the radius of the borehole at around 10 cm, the permeability at around 1 Darcy ( $=1 \mu\text{m}^2$ ), and the frequencies around 1–10 kHz, although obviously one could run the calculations for parameter sets appropriate to a scaled-down laboratory experiment. As a by-product, we derive a simple analytical formula for the self-demodulated signal which is valid for an arbitrary one-dimensional amplitude-modulated pulse, with an arbitrary dispersion/attenuation mechanism. Likewise, a formula for the energy contained within the second harmonic band, considered as a function of propagation distance, is derived analytically. Specific application, though, is made to the case at hand of tube wave attenuation/dispersion in permeable formations.

The article is organized as follows: The theory is developed in Sec. I, part A therein is a review of the Biot mechanism of tube wave attenuation and dispersion in permeable formations, whereas part B is a review of the nonlinear theory, in the absence of attenuation/dispersion. A generalized Burgers equation that incorporates these two effects is derived in part C and the method of numerical calculation is described in part D. Section II is devoted to the presentation of the numerical results for a few cases of interest. The conclusions of the article are summarized in Sec. III.

## I. THEORY

In this section, we derive an approximate equation of motion which governs the nonlinear propagation of a tube wave in the presence of attenuation and dispersion which, in turn, is caused by the fluid-flow permeability of the formation. The quasistatic/low-frequency limit is assumed; this means that the relevant wavelengths are all large compared to the borehole radius. In the first subsection, the results of the linear theory are reviewed; in the second subsection we review the results of the nonlinear theory in an impermeable elastic formation. Next, the two effects are considered simultaneously and an equation of motion analogous to the Burgers equation (see Hamilton and Blackstock, 1998) is derived. Finally, we give a brief description of the numerical method used to solve the nonlinear equation of motion.

### A. Permeability effects on tube waves: Linear theory

The geometry is idealized in terms of a homogeneous, isotropic formation for which the acoustic properties are described by the Biot theory of acoustics in porous media. There is a cylindrical borehole of radius  $b$ . Lining the borehole wall there is a mudcake, of thickness  $b-a$ , which is taken to be an elastic material. The remainder of the borehole is filled with an acoustic fluid. This model is described in detail by Liu and Johnson (1997). In the special case where there is no mudcake, this theoretical model had previously been shown to give an accurate description of experimentally observed tube wave dispersion and attenuation by Winkler *et al.* (1989). The dispersion and attenuation is due primarily to the ability of the tube wave to couple with the so-called slow acoustic wave in the formation, a mode whose properties are very much dependent upon the permeability of the formation. The following is a short synopsis of the relevant parts of Liu and Johnson (1997).

Tube wave characteristics (in the linear regime) are described by the complex-valued axial wave vector,  $k_z(\omega)$ ; its value is determined by the condition that the characteristic determinant, of size  $8 \times 8$ , vanishes. Results for reasonably chosen parameter values are given in Liu and Johnson (1997). The phase slowness,  $S_p(\omega)$ , the attenuation,  $\gamma(\omega)$ , and the specific attenuation,  $1/Q(\omega)$ , are related to the real and imaginary parts of  $k_z$  by

$$k_z(\omega) = \omega S_p(\omega) + i\gamma(\omega) = \omega S_p(\omega)[1 + i/2Q(\omega)]. \quad (1)$$

In general,  $k_z$  is determined from a numerical search of the aforementioned determinant, which is complicated. In the quasistatic limit, however, a simple analytical expression can be derived

$$k_z^2(\omega) \doteq \omega^2 \rho_{f0} \left[ \frac{1}{K_{f0}} + \frac{1}{M_F} + \frac{1}{[W_{mc} + W_p(\omega)] b^*} \right]. \quad (2)$$

Here,  $\rho_{f0}$ , and  $K_{f0}$  are the density and the bulk modulus, respectively, of the borehole fluid, and  $M_F$  is an effective shear modulus of the formation, corrected by mudcake effects. The mudcake membrane stiffness,  $W_{mc}$  describes the ease with which the mudcake can flex in and out of the surface pores of the formation: The special case  $W_{mc}=0$  corresponds to perfect ease of flexing and may be called the ‘‘open borehole’’ case, whereas  $W_{mc} \rightarrow \infty$  corresponds to a perfect sealing and may be called the ‘‘sealed borehole’’ case. For the purposes of the present article,  $W_{mc}$  is taken as an input parameter. There is, however, an interesting article by Tang (1994) in which he considers an elastic solid pressed against a porous medium consisting of holes of radius  $R$ . From his result, it is possible to show (Liu and Johnson, 1997) that reasonable values for  $W_{mc}$  lie in the range 1–10 GPa/cm for permeable sedimentary rocks. Indeed, comparison of the predictions of the full theory against measured field data indicates  $W_{mc} \approx 2.5$  GPa/cm (Pampuri *et al.*, 1998).

Permeability effects are contained within the factor  $W_p$ , which is given by

$$W_p(\omega) = -\frac{\eta C_D k_{SI} H_0^{(1)}(k_{SI} b)}{\kappa_0 H_1^{(1)}(k_{SI} b)}, \quad (3)$$

where  $\eta$  is the viscosity of the pore fluid,  $\kappa_0$  is the permeability of the formation, and  $H_{0,1}^{(1)}$  are Hankel functions. In this low-frequency regime, the slow wave is diffusive in character; the wave vector of the slow wave,  $k_{SI}$ , is related to its diffusivity,  $C_D$ , by

$$k_{SI}(\omega) = \sqrt{i\omega/C_D}. \quad (4)$$

The diffusivity, in turn, is related to the porosity,  $\phi$ , and the effective pore-fluid bulk modulus,  $K_f^*$ , by

$$C_D = \frac{\kappa_0 K_f^*}{\eta \phi}. \quad (5)$$

The parameter  $b^*$  is the borehole radius, slightly renormalized by mudcake elastic properties. See Liu and Johnson (1997) for details. For the purposes of the present article, we will assume  $b^* = b$  and  $K_f^* = K_f$ .

It is convenient to rewrite Eq. (2) in an obvious manner as

$$k_z^2 = \omega^2 [S_\infty^2 + 2S_\infty \tilde{\Theta}(\omega)], \quad (6)$$

where

$$S_\infty = \sqrt{\rho_{f0} \left[ \frac{1}{K_{f0}} + \frac{1}{M_F} \right]} \quad (7)$$

is the phase slowness of the tube wave predicted by Eq. (2) in the limit of high frequencies, and where

$$\tilde{\Theta}(\omega) = \frac{\rho_{f0}}{S_\infty b [W_{mc} + W_p(\omega)]}. \quad (8)$$

It is useful to use the approximation to Eq. (6)

$$k_z \doteq \omega [S_\infty + \tilde{\Theta}(\omega)], \quad (9)$$

which holds when the second term in Eq. (6) is small compared to the first. Written this way, it is clear that  $\tilde{\Theta}$  is a kind of additional complex-valued slowness due to permeability effects. Figure 1 shows the additional slowness,  $\Delta S_p = \Re[\tilde{\Theta}]$ , and the specific attenuation,  $1/Q = 2\Im[\tilde{\Theta}]/[S_\infty + \Delta S_p]$  predicted by this model for the parameter sets given by Table I, which will be used in the calculations presented later. [In Table I, the parameters  $S_\infty$ ,  $\eta$ ,  $\rho_{f0}$ , and  $K_f^*$  are assigned values appropriate to water, and the borehole radius is chosen to be typical of real oil-field values. The permeability values are characteristic of highly productive oil reservoirs ( $1 \text{ Darcy} = 1 \mu\text{m}^2$ ); the values of the membrane stiffness,  $W_{mc}$  are approximately those observed empirically by comparing the predictions of the full linear theory of tube wave characteristics against real measurements in oil-field boreholes (Pampuri *et al.*, 1998).] There are three observations worth making: (a)  $\tilde{\Theta}(\omega)$  depends upon permeability and frequency only through the ratio,  $\omega/\kappa_0$ , and this is immediately apparent in those two cases shown in Fig. 1 which differ only in the assumed value of the permeability. (b) Because  $\lim_{\omega \rightarrow 0} [W_p(\omega)] = 0$  and  $\lim_{\omega \rightarrow \infty} [W_p(\omega)] = \infty$ , there is a peak in  $1/Q$  [an inflection point in  $\Delta S_p(\omega)$ ] which

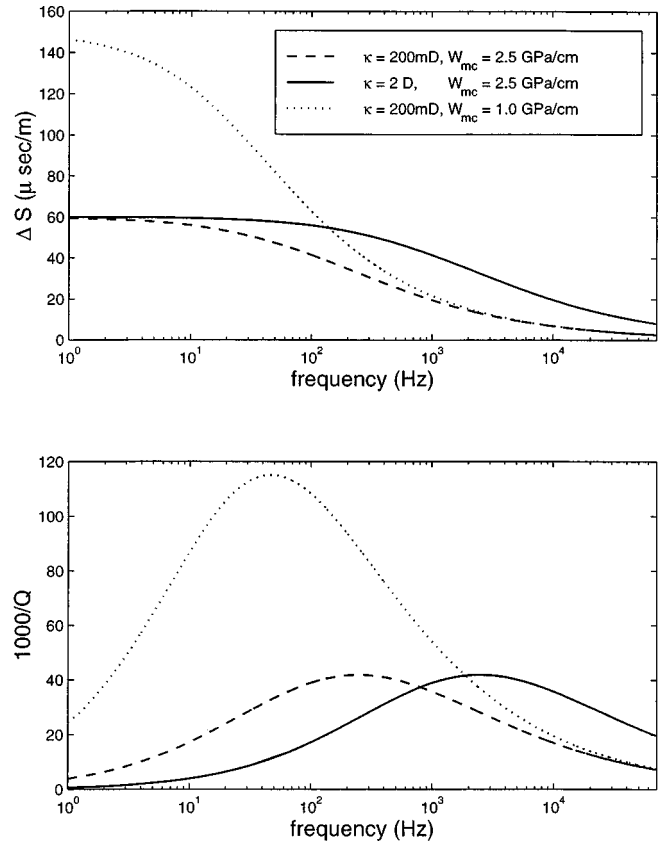


FIG. 1. Tube wave characteristics in a permeable formation calculated from Eqs. (8) and (9). Additional slowness (top) and specific attenuation (bottom). Parameter values not indicated in the figure are taken from Table I.

occurs when  $|W_p(\omega)| \approx W_{mc}$ . (c) The effects of permeability upon tube wave characteristics are increased and shifted to lower frequencies as the membrane-stiffness parameter,  $W_{mc}$ , is decreased, as seems intuitively obvious. In the limit that  $W_{mc} \approx 0$ , the slowness diverges and the specific attenuation attains extremely large values. In this case, the approxi-

TABLE I. Values of input parameters for the calculation of tube wave characteristics.

Sample	A	B	C
$\phi$	0.30	0.30	0.30
$\eta$ (poise)	0.010	0.010	0.010
$K_f^*$ (GPa)	2.25	2.25	2.25
$b$ (cm)	10	10	10
$\rho_f$ (gm/cm <sup>3</sup> )	1.0	1.0	1.0
Input parameters			
$S_\infty$ ( $\mu\text{s/m}$ )	667	667	667
$\beta$	50.5	50.5	50.5
$W_{mc}$ (GPa/cm)	2.5	2.5	1.0
$\kappa_0$ ( $\mu\text{m}^2$ )	2.0	0.200	0.200
$\omega_1/2\pi$ (Hz)	$10^2$	$10^4$	$10^4$
$T_W$ ( $10^{-3}$ s)	125	1.25	1.25
$P_0$ ( $10^4$ Pa)	8.1	8.0	9.2
Calculated quantities			
$C$ (m <sup>2</sup> /s)	15	1.5	1.5
$Z_{att}$ (m)	278	2.80	2.4
$Z_{shock}$ (m)	1323	13.4	11.6
$\Gamma = Z_{att}/Z_{shock}$	0.21	0.21	0.21
$\Delta S_p(\omega_1)$ ( $\mu\text{s/m}$ )	55.9	6.8	6.9
$\Delta S_g(\omega_1)$ ( $\mu\text{s/m}$ )	52.7	3.5	3.5



mation (9) may not hold at low frequencies, even though Eq. (6) is still valid.

## B. Tube waves in a nonlinear elastic formation

Here, the formation is considered to be hyperelastic and impermeable. In the quasistatic limit, and keeping all terms through quadratic nonlinearity, one can derive the following equation of motion for tube wave propagation [JKN: Johnson *et al.* (1994), Eq. (25)]:

$$p_{,zz} - S_{\infty}^2 p_{,tt} + \alpha (pp_{,t})_{,t} = 0, \quad (10)$$

where a comma denotes differentiation with respect to that variable. The lumped parameter  $\alpha$  depends upon the linear and nonlinear parameters of the borehole fluid as well as those of the formation. It is well-known (Winkler and Liu, 1996, and references therein) that the nonlinear parameters in rocks, the so-called third-order elastic constants, are orders of magnitude larger than those in nonporous media. However, if the shear modulus of the formation,  $\mu$ , is spatially uniform in the vicinity of the borehole, then it turns out that  $\alpha$  does not depend upon the third-order elastic constants of the formation at all, and it takes on a value similar to that of the borehole fluid alone. If, on the other hand, there is even a modest radial dependence,  $\mu(r)$ , then  $\alpha$  can easily be enhanced by an order of magnitude. For purposes of the later calculations, we somewhat arbitrarily take  $\alpha = 20 \times 10^{-20} \text{ s}^2/\text{dyne}$  [cf. Eq. (111) of JKN]. How the value of  $\alpha$  relates to the linear and nonlinear formation properties is precisely the subject matter of JKN.

It is worth noting that the lack of dependence of  $\alpha$  upon the third-order elastic constants of the formation when  $\mu$  is constant may be an artifact of the quasistatic limit (Sinha *et al.*, 1995). These authors considered how the speed of a tube wave changes as the static borehole pressure is increased, and they did this calculation as a function of frequency. An altered zone,  $\mu(r)$ , was not considered. They showed that the change in the tube wave speed does not depend upon the values of third-order elastic constants in the quasistatic limit, as proved in JKN, but does depend upon them as the frequency is increased to the point where the relevant wavelengths become comparable to the borehole diameter. One might therefore expect that the value of  $\alpha$  would depend upon the values of the third-order constants even for homogeneous formations, once one is out of the quasistatic limit.

In a real oil-field application, in which the borehole is drilled through various formations of geological interest,  $\alpha$  will, therefore, vary as a function of depth, though it may be relatively constant over the array of receivers. For the purposes of the present article, however, in which the nonlinear effects are weak in a way to be described later, all the nonlinear effects scale as  $\alpha P_0^2$  where  $P_0$  is the initial amplitude of the wave. Thus, a change in the value of  $\alpha$  is essentially the same as a change in the initial amplitude, *viz.*,

$$p(\xi\alpha, P_0/\xi; z, t) = p(\alpha, P_0; z, t)/\xi, \quad (11)$$

where  $\xi$  is an arbitrary scale factor. We return to this point later.

Equation (10) is the same as that for other one-dimensional nonlinear wave propagation problems (Hamilton and Blackstock, 1998). To make connection with more standard notation, it is rewritten as

$$p_{,zz} - S_{\infty}^2 p_{,tt} + \frac{\beta S_{\infty}^4}{\rho_f} (p^2)_{,tt} = 0, \quad (12)$$

where

$$\beta = \frac{\alpha \rho_{f0}}{2 S_{\infty}^4}. \quad (13)$$

This last parameter is dimensionless and, given the value of  $\alpha$  (above), has the value  $\beta = 50.5$ ; actual values may be larger than this by an order of magnitude or more. By comparison, water has the value  $\beta [\equiv 1 + (B/2A)] = 3.6$  (Hamilton and Blackstock, 1998).

## C. Tube wave propagation in a nonlinear, permeable formation

Equation (10) was derived from four fundamental equations: (a) The one-dimensional Navier–Stokes equation for the borehole fluid, (b) the constitutive relation between the pressure in the borehole fluid and its density, (c) the radial displacement of the borehole wall as a function of applied pressure, and (d) the equation of fluid conservation. There is an additional assumption that one is dealing with traveling waves. This assumption introduces errors in the treatment of so-called local nonlinear effects which shall be neglected in this article, precisely because the wave field is progressive (Naze Tjøtta and Tjøtta, 1987, and references therein). When these four equations are explicitly written, including all terms linear and quadratic in the field variables; Eq. (10) results.

In the present article, the same derivation is followed. The only difference is that the fluid-conservation equation must take into account the fact that some of the fluid enters/leaves the borehole by flowing in and out through the walls of the formation. Let  $U(z, t)$  represent the radial volumetric flow of fluid into the borehole per unit area of the borehole wall. Consider the mass conservation of fluid in a region between  $z$  and  $z + \Delta z$ . This equation is the same as Eq. (8) of JKN, but now modified by the radial fluid flow

$$\begin{aligned} & \pi [b + u(z + \Delta z)]^2 \rho(z + \Delta z) v(z + \Delta z) \\ & - \pi [b + u(z)]^2 \rho(z) v(z) + 2\pi b \Delta z \rho_{f0} U(z) \\ & + \frac{\partial}{\partial t} \{ \pi [b + u(z)]^2 \rho(z) \Delta z \} = 0, \end{aligned} \quad (14)$$

where  $\rho$ ,  $v$ , and  $u$  are, respectively, the fluid density, the fluid velocity, and the radial displacement of the borehole wall, each of which is a function of  $z$  and  $t$ . For the purposes of the present paper, it is only necessary to include attenuation/dispersion effects predicted by linear theory. For that reason, the term involving  $U$  is multiplied by the equilibrium borehole radius,  $b$ , and the equilibrium fluid density,  $\rho_{f0}$ .

The equations governing the other assumptions, (a)–(c) (above), are unaffected; they are, respectively, Eqs. (3), (4), and (9) of JKN. Combining these equations, one arrives at the following equation:

$$\frac{\partial^2 p}{\partial z^2} - S_\infty^2 \frac{\partial^2 p}{\partial t^2} - \frac{2\rho_{f0}}{b} \frac{\partial U}{\partial t} + (\alpha/2) \frac{\partial^2 p^2}{\partial t^2} + \mathcal{G} = 0, \quad (15)$$

where  $\mathcal{G}$  is a complicated expression which is quadratic in the field variables. It has the property that it vanishes for traveling waves, of the form  $p = p(t \pm S_\infty z)$ . Corrections to any derived results due to the neglect of this term are confined to the local nonlinear effects as discussed by Naze Tjøtta and Tjøtta (1987). In accordance with standard practice, we neglect it in this article.

Equation (15) (with  $\mathcal{G} = 0$ ) can be simplified if the wave is traveling in, say, the  $+z$  direction, by using a multiple-scales derivation. Let  $Z_1 = \epsilon z$ , and  $\tau = t - S_\infty z$ . The coordinate  $Z_1$  is referred to as the “slow scale” in the retarded time frame  $\tau$ . In the ensuing equation, one can neglect terms of order  $\epsilon^2$  as compared to those of order  $\epsilon$ . The result is

$$\frac{\partial p}{\partial z} + \frac{\partial F}{\partial \tau} - \frac{\beta S_\infty^3}{2\rho_{f0}} \frac{\partial p^2}{\partial \tau} = 0, \quad (16)$$

where the substitution  $U = (bS_\infty/\rho_{f0})(\partial F/\partial \tau)$  has been made;  $F$  represents essentially the relative fluid displacement into/out of the borehole wall. Equation (16) is a kind of generalized Burgers equation, in the sense of Blackstock (1985), for nonlinear tube wave propagation, once the connection between  $F$  and  $p$  has been made. Indeed, if  $F \propto -(\partial p/\partial \tau)$  one retrieves the Burgers equation exactly (Hamilton and Blackstock, 1998). In other contexts the relationship between  $F$  and  $p$  is not so simple;  $F$  is represented as a convolution over the past history of  $p$ , which lends itself to numerical solution only in special circumstances, as when the response is that of a “relaxing medium” (Webster and Blackstock, 1977; Cleveland *et al.*, 1996). In this special case there are even some exact analytical solutions: (Soluyan and Khokhlov, 1962; Polyakova *et al.*, 1962); see also Rudenko and Soluyan, 1977.

For the problem at hand,  $F$  is related to the borehole pressure by solving the following *gedanken* problem: Let the pressure at the borehole wall be axially uniform and vary sinusoidally in time. It is straightforward to solve for the sinusoidal flow through the formation, essentially following the derivation of a similar problem given by White (1983). The result is the relation between  $F$  and  $p$  in the Fourier transform domain

$$\tilde{F}(z, \omega) = \tilde{\Theta}(\omega) \tilde{p}(z, \omega), \quad (17)$$

where  $\tilde{\Theta}(\omega)$  is given by Eq. (8). Notice that Eqs. (15) and (17), in the small amplitude limit, reproduce Eq. (6), whereas Eqs. (16) and (17) reproduce Eq. (9), in that limit. Equations (16) and (17) constitute the basic equations of motion for tube wave evolution in a permeable formation, taking into account quadratic nonlinearity, in the quasistatic limit. Numerical results for representative sets of parameter values will be given in the next section. First, though, is a brief outline of the numerical method of calculation.

## D. Lax–Wendroff method of calculation

Equation (16) is rewritten as

$$\frac{\partial p}{\partial z} + \frac{\partial H}{\partial \tau} = 0, \quad (18)$$

where

$$H = F - \frac{\beta S_\infty^3}{2\rho_{f0}} p^2. \quad (19)$$

The idea is that  $p$ , and thereby  $H$ , is evaluated on a mesh of  $\tau$  values  $\{\tau_j = j\Delta\tau\}$  at the current value of  $z = z_i$ , and Eq. (18) is used to advance to the next value of  $z_{i+1} = z_i + \Delta z_i$ ; the process is started at  $z = 0$ . The quantity  $F$  is most efficiently calculated from  $p$  by means of the fast Fourier transform (FFT) and its inverse, as implied by Eq. (17), rather than as a convolution in the time domain. This method differs, then, from that of Cleveland *et al.* (1996), who investigated the effects of attenuation/dispersion for the specific case of a relaxing fluid, which is amenable to a time-domain calculation. Korpel (1980) and Cervenka and Alais (1990) have used frequency-domain techniques, but these were done via the method of coupled linear modes, which would be prohibitive to implement for the pulses considered in the present instance. Equations such as (19) are most effectively advanced by means of implicit methods, but this would defeat the purpose of the FFT, the matrix that would be inverted at each step would be of size  $N \times N$ , where  $N$  is the number of mesh points in the  $\tau$  domain. On the other hand, efficiency is not particularly necessary for this calculation, but accuracy and stability are. Thus, one is led to seek an explicit finite differencing scheme which is stable and which can be made as accurate as needed simply by refining the meshes. The Lax–Wendroff method is such a scheme, and is explained simply by Press *et al.* (1988). At each value of  $z_i$ , the values of  $p$  on the  $\tau$ -mesh are used to calculate values of  $H$ . Values of  $p$  at “midway” points are then calculated. From these midway points, a new set of values for  $H$  are constructed so that a centered-difference estimate of the derivative with respect to  $z$  can be performed. The local amplification factor,  $\xi(\omega)$ , for the growth of oscillations of temporal frequency,  $\omega$ , can be readily derived

$$\xi(\omega) = 1 + 2i \frac{\Delta z}{\Delta \tau} f g \sin(\omega \Delta \tau / 2), \quad (20)$$

where

$$f(\omega) = \tilde{\Theta}(\omega) - \frac{\beta S_\infty^3 P_0}{\rho_{f0}}, \quad (21)$$

and

$$g(\omega) = \cos(\omega \Delta \tau / 2) + i \frac{\Delta z}{\Delta \tau} f \sin(\omega \Delta \tau / 2). \quad (22)$$

$P_0$  is the assumed local value of the pressure,  $p$ . Stability, at least local stability, is guaranteed if  $|\xi| < 1$  for all relevant values of the parameters and this was straightforward to accomplish simply by making  $\Delta z$  small enough. It is worth noting that this scheme would be unstable,  $|\xi| > 1$ , if

the time derivative in Eq. (16) were evaluated in the frequency domain:  $\partial/\partial\tau \rightarrow -i\omega$ .

## II. RESULTS

The purpose of this article is to investigate the combined effects of nonlinearity and formation permeability on tube wave propagation, as applied to potentially realizable situations. Specifically, the initial pulse shape is chosen to be very narrow-band, of the form

$$p(z=0, \tau) = E_1(\tau) \sin(\omega_1 \tau), \quad (23)$$

where the envelope function,  $E_1(\tau)$ , is slowly varying over many periods of the fundamental. The following envelope function is most suitable for the present purposes:

$$E_1(\tau) = P_0 \exp[-(\tau/T_W)^{10}]. \quad (24)$$

There are two characteristic distances relevant to this problem. The first is the distance over which a pulse would be expected to decay, according to the predictions of the linearized theory. This is simply  $Z_{\text{att}} = 1/\gamma$ , where

$$\gamma = \omega_1 \Im[\tilde{\Theta}(\omega_1)]. \quad (25)$$

The second is the distance over which a shock front would develop, in the absence of attenuation. This is given by Hamilton and Blackstock (1998)

$$Z_{\text{shock}} = \frac{\rho f_0}{\beta S_\infty^3 \omega_1 P_0}. \quad (26)$$

The Gol'dberg number, the ratio of these two distances, is a measure of the importance of nonlinear effects relative to the linear

$$\Gamma = Z_{\text{att}}/Z_{\text{shock}}. \quad (27)$$

In all cases considered in the present article, the initial amplitude of the pulse was chosen such that the Gol'dberg number has the same value, namely  $\Gamma = 0.21$ . Thus, the amplitude is large enough that nonlinear effects are significant, but overall pulse propagation is dominated by linear attenuation/dispersion. The parameter sets are given in Table I. The relatively high permeabilities ( $1 \text{ Darcy} = 1 \mu\text{m}^2$ ) and the relatively large amplitudes guarantee that pulse self-demodulation will occur in a relatively short distance. The product  $\omega_1 \times T_W$  is the same, so that each pulse initially "looks" the same. The results of the numerical calculations based on the parameters of Table I are plotted in Figs. 2–4 in a manner which invites comparison against the three-dimensional results in Fig. 1 of Averkiou *et al.* (1993), who used the same shape initial pulse. At each distance,  $z$ , the waveform as well as its Fourier spectrum are plotted. Initially, the spectrum of the pulse is centered on the carrier frequency,  $\omega_1$ . The nonlinear term in Eq. (16) has Fourier components centered around  $\omega = 2\omega_1$  as well as those centered around  $\omega = 0$ , and therefore as the pulse evolves it acquires Fourier components centered on these two frequencies. Because the attenuation predicted by Eqs. (8) and (9) always increases with frequency, and because the Gol'dberg number is less than unity, the components centered at  $\omega = 2\omega_1$  never become large enough to be visible on the scale

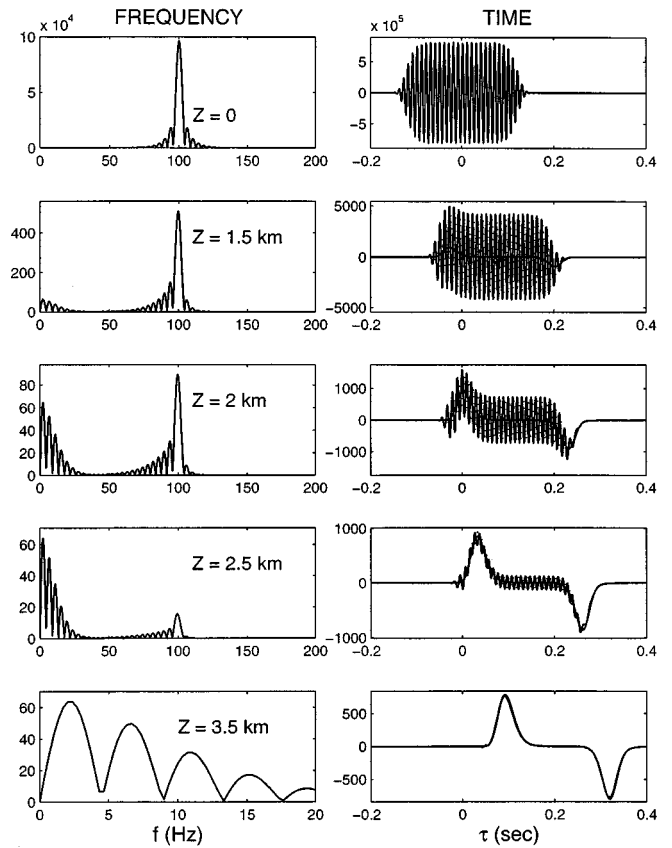


FIG. 2. Calculated pulse evolution using parameter set A, from Table I. The right column is the signal in the time domain, the left is the frequency spectrum. Also shown as a solid line on the right is the analytical expression for the self-demodulated pulse, Eq. (30), toward which the pulse evolves. Similarly, the analytical expression for the total signal, given by the first two terms of Eq. (28), is shown as a dashed line; it essentially overlies the numerically calculated signal.

of the figures (but see below). The component centered at  $\omega = \omega_1$  continuously decays. The components centered around  $\omega = 0$  initially grow, level off as their source, the original pulse, decays, and eventually they, too, suffer a very slow decay. At large enough distances, these low-frequency components generated by the nonlinear effects are the dominant ones; this is pulse self-demodulation in one dimension. The shapes of the self-demodulated pulses appear to be very different in the three cases; nonetheless, an analytical expression for the shape can be derived, as follows:

Because the Gol'dberg number is less than unity, the evolution of the pulse is dominated by linear acoustics. Because the pulse contains so many oscillations of the fundamental, it has a narrow-band frequency spectrum, as is clear from Figs. 2–4. According to linear theory, this means that the envelope function propagates with the group velocity, it decays exponentially with the decay rate  $\gamma$ , and each peak and trough travel with the phase velocity. This suggests the following *ansatz* for pulse evolution:

$$\begin{aligned} p(z, \tau) = & E_0(z, \tau) + E_1(\tau - \Delta S_g(\omega_1)z) \\ & \times \exp(-\gamma z) \sin[\omega_1(\tau - \Delta S_p(\omega_1)z)] \\ & + E_2(z, \tau) \sin[2\omega_1(\tau - \Delta S_p(2\omega_1)z) + \phi_2] \\ & + \dots \end{aligned} \quad (28)$$

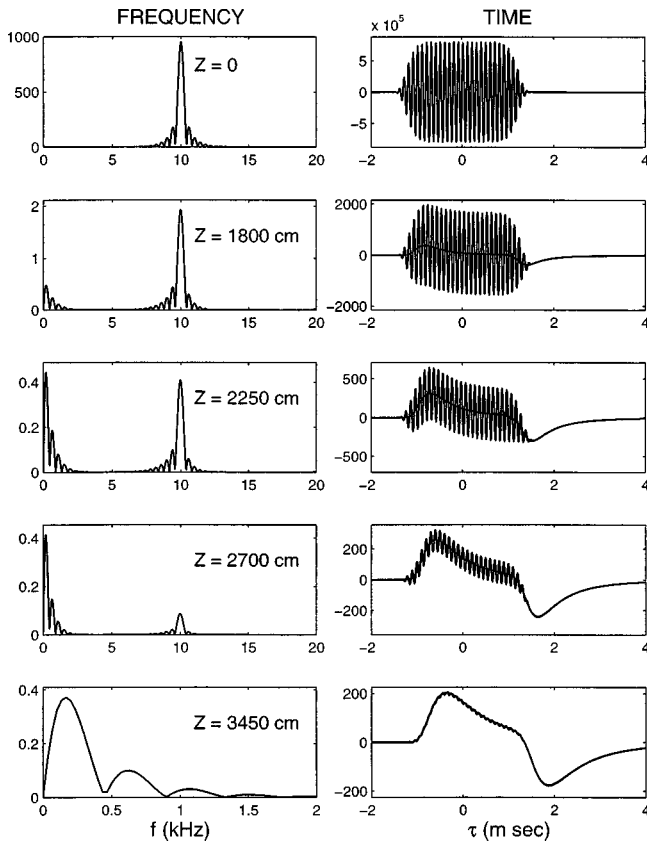


FIG. 3. Calculated pulse evolution using parameter set B, from Table I. Same conventions as Fig. 2.

$E_1$  is the original envelope function,  $\Delta S_p(\omega)$  is the additional phase slowness as defined earlier, and

$$\Delta S_g(\omega) = \frac{d\Re[\omega\tilde{\Theta}(\omega)]}{d\omega} \quad (29)$$

is the additional group slowness.  $E_0$  and  $E_2$  represent the nonlinearly generated components having Fourier structure near  $\omega=0$  and  $\omega=2\omega_1$ , respectively. In contrast to the  $z$ -dependence of the primary pulse, which is known from general considerations of linear acoustics and is therefore written explicitly, the  $z$ -dependence of these other two components must be determined from the equations of motion.

It is straightforward to substitute Eq. (28) into Eq. (16) and solve for  $E_0$ . Here, the  $E_1$  part of the signal forms the dominant part of the nonlinear term and acts as a source for the generation of both  $E_0$  and  $E_2$ , with the initial condition  $E_0(0, \tau) = E_2(0, \tau) = 0$ . The result for  $E_0$  is best obtained by means of the Fourier transform of Eq. (16)

$$E_0(z, \tau) = \frac{\beta S_\infty^3}{4\rho_{f0}} \int_{-\infty}^{+\infty} \frac{-i\omega\tilde{E}_1^2(\omega)}{i\omega\Delta S_g - 2\gamma - iq(\omega)} \times [e^{(i\omega\Delta S_g - 2\gamma)z} - e^{iq(\omega)z}] e^{-i\omega\tau} d\omega, \quad (30)$$

where

$$q(\omega) = \omega\tilde{\Theta}(\omega) \quad (31)$$

is the reduced wave number and

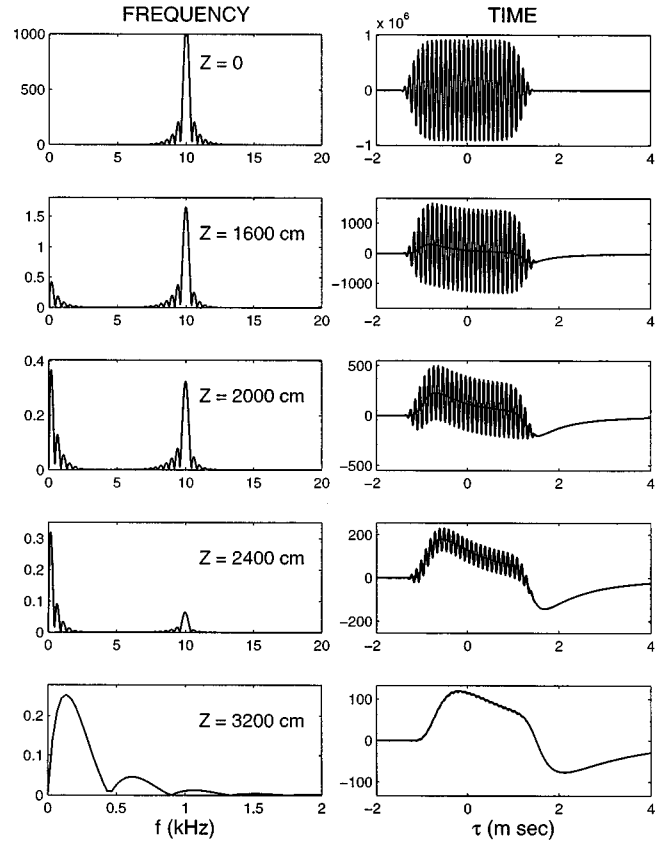


FIG. 4. Calculated pulse evolution using parameter set C, from Table I. Same conventions as Fig. 2.

$$\tilde{E}_1^2(\omega) = \frac{1}{2\pi} \int E_1^2(\tau) e^{i\omega\tau} d\tau \quad (32)$$

is the Fourier transform of the square of the envelope function (*not* the square of the Fourier transform).  $\Delta S_g$  and  $\gamma$  are evaluated at the carrier frequency,  $\omega_1$ , but otherwise the relevant spectrum appearing in Eq. (30) is confined to the low-frequency content of the envelope function and not to the original center frequency. Equation (30) is plotted in Figs. 2–4, where one can see that the numerically computed waveforms do indeed evolve to this solution. Moreover, the first two terms of Eq. (28) are also plotted as dashed lines, which dashed lines overlie the plot of the full numerical solutions.

There are several limiting cases of Eq. (30) which are of particular interest:

- (a) Suppose that in the range of frequencies relevant to Eq. (30), the attenuation at the carrier frequency is large:  $2\gamma(\omega_1) \gg |q(\omega) - \omega\Delta S_g(\omega_1)|$ . This can happen if, for example, there is negligible dispersion:  $q(\omega) \approx \Delta S_g(\omega_1)\omega$ . Therefore, in this limit

$$E_0(z, \tau) = \frac{\beta S_\infty^3}{8\rho_{f0}\gamma} [1 - e^{-2\gamma z}] \frac{\partial}{\partial \tau} [E_1^2(\tau - \Delta S_g(\omega_1)z)]. \quad (33)$$

The presence of the first order time derivative in Eq. (33) does not depend upon the details of the actual attenuation mechanism. This first derivative distinguishes pulse self-demodulation in one dimension from the second derivative, which occurs in the far field of a three-dimensional radiating pulse (Averkiou *et al.*, 1993, and references therein).

In the special case that one is dealing specifically with the Burgers equation,  $F = -a(\partial p/\partial \tau) \leftrightarrow \tilde{\Theta}(\omega) = +ia\omega, \leftrightarrow q(\omega) = ia\omega^2$ , Eq. (30) always reduces to Eq. (33).

- (b) Suppose that there is significant dispersion between the low-frequency content of Eq. (30) and  $\omega_1$ , but negligible attenuation throughout:  $q = \Delta S\omega$ , and  $\gamma \ll |\Delta S - \Delta S_g| \omega$ , where  $\Delta S$  is effectively frequency independent. Then

$$E_0(z, \tau) = \frac{\beta S_\infty^3}{4\rho_{f0}(\Delta S - \Delta S_g)} [E_1^2(\tau - \Delta S_g z) - E_1^2(\tau - \Delta S z)]. \quad (34)$$

- (c) Regardless of the dispersion, which may be complicated, the initial growth of the self-demodulated signal is simple

$$\lim_{z \rightarrow 0} E_0(z, \tau) = \frac{\beta S_\infty^3}{4\rho_{f0}} \frac{\partial E_1^2(\tau)}{\partial \tau} z + \mathcal{O}(z^2). \quad (35)$$

These analytical results allow one to gain insight into the results of Figs. 2–4. First, in Fig. 2 the carrier frequency,  $\omega_1/2\pi = 100$  Hz, is well to the left of the attenuation/dispersion effects, as can be seen from Fig. 1. The conditions assumed in the derivation of Eq. (33) are satisfied, as can also be seen from Fig. 1. Therefore, the self-demodulated signal is essentially that predicted by Eq. (33), as is more or less obvious from Fig. 2. Note that the additional group slowness,  $\Delta S_g(\omega_1)$ , is appreciable (as compared to  $S_\infty$ ), and so the self-demodulated signal progressively moves to later times, as  $z$  increases. Next, the frequency is increased and the permeability is decreased for Fig. 3. These two changes are in fact really the same change, done twice: Because  $\tilde{\Theta}$  is a function only of the ratio  $\omega/\kappa_0$ , it is simple to show that solutions to Eqs. (16) and (17) obey a scaling relationship. Let  $p = p_1(z, \tau; \kappa_0/\eta)$  be a solution to Eqs. (16) and (17). Consider the same problem, but with a permeability scaled up or down by a factor  $\zeta$ . It is straightforward to show that  $p_1$  with rescaled axes is also a solution to the new problem

$$p(z, \tau; \zeta \kappa_0/\eta) = p_1(\zeta z, \zeta \tau; \kappa_0/\eta). \quad (36)$$

In the present article, this means that if the frequency  $\omega_1$  is scaled up by the same factor as  $\kappa_0/\eta$  (holding  $\omega_1 \times T_W$  fixed), the resulting waveforms look the same, but with  $z$  and  $\tau$  rescaled. Figure 3 could also be considered the result of  $\kappa_0 = 0.02 \mu\text{m}^2$  and  $\omega_1/2\pi = 10^3$  Hz, etc. From Fig. 1, it is clear that  $q(\omega)$ , for frequencies in the spectrum of the demodulated signal of Fig. 3 as per Eq. (30), has a much more complicated dispersion relation than for Fig. 2. Consequently, the demodulated signal for Fig. 3 is not well represented by Eq. (33), although Eq. (30) is still valid.

To emphasize further the effects of dispersion and attenuation on the demodulated signal, the membrane parameter,  $W_{mc}$  is reduced relative to that of Fig. 3. The results are shown in Fig. 4; all the other parameters are the same as in Fig. 3. Here, it is clear from Fig. 1 that dispersion and attenuation effects are more important than before, and this is born out by the results: The demodulated signal in Fig. 4 is significantly more broadened than that in Fig. 3, especially for later times. One can see that it is a rough—very rough—

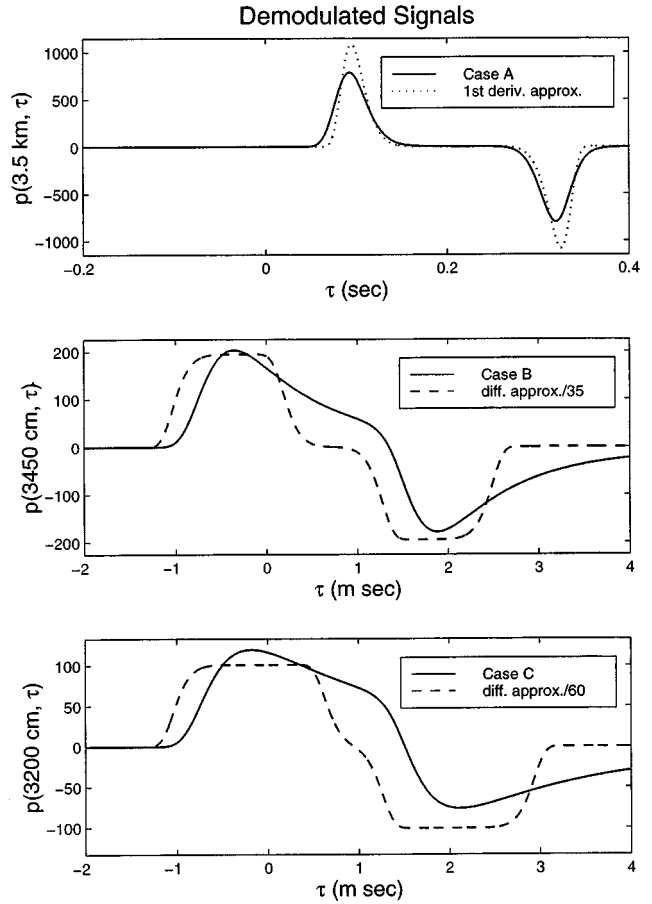


FIG. 5. A comparison of the demodulated signals, Eq. (30), for the three examples in Figs. 2–4. In case A, comparison is made against the approximation given by Eq. (33). In the other two cases, comparison is made against the difference equation, (34), which has been scaled down for the comparison.

approximation to Eq. (34). Note that in all three cases considered, the initial growth of the demodulated signal is given by Eq. (35).

Let us compare/contrast the demodulated signals predicted for the three different data sets, i.e., the last signals in Figs. 2–4. They are plotted in Fig. 5. In the top plot, the demodulated signal for case A is compared against the prediction of Eq. (33), which is seen to be in excellent agreement therewith. There is a small difference between the value of the additional group slowness at the carrier frequency,  $\Delta S_g(\omega_1)$ , and the low-frequency limit thereof  $\Delta S_g(\omega=0) \equiv \Delta S_p(\omega=0) = (\rho_{f0}/S_\infty W_{mc} b)$ ; the latter value was used in Eq. (33), plotted in Fig. 5. Cases B and C are shown in the middle and bottom plots, respectively, of Fig. 5; they are compared against the difference approximation, Eq. (34). The spectra of these two demodulated signals are peaked in the vicinity of 100 Hz, as can be seen from the last rows of Figs. 3 and 4. Accordingly, the value of the additional phase slowness,  $\Delta S_p$ , evaluated at the peak in the demodulated spectrum, was used in Eq. (34). Moreover, the demodulated signals are very much attenuated relative to the predictions of Eq. (34); accordingly, the latter has been reduced by a scale factor in order to make a comparison of the two shapes. Although Eq. (34) is a better approximation to these demodulated signals than is Eq. (33), there is no simple

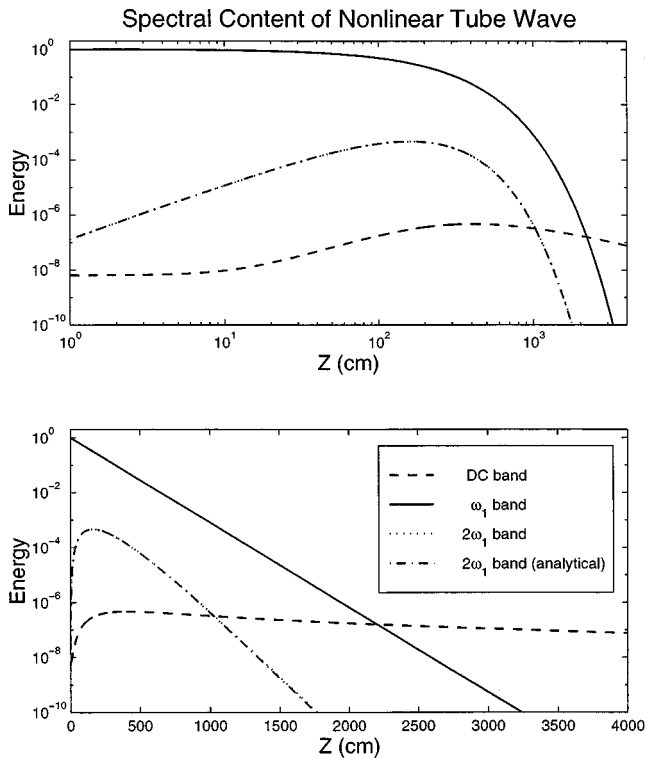


FIG. 6. Numerically calculated values of the energy in each of the three bands defined by Eq. (37) for the waveforms of Fig. 3, sample B. Each curve has been normalized by the value of  $\mathcal{E}_1(z=0)$ . The data are shown in both log-log and semilog plots, in order to emphasize the short and long distance behavior, respectively. Also shown is the analytical expression for  $\mathcal{E}_2(z)$  given by Eq. (39); the fact that it is indistinguishable from the numerically calculated curve is precisely the point.

way in which both of the assumptions leading to Eq. (34) can be satisfied accurately.

The distances over which tube wave demodulation can be expected to occur in a realistic borehole geometry are rather long, even though the amplitudes of the signals have been taken to be very large. A more practical way of detecting nonlinear effects would be to measure the amplitude of second-harmonic generation. As was mentioned earlier, the second-harmonic component of the propagating signal is not apparent in Figs. 2–4, but it is there, nonetheless, and is large enough in the examples given that it could be measured using appropriate filters.

It is convenient to define the energy in the various bands as

$$\begin{aligned}\mathcal{E}_0(z) &= \int_0^{\omega_1/2} |\bar{p}(z, \omega)|^2 d\omega, \\ \mathcal{E}_1(z) &= \int_{\omega_1/2}^{3\omega_1/2} |\bar{p}(z, \omega)|^2 d\omega, \\ \mathcal{E}_2(z) &= \int_{3\omega_1/2}^{5\omega_1/2} |\bar{p}(z, \omega)|^2 d\omega.\end{aligned}\quad (37)$$

For the case of sample B, whose waveforms were shown in Fig. 3, the energy in each of these three bands is plotted in Fig. 6, normalized against the value  $\mathcal{E}_1(z=0)$ . The plots

are shown in both log-log and semilog form in order to emphasize the short and the long distance behaviors, respectively. First, because the amplitude is relatively small,  $\Gamma < 1$ , the “carrier” energy decays nearly exponentially in accordance with the predictions of linear theory,  $\mathcal{E}_1(z) = \exp[-2\gamma(\omega_1)z]$ , as is obvious from the figure. Next, although the energies in the other two bands initially are non-zero because of the finite width of the initial pulse, their subsequent growth,  $\mathcal{E}_{0,2}(z) \propto z^2$ , is a manifestation of the nonlinearity of the system. At large distances, the carrier energy,  $\mathcal{E}_1(z)$ , decays away; the second-harmonic energy,  $\mathcal{E}_2(z)$ , decays even more rapidly, and the remaining low-frequency signal  $\mathcal{E}_0(z)$  dominates, although it, too, eventually decays. The region where  $\mathcal{E}_0(z) \approx \mathcal{E}_1(z)$  is where demodulation sets in; the plots in rows 2–4 of Figs. 2, 3, and 4 depict that crossover region. Because nonlinear effects are relatively small as compared to linear decay, the second-harmonic energy never gets larger than the fundamental before it decays. Nonetheless, for the values of the parameters considered here it is possible to measure this energy—only  $\sim 30$  dB down from the carrier for these examples, which are representative of field conditions.

It is also possible to derive a simple analytical expression for this second-harmonic energy, analogous to Eq. (30) for the demodulated signal (small nonlinearity, long pulses). It is easy to solve Eq. (16) for the Fourier components near  $2\omega_1$

$$\bar{p}(z, \omega) = \frac{\beta S_\infty^3 \omega_1 \tilde{E}_1^2(\omega - 2\omega_1) [e^{i2q(\omega_1)z} - e^{iq(2\omega_1)z}]}{4\rho_{f0} [2q(\omega_1) - q(2\omega_1)]}, \quad (38)$$

where all slowly varying prefactors have been evaluated at  $\omega = 2\omega_1$ . Therefore,

$$\mathcal{E}_2(z) = 2\pi \left( \frac{\beta S_\infty^3 \omega_1}{4\rho_{f0}} \right)^2 \left| \frac{e^{i2q(\omega_1)z} - e^{iq(2\omega_1)z}}{2q(\omega_1) - q(2\omega_1)} \right|^2 \int_{-\infty}^{+\infty} E_1^4(\tau) d\tau, \quad (39)$$

where use has been made of Parseval’s theorem by an obvious extension of certain limits of integration to  $\pm\infty$ . Equation (39) is plotted in Fig. 6, where one can see that it is an excellent approximation to the numerically calculated energy  $\mathcal{E}_2(z)$ . The role of the factor within the absolute value signs has long been recognized in the field of optical second-harmonic generation, where dispersion is extremely important, attenuation less so (Yariv, 1967). The role of the envelope function for pulses seems, however, not to have been noticed before. The reader is reminded that Eq. (39), like Eq. (30), presumes that nonlinear effects are small ( $\Gamma < 1$ ) and the pulses are narrow-band (e.g.,  $\omega_1 \times T_W \gg 1$ ), although Eq. (39) applies to any shape of pulse.

At this point it is qualitatively clear what happens when the value of either  $\alpha$  or  $P_0$  is changed. Because the Gol’dberg number is small,  $\Gamma < 1$ , the first three terms of Eq. (28) are an excellent approximation to all the effects calculated in the present article, and this approximation obeys the scaling relationship (11) by inspection. Referring to Fig. 6, we see that if either  $\alpha$  or  $P_0$  is increased by a multiplicative factor  $\xi$ , the energy  $\mathcal{E}_1(z)$  (scaled against its initial value at  $z=0$ ) does not change, whereas the scaled energies  $\mathcal{E}_0(z)$

and  $\mathcal{E}_2(z)$  increase by a factor  $\xi^2$ , with no change in shape. It is in this sense that a change in  $\alpha$  is equivalent to a change in  $P_0$ . This claim is not true when the Gol'dberg number is large.

### III. CONCLUSIONS

An approximate tube wave propagation equation, analogous to the Burgers equation, has been derived which takes into account quadratic nonlinearity as well as dispersion and attenuation due to fluid-flow permeability. Numerical results for narrow-band pulses in which the amplitude of the wave is relatively weak, and the pulse duration is long, demonstrate that self-demodulation does occur, as for the Burgers equation. The self-demodulated pulse shape can be significantly different from the derivative of the envelope function squared, predicted by the Burgers equation; these differences are due to the dispersive nature of tube wave propagation in the linear regime. Even so, an analytically derived formula for the self-demodulated pulse shape, Eq. (30), is in excellent agreement with the numerical results in the three cases considered. This result makes it easy to see that if the carrier frequency is considerably lower than that of the peak in the specific attenuation, the self-demodulated signal is essentially the first derivative of the square of the envelope function, Eq. (33), the same as for the Burgers equation. If the carrier frequency is considerably higher than the peak frequency, then the self-demodulated signal resembles a distorted version of the differenced envelope function, Eq. (34). Similarly, the amplitude of second-harmonic generation is calculated and compared with an analytical result which is seen to be numerically exact. This last result points the way to measuring the strength of the nonlinear coefficient in a borehole geometry. We remind the reader once again that these conclusions apply only in the limit of small Gol'dberg number  $\leftrightarrow$  weak nonlinearity.

### ACKNOWLEDGMENT

I am grateful to Mark Hamilton for several very informative conversations, and to T. J. Plona and K. W. Winkler for critically reading the manuscript.

- Averkiou, M. A., Lee, Y.-S., and Hamilton, M. F. (1993). "Self-demodulation of amplitude- and frequency-modulated pulses in a thermoviscous fluid," *J. Acoust. Soc. Am.* **94**, 2876–2883.
- Blackstock, D. T. (1985). "Generalized Burgers equation for plane waves," *J. Acoust. Soc. Am.* **77**, 2050–2053.
- Cervenka, P., and Alais, P. (1990). "Fourier formalism for describing nonlinear self-demodulation of a primary narrow ultrasonic beam," *J. Acoust. Soc. Am.* **88**, 473–481.
- Cleveland, R. O., Hamilton, M. F., and Blackstock, D. T. (1996). "Time-domain modeling of finite amplitude sound in relaxing fluids," *J. Acoust. Soc. Am.* **99**, 3312–3318.
- Hamilton, M. F., and Blackstock, D. T. (1998). *Nonlinear Acoustics* (Academic, New York).
- Johnson, D. L., Kostek, A. N., and Norris, A. N. (1994). "Nonlinear tube waves," *J. Acoust. Soc. Am.* **96**, 1829–1843.
- Korpel, A. (1980). "Frequency approach to nonlinear dispersive waves," *J. Acoust. Soc. Am.* **67**, 1954–1958.
- Liu, H.-L., and Johnson, D. L. (1997). "Effects of an elastic membrane on tube waves in permeable formations," *J. Acoust. Soc. Am.* **101**, 3322–3329.
- Naze Tjøtta, J., and Tjøtta, S. (1987). "Interaction of sound waves. Part I. Basic equations and plane waves," *J. Acoust. Soc. Am.* **82**, 1425–1428.
- Pampuri, F., Rovellini, M., Brie, A., and Fukushima, T. (1998). "Effective evaluation of fluid mobility from Stoneley waves using full Biot model inversion: Two case histories," in *SPE Annual Technical Conference and Exhibition*, Paper No. 49132, New Orleans (SPE, Richardson, TX, 1998).
- Polyakova, A. L., Soluyan, S. I., and Khokhlov, R. V. (1962). "Propagation of finite disturbances in a relaxing medium," *Sov. Phys. Acoust.* **8**, 78–82.
- Press, W. H., Flannery, B. P., Teukolsky, S. A., and Vetterling, W. T. (1988). *Numerical Recipes: The Art of Scientific Computing* (Cambridge University Press, New York).
- Rudenko, O. V., and Soluyan, S. I. (1977). *Theoretical Foundations of Nonlinear Acoustics* (Consultants Bureau/Plenum, New York).
- Sinha, B. K., Kostek, S., and Norris, A. N. (1995). "Stoneley and flexural modes in pressurized boreholes," *J. Geophys. Res.* **100**, 22375–22381.
- Soluyan, S. I., and Khokhlov, R. V. (1962). "Finite amplitude acoustic waves in a relaxing medium," *Sov. Phys. Acoust.* **8**, 170–175.
- Tang, X. M. (1994). "Effects of mudcake on the measurement of fluid flow properties using borehole acoustic waves," *SEG Expanded Abstracts, 64th Annual Meeting, Los Angeles (SEG, Tulsa, 1994)*, p. 78, BG 3.4.
- Webster, D. A., and Blackstock, D. T. (1977). "Finite-amplitude saturation of plane sound waves in air," *J. Acoust. Soc. Am.* **62**, 518–523.
- White, J. E. (1983). *Underground Sound—Application of Seismic Waves* (Elsevier, Amsterdam).
- Winkler, K. W., Liu, H.-L., and Johnson, D. L. (1989). "Permeability and borehole Stoneley waves: Comparison between experiment and theory," *Geophysics* **54**, 66–75.
- Winkler, K. W., and Liu, X. (1996). "Measurements of third-order elastic constants in rocks," *J. Acoust. Soc. Am.* **100**, 1392–1398.
- Yariv, A. (1967). *Quantum Electronics* (Wiley, New York).

# Static friction threshold effects on hysteresis and frequency spectra in rocks

Abraham Kadish

*Los Alamos National Laboratory, Los Alamos, New Mexico 87545*

(Received 14 July 1998; accepted for publication 5 March 1999)

Static friction in rocks is modeled as a threshold phenomenon. An inverse problem in ground motion is analyzed for propagation in a medium with an arbitrary number of fractures at which static friction thresholds determine the onset of slippage. It is shown that seismic source dynamics inferred from surface motion is unique. This result is used to show that the threshold property of static friction is not a mechanism for hysteresis in rocks. The coda for the surface motion is shown, in the case of a single fracture and a deep source of finite duration, to be either a periodic function in time, or to asymptote in time to such a function, depending on an algebraic relation between ground and fracture parameters. The fundamental period of the coda is independent of the source time series and is equal to four times the signal transit time between the single fracture and ground surface. Only odd harmonics of the fundamental frequency are present in a Fourier series of these periodic functions. © 1999 Acoustical Society of America. [S0001-4966(99)03106-9]

PACS numbers: 43.25.Dc [MFH]

## INTRODUCTION

Resonant bar and pulse propagation experiments at Los Alamos and the Institut du Petrol Francais reveal that nonlinear and plasticity effects in earth materials can be significant at even moderate strains.<sup>1,2</sup> In particular, rocks at low strain can exhibit effects attributable to hysteresis.<sup>3,4</sup> As a result of its threshold properties, static friction is a mechanism for accumulating and releasing stresses at fractures. Consequently, static friction is often thought of as a source of “memory” (i.e., hysteresis) in relations between stress and strain in rocks.

A main result of this article is that the threshold effect associated with static friction is not a source of hysteresis in rocks.

A theoretical study is presented that addresses effects of static friction thresholds at fractures on seismic source identification and ground surface motion. An idealized inverse problem is analyzed to show that static friction at any number of fractures is compatible with the unique determination of source dynamics from measurements of ground motion. This result was not obvious, since the modeling of static friction threshold effects uses discontinuous functions, and the appearance of “nonsmooth” functions of dependent variables in differential operators often leads to nonuniqueness of solutions.<sup>5</sup> (In order to isolate its effects, friction is the only nonlinear and dissipative mechanism considered, and a one-dimensional medium is employed.) The analogous problem for hysteretic materials does not have a unique solution.<sup>6,7</sup>

Rocks are materials of nonuniform composition containing pores, fractures, and, often, grains of nonuniform size and composition. In order to obtain practical mathematical descriptions of propagation in rocks, their components are usually treated collectively using an assumed state relation between stress and strain. If one imagines a “rock” composed of a dense set of very thin elastic layers separated by fractures, the inverse problem for this rock has a unique so-

lution. If this rock were hysteretic, the inverse problem would not have a unique solution: i.e., hysteresis cannot be attributed to static friction thresholds.

In this paper, the parameters that determine thresholds are taken to be constants. In real rocks, they may depend on physical variables (e.g., evolving grain geometry and temperature). Inverse problems associated with the determination of these variables may not be well posed. Nevertheless, the model used in this paper shows that the threshold property of static friction is not, by itself, a mechanism for hysteresis.

Using the same model, and assuming a single fracture and deeper seismic source of finite duration, one finds that surface codas have a quarter period equal to the signal transit time between the surface and single fracture. The relation between surface and source dynamics with slippage at fractures is strongly rate, as well as amplitude, dependent. After the seismic source terminates, friction brings the velocity at the fracture to zero while oscillations continue in the region between the fracture and surface: Only odd harmonics of the fundamental frequency survive. The details of its spectral content, and the way the coda is realized (i.e., whether in a finite time or asymptotically), depend on the source history and ground parameters. It should be noted that if one requires symmetry of hysteresis loops, the associated hysteretic state relations also imply the presence of only odd harmonics for rocks driven to periodic motion by a single (or odd harmonic) frequency source.

The model employed here, together with algebraic relations for velocities at fractures, is described in the next section. For simplicity of exposition, derivations of the velocity relations are given in Appendices A and B. The mathematical source of static friction threshold effects contained in the model is identified in Sec. II. The uniqueness of the determination of velocity at the source from measured surface velocity is proved in Sec. III. Connections between velocities and slippage are determined for the case of a single fault in



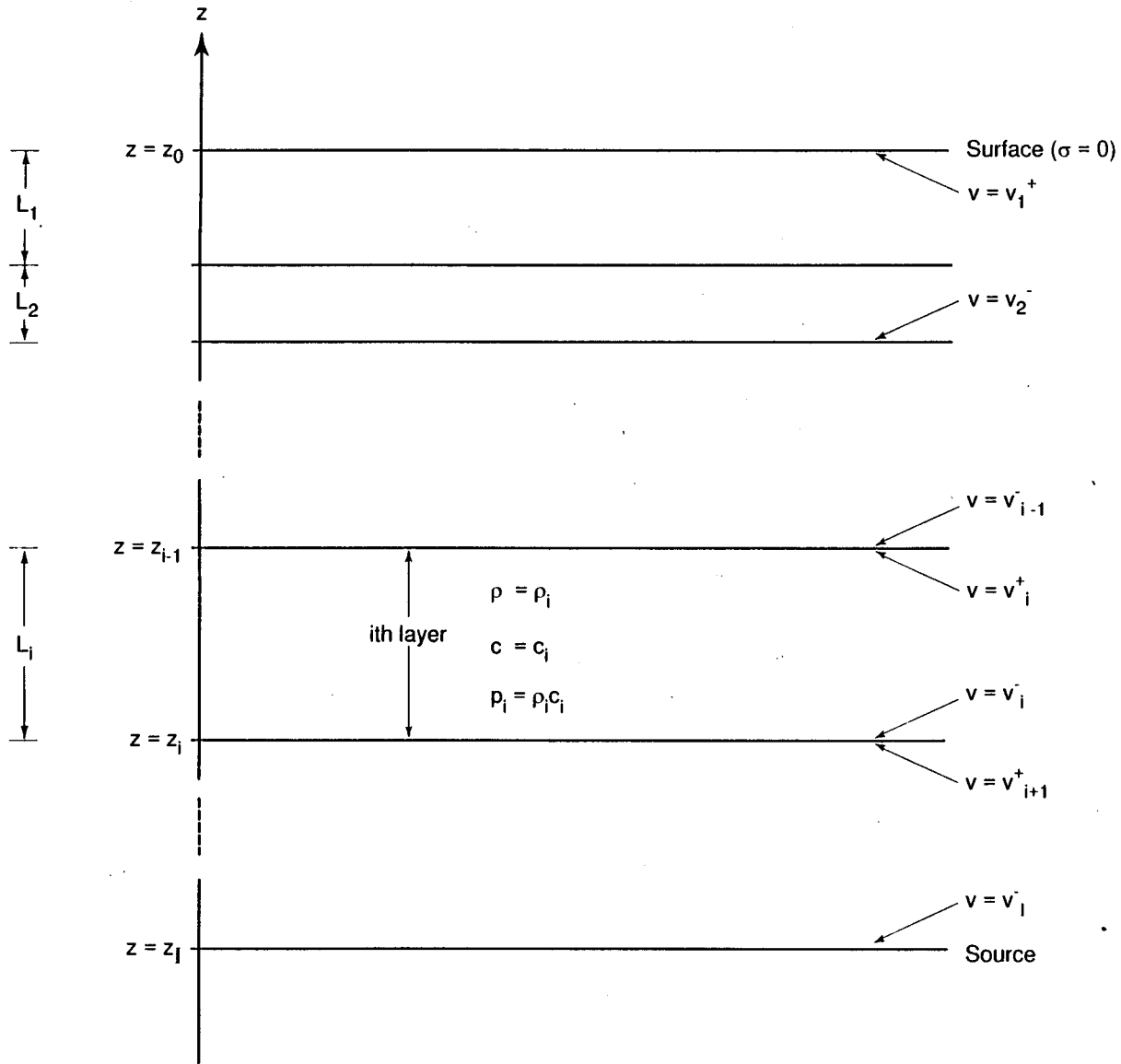


FIG. 1. The geometry and notation of the fractured medium. The ground surface is at  $z_0$ . Fractures are at  $z_i$  with  $i=1,2,\dots,I-1$ . The seismic source is at  $z=z_I$ . Mass densities and elastic moduli in the  $i$ th layer between  $z_{i-1}$  and  $z_i$  are  $\rho_i$  and  $c_i^2$ , respectively.

Sec. IV. Properties of the coda are also derived in this section. Some mathematical details of the coda analysis appear in Appendix C. Results are discussed in Sec. V.

### I. MODEL AND CONSTRAINT RELATIONS

We shall consider vertical shear wave propagation in a horizontally stratified earth material with fractures. The material is presumed to be initially free of stress and at rest. Each of the layers between fractures is taken to be uniform and linear in its response to stress. In the  $i$ th layer

$$\frac{\partial \sigma}{\partial t} = c_i^2 \frac{\partial \epsilon}{\partial t} = c_i^2 \frac{\partial v}{\partial z}, \quad \frac{\partial v}{\partial t} = \frac{\partial \sigma}{\partial z}, \quad (1)$$

where  $t$  is time,  $z$  is the ground coordinate,  $\sigma$  is the stress density,  $\epsilon$  is strain,  $v$  is shear displacement velocity, and  $c_i^2$  is a shear modulus in the layer. The fractures are located at  $z=z_k$ , for  $k=1,2,\dots,I-1$  (see Fig. 1). The surface is at

$z=z_0$ . The source of shear wave motion is at  $z=z_I$ . The density in the  $i$ th layer is indicated by  $\rho_i$ .

Each of the layers is acted on by body forces due to friction at the faults:

$$\frac{d}{dt} \int_{z=z_i}^{z=z_{i-1}} \rho_i v(z,t) dz = [F_{i-1,i}^* \operatorname{sgn}(\delta v_{i-1}) + f_{i-1,i}^* \delta v_{i-1}] - [F_{i,i+1}^* \operatorname{sgn}(\delta v_i) + f_{i,i+1}^* \delta v_i], \quad (2)$$

where  $F_{i,i+1}^*$  and  $f_{i,i+1}^*$  are positive coefficients of friction at the fault at  $z_i$  separating the  $i$ th and  $i+1$ st layer (except for  $F_{0,1}^* = f_{0,1}^* = 0$ , since the surface force is zero), and the  $\delta v_i$  is the difference in velocities above and below this fault. The  $\operatorname{sgn}$  functions take the value  $+1$  if their argument is positive and  $-1$  if it is negative. Their values for zero argument will be discussed in Sec. II where it will be shown that the  $F^{**}$ 's are coefficients of static friction. They define thresholds that must be exceeded if slippage is to occur: i.e.,  $|\delta v_i| > 0$ .

In the  $i$ th layer,  $v$  satisfies a wave equation with signal speed  $c_i$

$$\frac{\partial^2 v}{\partial t^2} = c_i^2 \frac{\partial^2 v}{\partial z^2}. \quad (3)$$

Superscripts “-” and “+” will be used to indicate values at lower and upper boundaries of a layer, respectively. Thus the  $i$ th layer (see Fig. 1) consists of  $z_i^- < z < z_i^+$ , and its thickness is

$$L_i = z_i^+ - z_i^-. \quad (4)$$

Note that  $z_{i+1}^+ = z_i^- = z_i$ . Velocities in the  $i$ th layer at upper and lower boundaries are written

$$\begin{aligned} v(z_i^+, t) &\equiv v_i(z_i^+, t) \equiv v_i^+(t), \\ v(z_i^-, t) &\equiv v_i(z_i^-, t) \equiv v_i^-(t). \end{aligned} \quad (5)$$

Inferring velocity at the source,  $z = z_I$ , from surface data requires determination of the velocities at the fractures. Consequently, determining the source velocity is a problem with  $2I - 1$  unknowns:  $v_i^-(t)$  and  $v_i^+(t)$  for  $i = 1, 2, \dots, I - 1$ .

The body forces on the first  $I - 1$  layers, Eq. (2), provide  $I - 1$  constraints for determining these  $2I - 1$  unknowns. An additional  $I$  constraints are provided by requiring that stress vanishes at the ground surface,  $z_1^+$ , and that stress is continuous across the  $I - 1$  fractures. The  $2I - 1$  mathematical constraint relations for the  $2I - 1$  unknowns are derived in Appendices A and B. They are Eq. (A6) of Appendix A, and Eqs. (B3), (B4), (B7), and (B8) of Appendix B.

We write these relations below in forms which are convenient for the treatments of Secs. II and III. The detailed composition of these expressions, displayed for completeness, is not needed for the treatments which follow. Only qualitative properties are used to derive the uniqueness of dynamics at the source determined from surface motion in Sec. III. For the analysis of Sec. IV, which treats the case of a single fracture and a deep source, many terms vanish, and much simpler expressions result.

The forms of Eqs. (A6) and (B7) are kept unchanged:

$$v_1^-(t) = \frac{1}{2} \left[ v_1^+ \left( t + \frac{L_1}{c_1} \right) + v_1^+ \left( t - \frac{L_1}{c_1} \right) \right] \equiv \langle v_1^+ \rangle(t), \quad (6)$$

$$\begin{aligned} p_1 \left[ -v_1^-(t) + 2 \sum_{n=1}^{\infty} (-1)^{n+1} v_1^- \left( t - 2n \frac{L_1}{c_1} \right) \right] \\ = F_{1,2}^* \operatorname{sgn}(v_1^-(t) - v_2^+(t)) + (v_1^-(t) - v_2^+(t)) f_{1,2}^*. \end{aligned} \quad (7)$$

Using Eq. (6) throughout the left-hand side of Eq. (7), one obtains

$$\begin{aligned} -\frac{p_1}{2} (\Delta v_1^+)(t) &\equiv -\frac{p_1}{2} \left[ v_1^+ \left( t + \frac{L_1}{c_1} \right) - v_1^+ \left( t - \frac{L_1}{c_1} \right) \right] \\ &= F_{1,2}^* \operatorname{sgn}(v_1^-(t) - v_2^+(t)) \\ &\quad + (v_1^-(t) - v_2^+(t)) f_{1,2}^*. \end{aligned} \quad (8)$$

It will be useful in Sec. IV to have noted that the left-hand side of Eq. (8) is the stress at the fracture at  $z = z_1$  [see Eq. (A8) of Appendix A].

Equation (B3), after isolating  $v_2^-(t - L_2/c_2)$ , which appears as the lead term in a summation, becomes

$$\begin{aligned} 2p_2 v_2^- \left( t - \frac{L_2}{c_2} \right) \\ = p_2 \left\{ v_2^+(t) - 2 \sum_{n=1}^{\infty} \left[ v_2^- \left( t - (2n+1) \frac{L_2}{c_2} \right) \right. \right. \\ \left. \left. - v_2^+ \left( t - 2n \frac{L_2}{c_2} \right) \right] \right\} \\ + p_1 \left[ v_1^-(t) - 2 \sum_{n=1}^{\infty} (-1)^{n+1} v_1^- \left( t - 2n \frac{L_1}{c_1} \right) \right]. \end{aligned} \quad (9)$$

Note that  $v_2^-(t - L_2/c_2)$  is given in terms of the velocities at the fracture at  $z = z_1$  at time  $t$  and times prior to  $t - L_2/c_2$ .

In analogy with Eq. (8), we isolate frictional force on the lower fault in Eq. (B8) to obtain

$$\begin{aligned} S_i(t) &\equiv +p_i \left\{ -v_i^-(t) + 2 \sum_{n=1}^{\infty} \left[ v_i^+ \left( t - (2n-1) \frac{L_i}{c_i} \right) \right. \right. \\ &\quad \left. \left. - v_i^- \left( t - 2n \frac{L_i}{c_i} \right) \right] \right\} - p_i \left\{ +v_i^+(t) - 2 \sum_{n=1}^{\infty} \right. \\ &\quad \left. \times \left[ v_i^- \left( t - (2n-1) \frac{L_i}{c_i} \right) - v_i^+ \left( t - 2n \frac{L_i}{c_i} \right) \right] \right\} \\ &\quad + [F_{i-1,i}^* \operatorname{sgn}(v_{i-1}^-(t) - v_i^+(t)) + (v_{i-1}^-(t) \\ &\quad - v_i^+(t)) f_{i-1,i}^*] \\ &= +[F_{i,i+1}^* \operatorname{sgn}(v_i^-(t) - v_{i+1}^+(t)) + (v_i^-(t) \\ &\quad - v_{i+1}^+(t)) f_{i,i+1}^*]. \end{aligned} \quad (10)$$

In analogy with Eq. (9), after isolating  $v_{i+1}^-(t - L_{i+1}/c_{i+1})$ , which appears as the lead term in a summation, Eq. (B4) becomes

$$\begin{aligned} 2p_{i+1} v_{i+1}^- \left( t - \frac{L_{i+1}}{c_{i+1}} \right) \\ = p_{i+1} \left\{ +v_{i+1}^+(t) - 2 \sum_{n=1}^{\infty} \left[ v_{i+1}^- \left( t - (2n+1) \frac{L_{i+1}}{c_{i+1}} \right) \right. \right. \\ \left. \left. - v_{i+1}^+ \left( t - 2n \frac{L_{i+1}}{c_{i+1}} \right) \right] \right\} - p_i \left\{ -v_i^-(t) + 2 \right. \\ \left. \times \sum_{n=1}^{\infty} \left[ v_i^+ \left( t - (2n-1) \frac{L_i}{c_i} \right) - v_i^- \left( t - 2n \frac{L_i}{c_i} \right) \right] \right\}. \end{aligned} \quad (11)$$

Comments analogous to those made after Eq. (9) apply.

Equations (6) and (8)–(11) are used in the treatments which follow. It should be noted that, at any given time, the summations appearing in these equations contain only a finite number of nonvanishing terms, since the ground is initially presumed to be at rest [i.e., for  $t < t_0$ ,  $v(z, t) = \sigma(z, t) = 0$ ].

## II. STATIC FRICTION THRESHOLDS

Equations (8) and (10) are of the form

$$k(t) = F^* \operatorname{sgn}[u(t)] + g([u(t)]), \quad (12)$$

where  $k(t)$  does not explicitly contain the differential slip velocity,  $u(t)$ ,  $F^*$  is a positive constant,  $g(u)$  is a continuous function of  $u$  satisfying  $g(0)=0$ , and  $\operatorname{sgn}(g)=\operatorname{sgn}(u)$  if  $u$  is not zero, and  $\operatorname{sgn}(u)=+1$  if  $u>0$  and  $-1$  if  $u<0$ . The value of the  $\operatorname{sgn}(u)$  if  $u=0$  will be defined below [see (iii)].

Suppose now that  $u$  takes a value different than zero. Multiplying Eq. (12) by  $\operatorname{sgn}[u(t)]$  (either plus or minus one) yields

$$k(t)\operatorname{sgn}[u(t)] - F^* = |g([u(t)])|. \quad (13)$$

From Eq. (13), we conclude that

- (i) it is necessary and sufficient that  $|k(t)| > F^*$  for  $u(t)$  to be different from zero;
  - (ii) if  $|k(t)| > F^*$ , then  $\operatorname{sgn}[u(t)] = \operatorname{sgn}[k(t)]$ ,
- and, therefore, from Eq. (12)
- (iii) if  $|k(t)| \leq F^*$ , then  $u(t) = 0$ , and  $F^* \operatorname{sgn}[u(t)] \equiv k(t)$  [since  $g(0) = 0$ ].

In Eqs. (8) and (10),  $u(t)$  is the difference in velocities above and below a fracture. Consequently, the constants  $F_{i,i+1}^*$  are threshold values that must be exceeded by  $\sigma_i^-(t)$  (if  $i=1$ ) or  $|S_i(t)|$  (if  $i=2,3,\dots,I-1$ ) if slippage is to occur at the  $i$ th fracture.

It should be noted that the arguments used in the derivation of (i)–(iii) apply equally well if  $F^*$  depends on  $u$ : e.g., if  $F^*$  has different values,  $F^{*+}$  for  $u=0+$  and  $F^{*-}$  for  $u=0-$ , or if  $F^*$  contains a factor which restricts its influence to small values of  $|u|$  (e.g.,  $\exp[-(u(t)/\mu U_0)^2]$ ; where  $|\mu| \leq 1$  and  $U_0$  is a scale velocity).

## III. SURFACE AND SOURCE VELOCITIES

It will now be shown that for the configurations described in Sec. I and illustrated in Fig. 1, the velocity at the source,  $v_1^-(t)$ , is uniquely determined by a measured ground motion,  $v_1^+(t)$ .

The procedure uses qualitative properties of Eqs. (6) and (8)–(11) and consists of two parts. In the first part it is shown how  $v_1^+(t)$  uniquely determines the velocities,  $v_1^-(t)$  and  $v_2^+(t)$  at the fracture between the first and second layers, and the velocity,  $v_2^-(t)$ , at the lower boundary of the second layer. The second part consists of showing that if the velocity is known in the first  $i$  layers, the velocities at the fracture between the  $i$ th and  $i+1$ st layers and at the lower boundary of the  $i+1$ st layer are uniquely determined. Since the surface velocity will have been shown, in the first part, to uniquely determine the velocity throughout the first two layers, the result of second part will imply, by induction, that the velocity at the source,  $v_I^-(t)$ , is uniquely determined by the surface velocity for any number of fractures. [Determination of source dynamics from  $v_1^-(t)$  is then completed by boundary conditions at the source: e.g., an outgoing radiation condition for propagation to the source for signals from the  $I-1$ st fault. Prior to the arrival of such signals,  $v_I^-(t)$  is the source velocity.]

Equation (6) gives  $v_1^-(t)$  for all times uniquely in terms of values of  $v_1^+$ , so knowledge of the surface motion determines the velocity throughout the first layer.

Since  $v_1^-$  is known for all times in terms of values of  $v_1^+$ , Eq. (8) may be treated as an equation for  $v_2^+(t)$  involving only values of  $v_1^+$ . Since Eq. (8) is an equation of the type treated in Sec. II, it follows that the necessary and sufficient condition for threshold to be exceeded at time  $t$  is that  $p_1|(\Delta v_1^+)(t)|/2 > F_{1,2}^*$ . If threshold is exceeded,  $\operatorname{sgn}[v_1^-(t) - v_2^+(t)] = -\operatorname{sgn}[(\Delta v_1^+)(t)]$ . Defining the nonnegative ‘‘threshold excess’’ to be

$$\delta F_{1,2}(t) \equiv \left( \frac{p_1}{2} |(\Delta v_1^+)(t)| - F_{1,2}^* \right) \times \begin{cases} 1 & \text{if } \frac{p_1}{2} |(\Delta v_1^+)(t)| - F_{1,2}^* \geq 0 \\ 0 & \text{if } \frac{p_1}{2} |(\Delta v_1^+)(t)| - F_{1,2}^* \leq 0 \end{cases}, \quad (14)$$

it follows that

$$v_1^-(t) - v_2^+(t) = -\frac{\delta F_{1,2}(t)}{f_{1,2}^*} \operatorname{sgn}(\Delta v_1^+)(t) \quad (15)$$

or

$$v_2^+(t) = \langle v_1^+ \rangle(t) + \frac{\delta F_{1,2}(t)}{f_{1,2}^*} \operatorname{sgn}(\Delta v_1^+)(t). \quad (16)$$

Consequently,  $v_2^+(t)$  is uniquely determined for all times in terms of values of  $v_1^+$ .

Therefore, using Eq. (16) in Eq. (9) uniquely determines  $v_2^-(t)$  in terms of its values at earlier times (e.g., for sufficiently early times, all these values vanish) and values of  $v_1^+$ .

This completes the first part of the proof.

For the case of a single fracture, or two-layer configuration,  $v_2^-$  is the velocity at the source, which has, therefore, been uniquely determined. For multiple fractures, assume that  $v_{i-1}^-$ ,  $v_i^+$ , and  $v_i^-$  have been uniquely determined for all times (e.g.,  $v_1^-$ ,  $v_2^+$ , and  $v_2^-$ ). To show that these uniquely determine  $v_{i+1}^+$  and  $v_{i+1}^-$ , one simply parallels the procedure of the first part using Eq. (10) [which is of the same form as Eq. (8)] to determine  $v_{i+1}^+(t)$  and then Eq. (11) [which is of the same form as Eq. (9)] to determine  $v_{i+1}^-(t - L_{i+1}/c_{i+1})$  for all times.

Consequently, the source velocity,  $v_I^-(t)$ , is uniquely determined for all times as a functional of the surface velocity.

## IV. THE CASE OF A SINGLE FRACTURE

For the case of a single fracture, the velocity at the source,  $v_2^-(t)$ , is determined from measurements of the surface velocity,  $v_1^+(t)$ , using Eq. (9), together with Eqs. (14) and (16). Taking the surface coordinate to be  $z=0$ , the fracture location is  $z=-L_1$  and the source location is  $z=-(L_1+L_2)$  (see Fig. 1). Using a clock set to  $t=0$  when the first indication of source motion arrives at the fracture, source motion begins at  $t=-L_2/c_2$ , and the first motion at

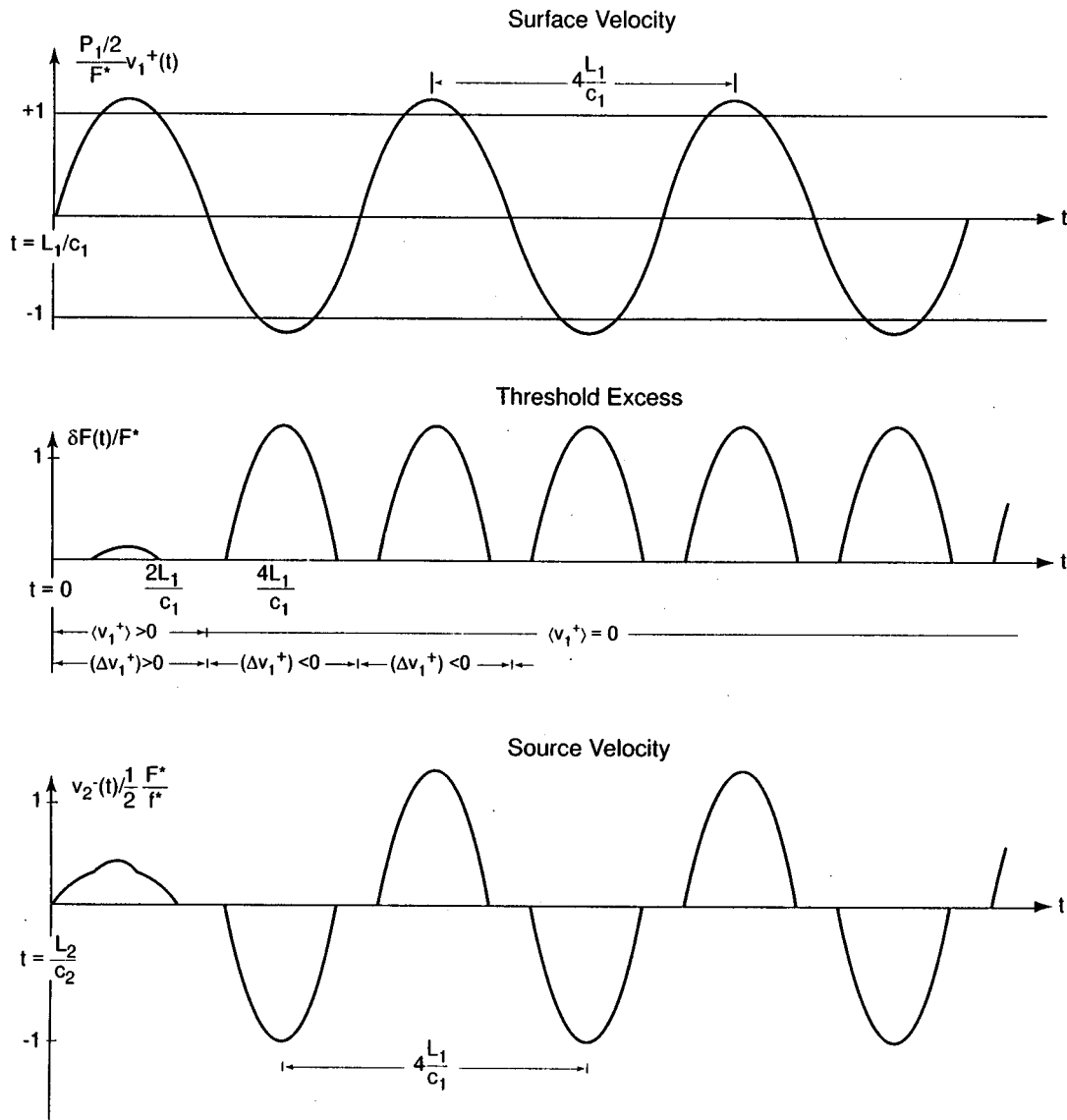


FIG. 2. Illustrating the relations between a hypothetical periodic measured surface velocity initiated at  $t=L_1/c_1$ , slippage and the uniquely inferred source velocity for the case of a single fracture and a deep source. The period of the assumed ground motion, initiated at  $t=L_1/c_1$ , is  $4L_1/c_1$ .

the surface occurs at  $t=L_1/c_1$ . For simplicity, we assume a deep source (i.e.,  $L_2/c_2 \gg L_1/c_1$ ) and consider times earlier than  $2L_2/c_2$ . This implies that all  $v_2^+$  and  $v_2^-$  terms under the summation sign in Eq. (9) vanish for the times under consideration. Consequently,

$$\begin{aligned}
 2v_2^-\left(t - \frac{L_2}{c_2}\right) &= v_2^+(t) + \frac{p_1}{p_2}v_1^-(t) \\
 &= \left(1 + \frac{p_1}{p_2}\right)\langle v_1^+\rangle(t) + \frac{\delta F_{1,2}(t)}{f_{1,2}^*} \operatorname{sgn}(\Delta v_1^+)(t).
 \end{aligned}
 \tag{17}$$

### A. Examples of surface and source velocity pairs

The relation between the threshold condition and source and surface velocities is illustrated for two hypothetical “measured” periodic surface velocities (not necessarily sinusoidal) after  $t=L_1/c_1$ , with periods four times (see Fig. 2) and twice (see Fig. 3) the transit time,  $L_1/c_1$ , for signals

propagating between the surface and fracture. For  $t > 2L_1/c_1$ , in the longer period example we have required  $v_1^+(t+L_1/c_1) = -v_1^+(t-L_1/c_1)$ , while for the shorter period example we have required  $v_1^+(t+L_1/c_1) = +v_1^+(t-L_1/c_1)$ . These illustrations show that the effects of slippage are rate, as well as amplitude, dependent. The illustrations use  $p_1/p_2 \ll 1$  in the graphs of source velocity to emphasize the effect of the threshold. For the times shown, all terms in the expression for the source velocity associated with  $t > 2L_2/c_2$  are assumed to vanish.

The longer period example corresponds to destructive interference of the surface motion with reflections from the fracture. In this case periodic slippage (see graph of  $\delta F_{1,2}$  in Fig. 2) at the fracture persists if the amplitude of the surface motion is sufficiently large:  $(\Delta v_1^+)(t) (= 2v_1^+(t-L_1/c_1))$  for  $t > 2L_1/c_1$  and  $\delta F_{1,2}(t)$  vary in time. However,  $\langle v_1^+\rangle(t)$  and  $v_1^-(t)$  become zero. The stress just above the fracture,  $\sigma_1^-(t)$ , is time varying. No power flows into or out

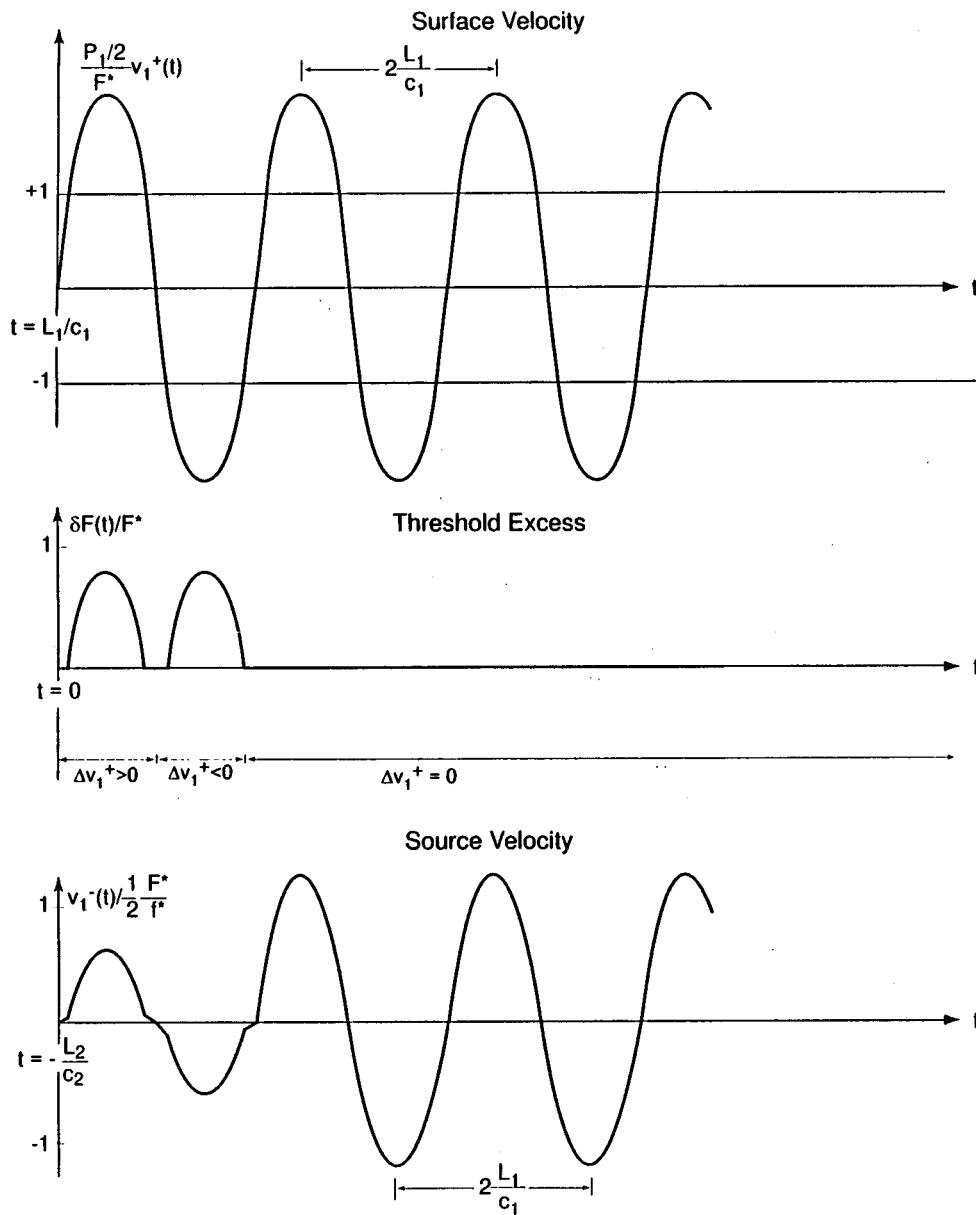


FIG. 3. Illustrating the relations between a hypothetical periodic measured surface velocity initiated at  $t = L_1/c_1$ , slippage and the uniquely inferred source velocity for the case of a single fracture and a deep source. The period of the assumed ground motion, initiated at  $t = L_1/c_1$ , is  $2L_1/c_1$ .

of the region between the ground surface and fracture. Slippage at the fracture can still be driven by the source.

For the shorter period example, the thickness of the layer between the fracture and surface corresponds to resonance, or constructive interference, between the surface motion and reflections from the fracture, and vice versa. The stresses immediately above and below the fracture,  $\sigma_1^-(t)$ , become zero, but the stress gradient,  $(\partial\sigma/\partial z)_1^-(t)$ , does not. Again, for the assumed motion, power neither flows into or out of the region between the fracture and ground surface. The velocity,  $v_1^-(t)$  becomes periodic, and slippage very quickly terminates, independently of the amplitude of surface motion: Both  $(\Delta v_1^+(t))$  and  $\delta F_{1,2}(t)$  vanish after a single period of the surface oscillation (see graph of  $\delta F_{1,2}$  in Fig. 3).

In both cases, discontinuities in acceleration at the source are generated at times corresponding to the onset and termination of slippage at the fracture because surface motions with continuous accelerations have been assumed, and

static friction at the fracture produces discontinuities in relative accelerations. Conversely, if source accelerations are continuous and slippage occurs, discontinuities in surface accelerations will be observed at times associated with the onset and termination of slippage.

### B. Surface codas

Equation (17) is in a form which facilitates an analysis of codas. Considering long time intervals after source termination for which  $v_2^-(t - L_2/c_2) = 0$ , Eq. (17) becomes

$$0 = \left(1 + \frac{p_1}{p_2}\right) \langle v_1^+ \rangle(t) + \frac{\delta F_{1,2}(t)}{f_{1,2}^*} \text{sgn}(\Delta v_1^+)(t). \quad (18)$$

Equation (18) implies that at or below threshold, since  $\delta F_{1,2} = 0$ ,

$$v_1^+ \left( t + \frac{L_1}{c_1} \right) = -v_1^+ \left( t - \frac{L_1}{c_1} \right), \quad (19)$$

while at or above threshold

$$v_1^+ \left( t + \frac{L_1}{c_1} \right) = - \frac{[(1+p_1/p_2) - p_1/f_{1,2}^*]}{[(1+p_1/p_2) + p_1/f_{1,2}^*]} v_1^+ \left( t - \frac{L_1}{c_1} \right) + \frac{2(F_{1,2}^*/f_{1,2}^*)}{[(1+p_1/p_2) + p_1/f_{1,2}^*]} \text{sgn}(\Delta v_1^+)(t). \quad (20)$$

From Eq. (19), it is evident that if  $|(\Delta v_1^+)(t)|$  is below threshold on a time interval of length  $2L_1/c_1$ , then it remains below threshold, and the velocity is periodic with period  $4(L_1/c_1)$ . Since Eq. (19) implies that  $\langle v_1^+ \rangle(t) = 0$ , the velocity just above the fracture,  $v_1^-(t)$ , also vanishes. Consequently, no power flows into or out of the region between the fracture and ground surface vanishes at either boundary [i.e.,  $v_1^-(t) = \sigma_1^+(t) = 0$ ], so energy is conserved, and periodic motion results.

These conclusions remain valid if  $|(\Delta v_1^+)(t)|$  is either below or at threshold throughout a time interval of length  $2L_1/c_1$ . Moreover, the class of periodic motions for this latter case are attractors if  $|(\Delta v_1^+)(t)|$  exceeds threshold somewhere on the time interval. Heuristically, one expects frictional drag to bring the velocity to threshold or below, so Eq. (19) becomes a good approximation to the velocity evolution. The mathematical proof given in Appendix C shows how this occurs, and reveals the elegant dependence of the periodic motion on its earlier time series.

Note that since Eq. (19) is satisfied, if the time series for the periodic coda is expanded in a Fourier series in  $\exp[2\pi ni(4L_1/c_1)^{-1}t]$ , only odd values of  $n$  appear.

## V. SUMMARY AND DISCUSSION

One-dimensional shear wave propagation in elastic earth materials with an arbitrary number of fractures was studied using a propagation model which included the effects of stress gradients and both static and sliding friction. It was shown for the friction model used (i.e., constant coefficients of static and sliding friction) that a history of the ground surface motion, together with a knowledge of fracture and source locations, allows a unique determination of dynamics at the source. The determined source dynamics is stable against small errors in measurements of surface motion. Since the number of fractures is arbitrary, one may apply the results of the analysis to materials of arbitrary fracture density. The inverse problem for any "rock" so constructed has a unique solution. Since the corresponding inverse problem for a hysteretic rock does not have a unique solution, hysteresis cannot be attributed to static friction thresholds.

The relation between source and surface histories was studied in detail for the case of a single fracture and a much deeper source. Comparing the different effects of slippage in the two examples used in the illustrations shows that slippage at faults and the coupling between surface and source velocities is rate dependent. A large amplitude oscillating source may not drive slippage at a fracture, while a smaller

amplitude source at a different frequency may. (This is because a sufficiently slow source will produce small stress gradients.) If the region between source and surface were to be modeled as a single entity, the bulk modulus of any assumed state relation would necessarily be frequency, as well as amplitude, dependent. The reason for the need to include rate dependence in the modeling on the bulk modulus is the failure of slippage at the fault to terminate or time average to zero on a time scale which is short compared to that of the surface motion.

For the case of a single fracture and a very deep source, the coda for the surface motion was shown to be either periodic or asymptotic to a periodic function in time. The quarter period of the coda is the signal transit time between the fracture and the surface. Even multiples of the fundamental frequency of the periodic function do not appear. The spectral amplitudes of the odd harmonics depend on the source history.

## ACKNOWLEDGMENTS

The author wishes to thank J. N. Albright and C. W. Myers for their encouragement and P. A. Johnson and J. Ten Cate for helpful technical discussions. This work was supported by the U.S. Department of Energy Office of Basic Energy Research: Scientific Computing.

## APPENDIX A: VELOCITIES AND STRESSES AT LAYER BOUNDARIES

In each layer, the velocity may be written as a sum over upward and downward propagating components satisfying a wave equation with constant signal speed, Eq. (3) of Sec. I.

### Layer $i=1$

At the stress-free surface of the uppermost layer, derivatives of the sum of the components must satisfy

$$\frac{\partial}{\partial z} v_1(z_1^+, t) = 0. \quad (A1)$$

In this layer, the velocity is given in terms of its velocity at the fault at  $z=z_1^-$  by

$$v(z, t) = v_1^- \left( t - \frac{z - z_1^-}{c_1} \right) + \sum_{n=1}^{\infty} (-1)^{n+1} v_1^- \left( t - \frac{(z_1^- + 2nL_1) - z}{c_1} \right) + \sum_{m=1}^{\infty} (-1)^m v_1^- \left( t - \frac{z - (z_1^- - 2mL_1)}{c_1} \right). \quad (A2)$$

Note that the right hand side of Eq. (A2) becomes  $v_1^-(t)$  at  $z=z_1^-$ . Verification that the stress vanishes at  $z=z_1^+$  follows from Eq. (A7) below.

Using Eq. (1) of Sec. I,

$$\frac{\partial \sigma}{\partial t} = c_1^2 \frac{\partial v}{\partial z} = c_1 \frac{\partial}{\partial t} \left[ -v_1^- \left( t - \frac{z - z_1^-}{c_1} \right) + \sum_{n=1}^{\infty} (-1)^{n+1} v_1^- \left( t - \frac{(z_1^- + 2nL_1) - z}{c_1} \right) - \sum_{m=1}^{\infty} (-1)^m v_1^- \left( t - \frac{z - (z_1^- - 2mL_1)}{c_1} \right) \right]. \quad (\text{A3})$$

The motion at the ground surface,  $z = z_1^+$ , satisfies

$$v_1^+(t) = v_1^- \left( t - \frac{L_1}{c_1} \right) + \sum_{n=1}^{\infty} (-1)^{n+1} \times \left[ v_1^- \left( t - \frac{(2n-1)L_1}{c_1} \right) - v_1^- \left( t - \frac{(2n+1)L_1}{c_1} \right) \right] = 2v_1^- \left( t - \frac{L_1}{c_1} \right) + 2 \sum_{n=1}^{\infty} (-1)^n v_1^- \left( t - \frac{(2n+1)L_1}{c_1} \right). \quad (\text{A4})$$

Replacing  $t$  by  $t + L_1/c_1$  and  $t - L_1/c_1$  in Eq. (A2) yields

$$v_1^+ \left( t + \frac{L_1}{c_1} \right) = 2v_1^-(t) + 2 \sum_{n=1}^{\infty} (-1)^n v_1^- \left( t - 2n \frac{L_1}{c_1} \right), \quad (\text{A5})$$

$$v_1^+ \left( t - \frac{L_1}{c_1} \right) = 2v_1^- \left( t - 2 \frac{L_1}{c_1} \right) + 2 \sum_{n=1}^{\infty} (-1)^n \times v_1^- \left( t - 2(n+1) \frac{L_1}{c_1} \right) = -2 \sum_{n=1}^{\infty} (-1)^n v_1^- \left( t - 2n \frac{L_1}{c_1} \right).$$

Therefore,

$$v_1^-(t) = \frac{1}{2} \left[ v_1^+ \left( t + \frac{L_1}{c_1} \right) + v_1^+ \left( t - \frac{L_1}{c_1} \right) \right] \equiv \langle v_1^+ \rangle (t). \quad (\text{A6})$$

Integrating Eq. (A3) with respect to time between minus infinity and  $t$  yields

$$\frac{\partial \sigma}{\partial t} = c_i^2 \frac{\partial v}{\partial z} = c_i \frac{\partial}{\partial t} \left\{ -v_i^- \left( t - \frac{z - z_i^-}{c_i} \right) - \sum_{n=1}^{\infty} \left[ v_i^- \left( t - \frac{2nL_i + (z - z_i^-)}{c_i} \right) + v_i^- \left( t - \frac{2nL_i - (z - z_i^-)}{c_i} \right) \right] + v_i^+ \left( t - \frac{z_i^+ - z}{c_i} \right) + \sum_{m=1}^{\infty} \left[ v_i^+ \left( t - \frac{2mL_i + (z_i^+ - z)}{c_i} \right) + v_i^+ \left( t - \frac{2mL_i - (z_i^+ - z)}{c_i} \right) \right] \right\}. \quad (\text{A10})$$

Integrating with respect to time as before, one obtains

$$\sigma_i^-(t) \equiv \sigma(z_i^-, t) = c_i \left\{ -v_i^-(t) + 2 \sum_{n=1}^{\infty} \left[ v_i^- \left( t - (2n-1) \frac{L_i}{c_i} \right) - v_i^- \left( t - 2n \frac{L_i}{c_i} \right) \right] \right\}, \quad (\text{A11})$$

$$\sigma(z, t) = c_1 \left[ -v_1^- \left( t - \frac{z - z_1^-}{c_1} \right) + \sum_{n=1}^{\infty} (-1)^{n+1} v_1^- \left( t - \frac{(z_1^- + 2nL_1) - z}{c_1} \right) - \sum_{m=1}^{\infty} (-1)^m v_1^- \left( t - \frac{z - (z_1^- - 2mL_1)}{c_1} \right) \right]. \quad (\text{A7})$$

At  $z = z_1^-$

$$\sigma(z_1^-, t) \equiv \sigma_1^-(t) = c_1 \left[ -v_1^-(t) + 2 \sum_{n=1}^{\infty} (-1)^{n+1} v_1^- \left( t - 2n \frac{L_1}{c_1} \right) \right] = -\frac{c_1}{2} \left[ v_1^+ \left( t + \frac{L_1}{c_1} \right) - v_1^+ \left( t - \frac{L_1}{c_1} \right) \right] \equiv -\frac{c_1}{2} (\Delta v_1^+)(t). \quad (\text{A8})$$

At  $z = z_1^+$ , the right-hand side of Eq. (A7) vanishes.

### Layers $i = 2, 3, \dots, I$

The velocity in all other layers is given in terms of values at the faults (and in the case  $i = I$ , the source) by

$$v(z, t) = v_i^- \left( t - \frac{z - z_i^-}{c_i} \right) + \sum_{n=1}^{\infty} \left[ v_i^- \left( t - \frac{2nL_i + (z - z_i^-)}{c_i} \right) - v_i^- \left( t - \frac{2nL_i - (z - z_i^-)}{c_i} \right) \right] + v_i^+ \left( t - \frac{z_i^+ - z}{c_i} \right) + \sum_{m=1}^{\infty} \left[ v_i^+ \left( t - \frac{2mL_i + (z_i^+ - z)}{c_i} \right) - v_i^+ \left( t - \frac{2mL_i - (z_i^+ - z)}{c_i} \right) \right], \quad (\text{A9})$$

## APPENDIX B: CONSTRAINTS AT FAULTS

We require stress to be continuous across fractures. Since  $\sigma$ , in each layer, is stress divided by density, defining

$$p_i \equiv \rho_i c_i, \quad (\text{B1})$$

continuity of stress requires that for  $i=1,2,\dots,I-1$ ,

$$\rho_i \sigma_i^-(t) = \rho_{i+1} \sigma_{i+1}^+(t). \quad (\text{B2})$$

At the first fracture (between the first and second layers), Eqs. (A8) and (A11) of Appendix A give

$$\begin{aligned} p_1 & \left[ -v_1^-(t) + 2 \sum_{n=1}^{\infty} (-1)^{n+1} v_1^- \left( t - 2n \frac{L_1}{c_1} \right) \right] \\ & = p_2 \left\{ +v_2^+(t) - 2 \sum_{n=1}^{\infty} \left[ v_2^- \left( t - (2n-1) \frac{L_2}{c_2} \right) \right. \right. \\ & \quad \left. \left. - v_2^+ \left( t - 2n \frac{L_2}{c_2} \right) \right] \right\}. \quad (\text{B3}) \end{aligned}$$

At all other fractures between the  $i$ th and  $i+1$ st layer, Eq. (A11) gives

$$\begin{aligned} p_i & \left\{ -v_i^-(t) + 2 \sum_{n=1}^{\infty} \left[ v_i^+ \left( t - (2n-1) \frac{L_i}{c_i} \right) - v_i^- \left( t - 2n \frac{L_i}{c_i} \right) \right] \right\} \\ & = p_{i+1} \left\{ +v_{i+1}^+(t) - 2 \sum_{n=1}^{\infty} \left[ v_{i+1}^- \left( t - (2n-1) \frac{L_{i+1}}{c_{i+1}} \right) \right. \right. \\ & \quad \left. \left. - v_{i+1}^+ \left( t - 2n \frac{L_{i+1}}{c_{i+1}} \right) \right] \right\}. \quad (\text{B4}) \end{aligned}$$

Using the momentum conservation component of Eq. (1) of Sec. I

$$\frac{\partial v}{\partial t} = \frac{\partial \sigma}{\partial z} \quad (\text{B5})$$

in Eq. (3) of that section, to replace the time derivative of  $v$  with the space derivative of  $\sigma$ ,

$$\begin{aligned} \frac{d}{dt} \int_{z=z_i}^{z=z_{i-1}} \rho_i v(z,t) dz & \\ & = \rho_i [\sigma_i^+(t) - \sigma_i^-(t)] \\ & = [F_{i-1,i}^* \text{sgn}(v_{i-1}^-(t) - v_i^+(t)) \\ & \quad + (v_{i-1}^-(t) - v_i^+(t)) f_{i-1,i}^*] - [F_{i,i+1}^* \text{sgn}(v_i^-(t) \\ & \quad - v_{i+1}^+(t)) + (v_i^-(t) - v_{i+1}^+(t)) f_{i,i+1}^*]. \quad (\text{B6}) \end{aligned}$$

For the first layer, Eqs. (A8) and (B6), together with  $F_{0,1}^* = f_{0,1}^* = \sigma_1^+ = 0$  gives

$$\begin{aligned} p_1 & \left[ -v_1^-(t) + 2 \sum_{n=1}^{\infty} (-1)^{n+1} v_1^- \left( t - 2n \frac{L_1}{c_1} \right) \right] \\ & = F_{1,2}^* \text{sgn}(v_1^-(t) - v_2^+(t)) + (v_1^-(t) - v_2^+(t)) f_{1,2}^*. \quad (\text{B7}) \end{aligned}$$

For the other layers bounded by fractures ( $i=2,3,\dots,I-1$ ), Eqs. (A11) and (B6) give

$$\begin{aligned} p_i & \left\{ +v_i^+(t) - 2 \sum_{n=1}^{\infty} \left[ v_i^- \left( t - (2n-1) \frac{L_i}{c_i} \right) - v_i^+ \left( t - 2n \frac{L_i}{c_i} \right) \right] \right\} \\ & - p_i \left\{ -v_i^-(t) + 2 \sum_{n=1}^{\infty} \left[ v_i^+ \left( t - (2n-1) \frac{L_i}{c_i} \right) - v_i^- \left( t - 2n \frac{L_i}{c_i} \right) \right] \right\} \\ & = [F_{i-1,i}^* \text{sgn}(v_{i-1}^-(t) - v_i^+(t)) + (v_{i-1}^-(t) - v_i^+(t)) f_{i-1,i}^*] - [F_{i,i+1}^* \text{sgn}(v_i^-(t) - v_{i+1}^+(t)) \\ & \quad + (v_i^-(t) - v_{i+1}^+(t)) f_{i,i+1}^*]. \quad (\text{B8}) \end{aligned}$$

Equations (A6) and (B3), (B4), (B7), and (B8) provide  $2I-1$  nonlinear algebraic constraint relations for the  $2I-1$  unknowns,  $v_i^-$  (for  $i=1,2,\dots,I$ ) and  $v_i^+(t)$  (for  $i=2,3,\dots,I$ ) of the inverse problem treated in Sec. II.

## APPENDIX C: PERIODICITY OF CODAS

It was shown in Sec. IV that the coda is periodic if  $|(\Delta v_1^+)|$  is below threshold on an entire closed time interval of length  $2L_1/c_1$ . In this Appendix we show that the coda is periodic with the same periodicity even if  $|(\Delta v_1^+)(t)|$  is either at or below threshold on such a time interval and that the class of the periodic motions for this latter case are attractors

if  $|(\Delta v_1^+)(t)|$  exceeds threshold somewhere on the time interval. [It is obvious that a periodic solution can be constructed using Eq. (19) if  $|(\Delta v_1^+)(t)|$  is either at or below threshold on such a time interval. However, it must be shown for this case that  $|(\Delta v_1^+)(t)|$  cannot later exceed threshold: i.e., that the construction describes a unique continuation of the motion.]

The proof is in two parts. First it will first be shown that  $|(\Delta v_1^+)(t)|$  cannot remain at or above threshold for all times. Then it will then be shown that if it is ever below threshold, any excursions above threshold occur periodically in time but with decreasing in amplitude. Asymptotically, the velocity has period  $4(L_1/c_1)$ .



Consideration of the various combinations of signs in Eq. (20) shows that for any time interval on which  $|(\Delta v_1^+) \times(t)|$  is at or above threshold

$$\text{sgn}[\delta|(\Delta v_1^+)(t)|] = -\text{sgn}[(\Delta v_1^+)(t)] \cdot \text{sgn}\left[\delta v_1^+\left(t - \frac{L_1}{c_1}\right)\right], \quad (\text{C1})$$

in which  $\text{sgn}(0)=0$ , and for any function  $q$  and positive time differential  $\delta t$ , we define  $\delta q(t) \equiv q(t + \delta t) - q(t)$ . Consequently, if  $|(\Delta v_1^+)(t_{\text{threshold}})|$  is at threshold, the prior history of  $v_1^+$  determines whether  $\delta|(\Delta v_1^+)(t)| > 0$  [i.e., whether  $|(\Delta v_1^+)(t)|$  will move above threshold for neighboring positive times]. In particular, if

$$\left|v_1^+\left(t_{\text{threshold}} \pm \frac{L_1}{c_1}\right)\right| = \frac{F_{1,2}^*}{p_1}$$

(with, of course, the two values of  $v_1^+$  having opposite sign), then for  $|(\Delta v_1^+)(t)|$  to move above threshold, it is necessary and sufficient that

$$\delta\left|v_1^+\left(t_{\text{threshold}} - \frac{L_1}{c_1}\right)\right| > 0,$$

i.e., that  $|v_1^+(t - L_1/c_1)| > F_{1,2}^*/p_1$ .

The first part of the proof is by contradiction. Assume that  $|(\Delta v_1^+)(t)|$  is at or above threshold on all of a time interval  $[t_\alpha, t_\alpha + 2L_1/c_1]$  for some  $t_\alpha$  later than the last time a signal from the active source arrives at the fracture. Evaluating Eq. (20) at the endpoints of the interval and forming the difference yields, using  $\text{sgn}(\Delta v_1^+)(t) = \text{constant}$  on the interval [since, by assumption  $|(\Delta v_1^+)(t)|$  does not fall below threshold],

$$\begin{aligned} (\Delta v_1^+)\left(t_\alpha + 2\frac{L_1}{c_1}\right) &\equiv (\Delta v_1^+)\left(t_\alpha + 3\frac{L_1}{c_1}\right) - (\Delta v_1^+)\left(t_\alpha + \frac{L_1}{c_1}\right) \\ &= -\frac{[(1+p_1/p_2) - p_1/f_{1,2}^*]}{[(1+p_1/p_2) + p_1/f_{1,2}^*]} (\Delta v_1^+)(t_\alpha). \end{aligned} \quad (\text{C2})$$

If it is at or above threshold on  $n$  consecutive intervals, by induction,

$$\begin{aligned} (\Delta v_1^+)\left(t_\alpha + 2n\frac{L_1}{c_1}\right) &= \left(-\frac{[(1+p_1/p_2) - p_1/f_{1,2}^*]}{[(1+p_1/p_2) + p_1/f_{1,2}^*]}\right)^n \\ &\quad \times (\Delta v_1^+)(t_\alpha). \end{aligned} \quad (\text{C3})$$

Clearly, if  $n$  is sufficiently large the coefficient of  $|(\Delta v_1^+) \times(t_\alpha)|$  in Eq. (C3) becomes arbitrarily small, and  $|(\Delta v_1^+) \times(t)|$  is below threshold somewhere in  $[t_\alpha, t_\alpha + 2nL_1/c_1]$ .

This contradicts the assumption and completes the first part of the proof. Consequently, for some time  $t_\beta$ ,  $|(\Delta v_1^+) \times(t)|$  is below threshold somewhere in an interval  $[t_\beta, t_\beta + 2nL_1/c_1]$ .

For the second part of the proof, we use the result of the first part, and consider the nature of pairs of excursions above threshold. Suppose that after being below threshold for neighboring earlier times,  $|(\Delta v_1^+)(t_{\text{threshold}})|$  is at threshold. Then

$$\left|v_1^+\left(t_{\text{threshold}} \pm \frac{L_1}{c_1}\right)\right| = \frac{F_{1,2}^*}{p_1}.$$

If  $|(\Delta v_1^+)(t)|$  subsequently exceeds threshold, Eq. (20) implies

$$\begin{aligned} (\delta v_1^+)\left(t + \frac{L_1}{c_1}\right) &= -\frac{[(1+p_1/p_2) - p_1/f_{1,2}^*]}{[(1+p_1/p_2) + p_1/f_{1,2}^*]} \\ &\quad \times (\delta v_1^+)\left(t - \frac{L_1}{c_1}\right), \end{aligned} \quad (\text{C4})$$

so while  $|(\Delta v_1^+)(t)|$  is above threshold

$$\left|(\delta v_1^+)\left(t + \frac{L_1}{c_1}\right)\right| < \left|(\delta v_1^+)\left(t - \frac{L_1}{c_1}\right)\right|.$$

Clearly, if

$$\left(1 + \frac{p_1}{p_2}\right) - \frac{p_1}{f_{1,2}^*}$$

is positive, then the variations of  $v_1^+(t - L_1/c_1)$  and  $v_1^+(t - L_1/c_1)$  are of opposite sign, so

$$\left|v_1^+\left(t \pm \frac{L_1}{c_1}\right)\right| > \frac{F_{1,2}^*}{p_1}$$

and

$$\left|v_1^+\left(t + \frac{L_1}{c_1}\right)\right| - \frac{F_{1,2}^*}{p_1} < \left|v_1^+\left(t - \frac{L_1}{c_1}\right)\right| - \frac{F_{1,2}^*}{p_1}.$$

Thus excursions above threshold occur periodically in time with a periodicity equal to  $2L_1/c_1$ , and the amplitude of the excursions tend to zero exponentially in time. Since the signs of  $v_1^+(t + L_1/c_1)$  and  $v_1^+(t - L_1/c_1)$  are opposed,  $4L_1/c_1$  is an asymptotic period for the limiting motion.

If

$$\left(1 + \frac{p_1}{p_2}\right) - \frac{p_1}{f_{1,2}^*}$$

is negative, then the variations of  $v_1^+(t - L_1/c_1)$  and  $v_1^+(t - L_1/c_1)$  are of the same sign. Then in order to exceed threshold, it is necessary that

$$\left|v_1^+\left(t - \frac{L_1}{c_1}\right)\right| - \frac{F_{1,2}^*}{p_1} > 0$$

but

$$\left|v_1^+\left(t + \frac{L_1}{c_1}\right)\right| - \frac{F_{1,2}^*}{p_1} < 0.$$

Consequently, since  $|v_1^+((t + 2L_1/c_1) - L_1/c_1)| = |v_1^+(t + L_1/c_1)| < F_{1,2}^*/p_1$ ,  $|(\Delta v_1^+)(t + 2L_1/c_1)|$  cannot be above threshold. Consequently, Eq. (19) applies after a finite time, and  $4L_1/c_1$  is a period for the motion.

<sup>1</sup>P. A. Johnson, B. Zinszner, and P. N. J. Rasolofosaon, "Resonance and elastic nonlinear phenomena in rock," *J. Geophys. Res.* **101**, 11553–11564 (1996).

<sup>2</sup>P. A. Johnson and P. N. J. Rasolofosaon, "Manifestation of nonlinearity in rock: Convincing evidence over large frequency and strain levels from laboratory studies," *Nonlinear Processes in Geophysics* **3**, 77–88 (1996).

<sup>3</sup>R. A. Guyer, K. R. McCall, and G. N. Boitnott, "Hysteresis, discrete

memory and nonlinear wave propagation in rock,” *Phys. Rev. Lett.* **74**, 3491–3494 (1995).

<sup>4</sup>A. Kadish, P. A. Johnson, and B. Zinszner, “Evaluating hysteresis in earth materials under dynamic resonance,” *J. Geophys. Res.* **101**, 25139–25147 (1996).

<sup>5</sup>E. A. Coddington and N. Levinson, *Theory of Ordinary Differential Equations* (McGraw-Hill, New York, 1995), Chap. 1, Sec. 1.

*tions* (McGraw-Hill, New York, 1995), Chap. 1, Sec. 1.

<sup>6</sup>A. Kadish, “On the problem of source identification from far-field observations in hysteretic materials,” *J. Acoust. Soc. Am.* **103**, 3196–3202 (1998).

<sup>7</sup>M. Brokate and J. Sprekels, *Hysteresis and Phase Transitions* (Springer-Verlag, New York, 1996), Chap. 2.

# The secondary field of a parametric source following interaction with sea surface

Lian Sheng Wang, Brian V. Smith, and Rodney Coates

*School of Electronic and Electrical Engineering, University of Birmingham, Edgbaston, Birmingham B15 2TT, United Kingdom*

(Received 16 April 1997; accepted for publication 26 January 1999)

The effects of a flat sea surface on the secondary wave field for a nonsaturation limited parametric source with a piston-type transmit transducer have been assessed both theoretically and experimentally. When the virtual array of the parametric source interacts with the sea surface, the amplitude of the secondary signal is reduced. In addition to the normal interference between the direct and surface reflection paths, there is also one extra mechanism which causes the reduction of the secondary signal level. This is the destructive summation of the secondary field generated before and after intersection with the water surface. A numerical model and the Westervelt model are used to predict the secondary field which is a function of grazing angle. Experiments in an in-door laboratory tank have been carried out to measure the primary and secondary fields of a parametric array under these circumstances. It is shown that the surface reflection affects the beam pattern of the secondary field and the induced reduction in the signal level depends on the characteristics of the parametric source and the geometry of the problem. The level of reduction caused by the effects of the surface interaction approaches a constant at large grazing angles. © 1999 Acoustical Society of America. [S0001-4966(99)00206-4]

PACS numbers: 43.25.Lj, 43.30.Lz [DLB]

## INTRODUCTION

Since Westervelt<sup>1</sup> first published the theory of the parametric array, some 30 years ago, there has been much interest in this phenomenon. The significant property of the parametric array is that it produces a much narrower beam from an array, at the secondary frequency, than is possible by transmitting directly, at this frequency, in a conventional manner from the same sized array. Therefore the parametric array has the potential of producing narrow beams, at relative low frequencies, from small arrays. It is this feature which has been most commonly exploited in practice.<sup>2-6</sup> A comprehensive introduction of the applications of parametric sonars has been given by Konrad.<sup>7</sup> The characteristics of the parametric array are also very attractive for the design of underwater communication systems<sup>7-10</sup> that can be used in very shallow-water channels.

The main problems with acoustic communication in a shallow-water channel are signal fluctuations and severe inter-symbol interference, ISI, caused by multipath. When the channels are time invariant or slowly time varying with stationary properties, then the adverse effects of multipath can be reduced by employing various equalization methods at the receiver end. Unfortunately, in most cases the underwater channels are far away from the ideal, for example, channels with surface reflection path introduce rapid fluctuations in received signals, therefore it is very difficult to combat the multipath with signal processing methods only. Ideally, the boundary interaction could be avoided by using a narrow beam and, to achieve this in practice, an array with a large aperture in comparison with wavelength should be used. Although it is straightforward to realize this at high frequencies, because the array will be physically small, at low frequencies, which are essential for long-range commu-

nication, the array size becomes unmanageable if conventional transduction methods are employed. However, with a parametric array, as mentioned previously, a narrow beam can be obtained at low frequencies from a physically small array. In addition to the narrow beamwidth property, a parametric array can also offer a very wide bandwidth relative to the central frequency of the secondary signal. This is also very crucial if a data link with a relative high data rate is to be achieved at low frequencies. In summary, the parametric array is a suitable tool for obtaining long-range underwater acoustic communication in a shallow-water channel.

Since the secondary field of a parametric array is generated by the interaction of two primary frequency signals within the medium they ensonify, any change in the propagation conditions for the primary signals will affect the secondary signal. It is inevitable that the sound field will interact with both sea surface and bottom in shallow channels when long-range propagations are being sought, even with the use of narrow beams. After reflection, the primary and secondary waves are all subject to phase change at these boundaries, as predicted by conventional linear acoustic theory. However, if a boundary occurs in the region of the virtual array it is interesting to note that the secondary wave generated after the reflection from boundaries may have a different phase to that component before the reflection. For example, in the case of pressure-release boundary, such as sea surface, these phases are exactly opposite. Consequently, any difference in phases before and after the reflection of the virtual array will result in a change in the overall secondary signal level. When a boundary is rough, there will be some loss in the secondary signal level associated with the loss of coherence of the primary signal after reflection. In the extreme case where a boundary is very rough, it seems that the virtual array gen-

erating the secondary wave is essentially truncated and terminated after the reflection from which the primary signals are scattered randomly in all directions.

The effects of the pressure-release boundary on a parametric array have been investigated previously under a predominantly saturation limited condition in order to enhance the secondary signal level at long distances by reversing the distortion process of the primary wave within the virtual array of the parametric source to postpone acoustic saturation in finite-amplitude waveforms.<sup>11-14</sup> Partial cancellation in the secondary signal level has been confirmed experimentally and theoretically. Good agreement is shown between the measured on-axis secondary sound pressure level and that predicted by theory, except at the highest primary sound pressure level.<sup>14</sup>

In order to achieve reliable communication links, the parametric source for underwater acoustic communications should be used with a source level well below the threshold of shock wave limit to avoid possible distortion caused by saturation.<sup>15</sup> It is obvious there will be a reduction of the sound pressure level at the secondary frequency due to the surface reflection under such circumstances. It is necessary to estimate the possible signal level reduction caused by boundary interaction when a parametric source is used in a channel where boundary interaction within the length of virtual array is inevitable. Furthermore, knowledge about the secondary field as a function of grazing angle will be of great practical interest to the designer using the parametric transduction method in a shallow-water channel. In this paper, we will be concerned with these issues. In order to quantify the effects of this boundary interaction, some experimental and theoretical studies have been carried out in a simple scenario; a circular piston-type transmit transducer with its virtual array interacting with a flat water surface. Both the primary and secondary fields were measured in experiments in an indoor tank. The Westervelt model<sup>1</sup> and a numerical solution are used to calculate the secondary field to compare with the experimental results. It is shown from this study that the most important effects of the surface interaction are to reduce the on-axis secondary signal level and to result in a slow roll-off of the beam pattern of the secondary field. The Westervelt model tends to overestimate the level of reduction in comparison with the experimental results, while the numerical method shows close agreement at a grazing angle greater than 5 degrees.

## I. THEORY

In order to avoid possible wave form distortion caused by saturation, only the parametric transduction process in the absorption and spreading limited cases will be considered here. The secondary sound pressure field generated by two primary signals at frequency  $f_1$  and  $f_2$  in a free space is given by Westervelt<sup>1</sup>

$$P_d(\mathbf{R}, t) = \rho_0 \frac{\partial}{\partial t} \left[ \frac{1}{4\pi} \int_V \frac{q_d e^{-(\alpha + jk_d)|\mathbf{R} - \mathbf{R}'|}}{|\mathbf{R} - \mathbf{R}'|} dV \right], \quad (1)$$

where  $\rho_0$  is the density of the water,  $q_d$  is the source density function for the secondary wave,  $\alpha$  is the absorption coefficient

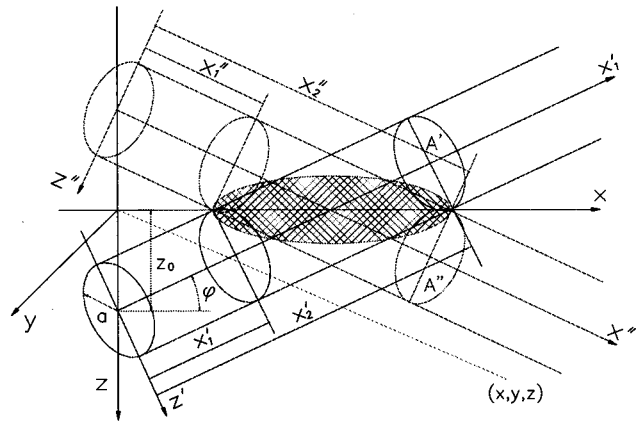


FIG. 1. A Westervelt parametric source reflected by the water surface.

at the secondary frequency,  $k_d = \omega_d / c_0$  is the wave number at the secondary frequency,  $\omega_d$  is the angular frequency of the secondary signal, and  $c_0$  is the sound speed in water. The integration in Eq. (1) is over the whole volume of space in which the nonlinear interaction takes place.

The secondary field of a parametric array in Eq. (1) can be evaluated mainly in two different ways depending on whether the nonlinear interaction taking place is within the collimated zone of the primary signals or in a spreading beam. The former is the absorption limited case, while the latter is the spreading limited case. The criteria for the two regions are determined by the parameter<sup>16</sup>

$$R_F = \frac{\alpha_T R_0 f_0}{f_d}, \quad (2)$$

where  $\alpha_T = \alpha_1 + \alpha_2$  is the sum of the absorption coefficients  $\alpha_1$  and  $\alpha_2$  at  $f_1$  and  $f_2$ , respectively,  $R_0 = d^2 / \lambda_0$  is the Rayleigh distance,  $\lambda_0$  is the wavelength of the primary signal,  $d$  is the aperture dimension of the source,  $f_0 = (f_1 + f_2) / 2$ , and  $f_d$  is the secondary frequency. For  $R_F \gg 1$ , the primary wave is highly attenuated within the collimated zone; then the Westervelt model<sup>1</sup> should be used. For  $R_F \ll 1$ , other models<sup>17-19</sup> have to be used to take the spherical spreading of the primary wave into account.

### A. Westervelt model

There are different expressions for the source density function, depending upon the parameters of the parametric source. The most simple form is for the collimated source, where the secondary signal is generated well within the Rayleigh distance at primary frequency. Assuming the acoustic axis is along the  $x$  direction, the source density function can be expressed as

$$q_d = j \frac{\omega_d P_1 P_2 \beta}{\rho_0^2 c_0^4} e^{-(\alpha_1 + \alpha_2)x} e^{j(\omega_d t - k_d x)}, \quad (3)$$

where  $P_1$  and  $P_2$  are the amplitudes of the sound pressure of the primary signals, and  $\beta$  is the parameter of nonlinearity.

Consider that the virtual array of a parametric source is intercepted by a water surface as shown in Fig. 1. The surface is chosen as the  $x$ - $y$  plane with the  $z$  coordinate pointing downward. For the sake of simplicity, a circular piston-type

transmit transducer is used as the source. It is placed below the water surface at depth  $z_0$ , and the acoustic axis of the transducer is in the  $x$ - $z$  plane with a grazing angle  $\varphi$  with respect to the surface. The area of the active face of the transducer is  $S_0 = \pi a^2$ , where  $a$  is the radius of the transducer. To satisfy the boundary condition at the surface, the total field generated by the transducer should be the summation of the real source and the corresponding virtual source, the image of the real source mirrored by the surface with an opposite sign, as demonstrated in Fig. 1. As a consequence of this reflection from the surface, the secondary field of the parametric array is formed by the summation of the two total virtual arrays with opposite signs. When the virtual array of the parametric source intercepts with the surface, the phase of the secondary field generated before the interception is reversed by the reflection from the surface. Since the phases of both primary signals are reversed simultaneously upon reflection from the surface, the secondary field so generated after the reflection has the same phase as that prior to the reflection. Thus it can be appreciated that the contribution to the secondary field from the real virtual array is from two parts. The first part from the source is a cylinder chopped at the water-air interface. The second part starts from the chopped end on the surface and continues unimpeded in the water in the specula direction. It is the same for the image of the virtual array.

Provided the observation point  $(x, y, z)$  is far away from the source, the total secondary field in the far field region can be described by

$$P(\mathbf{R}) = -\frac{\omega_d^2 P_1 P_2 \beta}{4\pi\rho_0 c_0^4 R} e^{-(\alpha_d + k_d)R} I, \quad (4)$$

where  $R = \sqrt{x^2 + y^2 + z^2}$  and,

$$I = I_1|_0^{x'_2} - I_1|_{x'_2}^\infty - 2I_1|_{x'_1}^{x'_2} - I_2|_0^{x''_2} + I_2|_{x''_2}^\infty + 2I_2|_{x''_1}^{x''_2} \quad (5)$$

and

$$I_1|_0^{x'_2} - I_1|_{x'_2}^\infty = \frac{S_0}{2} \frac{2e^{-2(\alpha_1 + jk_d \sin^2 \theta/2)x'_2} - 1}{\alpha_1 + jk_d \sin^2 \theta/2} D(\theta), \quad (6)$$

$$I_2|_0^{x''_2} - I_2|_{x''_2}^\infty = \frac{S_0}{2} \frac{2e^{-2(\alpha_1 + jk_d \sin^2 \gamma/2)x''_2} - 1}{\alpha_1 + jk_d \sin^2 \gamma/2} D(\gamma), \quad (7)$$

$$x'_1 = x''_1 = \frac{z_0 - a \cos \varphi}{\sin \varphi}, \quad (8)$$

$$x'_2 = x''_2 = \frac{z_0 + a \cos \varphi}{\sin \varphi}, \quad (9)$$

$$\cos \theta = \frac{x \cos \varphi - (z - z_0) \sin \varphi}{\sqrt{x^2 + y^2 + (z - z_0)^2}}, \quad (10)$$

$$\cos \gamma = \frac{x \cos \varphi + (z + z_0) \sin \varphi}{\sqrt{x^2 + y^2 + (z + z_0)^2}}, \quad (11)$$

$$D(\theta) = \frac{2J_1(k_d a \sin \theta)}{k_d a \sin \theta}. \quad (12)$$

$D(\theta)$  is the directional pattern for a piston-type transducer.

The contribution to the secondary field from the chopped part of the cylinder between the hatched area and  $A'$  for the real source is given by the following integral expression:

$$I_{1s} = c_1 D(\theta) \int_{-a}^a \left( 1 - \exp \left\{ - \left[ 2\alpha_1 + j \frac{k_d}{R} (R - x') \right] \times \left( \frac{x_1 - x_2}{2} + \frac{z_1}{\tan \varphi} \right) \right\} \right) \exp \left( j \frac{k_d}{R} z' z_1 \right) \times \sin \frac{k_d y}{R} \sqrt{a^2 - z_1^2} dz_1, \quad (13)$$

and the counterpart for the image source is expressed by

$$I_{2s} = c_2 D(\gamma) \int_{-a}^a \left( 1 - \exp \left\{ - \left[ 2\alpha_1 + j \frac{k_d}{R} (R - x'') \right] \times \left( \frac{x_1 - x_2}{2} - \frac{z_2}{\tan \varphi} \right) \right\} \right) \exp \left( j \frac{k_d}{R} z'' z_2 \right) \times \sin \frac{k_d y}{R} \sqrt{a^2 - z_2^2} dz_2, \quad (14)$$

where

$$c_{1,2} = -\frac{2R}{k_d y} \frac{\exp \left\{ - \left[ 2\alpha_1 + j \frac{k_d}{R} (R - x'') \right] x'_2 \right\}}{2\alpha_1 + j \frac{k_d}{R} (R - x'')} \quad (15)$$

and

$$x' = x \cos \varphi - (z - z_0) \sin \varphi, \quad (16)$$

$$z' = x \sin \varphi + (z - z_0) \cos \varphi,$$

$$x'' = x \cos \varphi + (z + z_0) \sin \varphi,$$

$$z'' = x \sin \varphi + (z + z_0) \cos \varphi. \quad (17)$$

When the source is positioned such that the distance to the surface is much greater than the length of the virtual array, then there is little cancellation of the virtual array before and after reflection from the surface so that Eq. (5) can be reduced to the summation of two parametric arrays (real and image) with opposite signs.

## B. Numerical method (Spherical spreading model)

In practice, most of the parametric arrays are used in the region  $R_F \leq 1$ . The secondary signal in such cases is generated not only within the collimated zone, but also substantially outside it in a spherically spreading beam. Under these circumstances, the source density function in Eq. (1) should be

$$q_d = j \frac{\omega_d P_1 P_2 \beta}{\rho_0^2 c_0^4} e^{-(\alpha_1 + \alpha_2)r} \frac{e^{j(\omega_d t - k_d r)}}{r^2} D_1 D_2, \quad (18)$$

where  $D_1$  and  $D_2$  are the beam patterns of the parametric source at the two primary signals. Equation (1) becomes

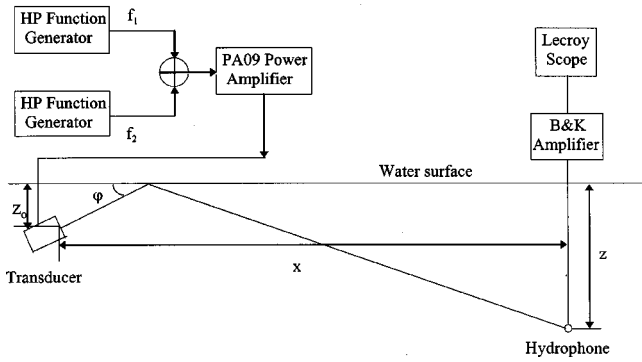


FIG. 2. The experimental setup for measuring the sound fields from a parametric source.

$$P_d(\mathbf{R}) = -\frac{\omega_d^2 P_1 P_2 \beta}{4\pi\rho_0 c_0^4} \int \int \int \frac{D_1 D_2}{r_v} e^{-[\alpha_T r' + jk_d(r_v + r')]} \times \sin \theta dr' d\theta d\phi', \quad (19)$$

where  $r_v = (r^2 + r'^2 - 2rr' \cos \nu)^{1/2}$  and  $\cos \nu = \cos \theta' \cos \theta + \sin \theta' \sin \theta \cos(\phi' - \phi)$ .

For the parametric array, which interacts with the water surface, the field of the secondary signal is given by

$$P(\mathbf{R}) = P'_1 - P'_2 - P''_1 + P''_2, \quad (20)$$

where  $P'_1$  and  $P''_1$  are the contributions from the parametric source and its image before reflection, and  $P'_2$  and  $P''_2$  are the contributions from the parametric source and its image after reflection. They can only be evaluated by numerical methods. The Gaussian-Legendre quadrature is used to evaluate the integration in Eq. (20). Notice that Eq. (20) can be applied to a transmit source of arbitrary shape.

## II. EXPERIMENTS

Experiments were carried out in a laboratory tank to examine the effects of the water surface on the secondary wave field. The size of the tank is about 8 m long, 5 m wide, and 3 m deep. The setup of the experiment is shown in Fig. 2. Two HP 8116A function generators were used to generate two primary signals at frequency  $f_1$  and  $f_2$ , respectively. Pulse signals were used, and the length of the pulse was chosen so that there was no interference from the reflection on the back wall of the tank. The reflections from the side walls and bottom were negligible because of the narrow beam width of the transducer. The two primary signals were added together and fed into a PA09 power amplifier to drive a disk transmit transducer. The transmit transducer was mounted on a tilt mechanism driven by a stepper motor, which was controlled by a PC via its parallel printer port. This allowed the incident angle  $\phi$  from the transducer to the surface to vary accordingly. The receive hydrophone was a B&K 8103 hydrophone. In order to increase the signal-to-noise ratio, a B&K wide-band conditioning amplifier 2638 was used with a pass band from 200 kHz to 2 MHz and a gain of 60 dB. The received signals after amplification were captured by a Lecroy 9310L digital oscilloscope and the power spectra of the signals was recorded so as to find the signal amplitudes at the two primary and secondary frequen-

TABLE I. Parameters used in the experiments.

	$f_1 = 2.44$ MHz	$f_2 = 2.71$ MHz	$f_d = 270$ kHz
Absorption coefficient (Neper/m)	0.2082	0.2563	$7.705 \times 10^{-3}$
Virtual array length (m)	2.4	1.95	N/A
Beam width (degrees)	1.23	1.11	3.28

cies. All the measurements were relative. A power spectrum average was used to improve the signal-to-noise ratio in the experiments.

The linearity of the equipment, especially that of the power amplifier, is critical to the experiment, therefore it was carefully measured to determine a proper transmit signal level, at which the maximum signal-to-noise ratio was adequate for the measurements, while the nonlinear effects from the equipment was kept just below the noise background.

In order to clearly observe the truncation effects in the tank, the frequencies of the signals has to be high enough to generate a virtual array of the secondary wave well within the tank. The transducer was made of a thin disk of PZT-4 ceramic, 0.75 mm thick with a diameter of 3 cm. The resonant frequency of the transducer was about 2.6 MHz. The primary and secondary frequencies and some other important parameters used in the experiments are listed in Table I below. The beam width at the secondary frequency is predicted by Westervelt model.<sup>1</sup>

In the experiments, both the primary and secondary signals were measured and recorded simultaneously. Since the level of the secondary signal is proportional to the primary signals, monitoring the primary signals provided means to crosscheck if the measurements were made correctly.

A B&K 8103 hydrophone was used as a receive hydrophone in the experiments. Since the primary and secondary frequencies were well above the calibration frequency range, the directionality of the hydrophone was measured at these frequencies in its horizontal plane. The results are plotted in Fig. 3 in a polar coordinate system. The responses are marked with up and down triangles at the primary frequencies  $f_1$  and  $f_2$ , respectively, and squares at the secondary frequency  $f_d$ . As expected, the response of the hydrophone was almost omnidirectional at the secondary frequency, 270 kHz. There was only a 1.7 dB change from the minimum to maximum in the angular response. However, there were significant changes in the angular response at the two primary frequencies (2.44 and 2.71 MHz). The orientation of the hydrophone in the experiment setup was chosen so that the angle of interests was within  $\pm 20^\circ$ , around 0 degree. It can be seen that the angular response of the hydrophone was quite uniform within the angle of interest at the primary and secondary frequencies.

The measured beam patterns at  $f_1$  and  $f_2$ , respectively, for the transmit transducer are plotted in Fig. 4. The agree-

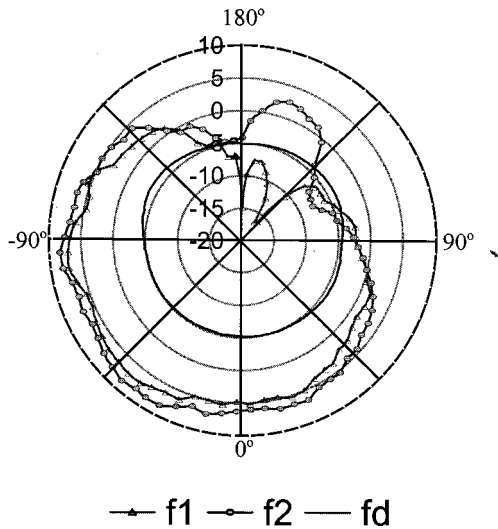


FIG. 3. Measured polar response of the B&K 8103 hydrophone at the primary and secondary frequencies.

ment between the measured and theoretical beam patterns is very good for the primary signals.

Figure 5 shows the measured beam pattern and predicted results by the Westervelt model<sup>1</sup> and the Mellon and Moffett numerical model<sup>20</sup> at the secondary frequency. The aperture correction to the parametric source<sup>21</sup> was applied to both predicted results. There are no sidelobes in the beam pattern at the secondary frequency. The predicted beam width by Westervelt model is  $3.28^\circ$ , which is much greater than the measured one, as illustrated in the figure. There are two reasons for the discrepancy. First, the measurements were not taken far enough away from the source to be strictly in the far field of the virtual array. Although the measurement was made at a distance equal to about three times the length of the virtual array, this is still not long enough to satisfy the approximation made in deriving the Westervelt model. The other reason is that the Westervelt model predicts the secondary field with a virtual array well within the Rayleigh distance of the primary field, i.e., a collimated array; whereas in our experiment, the virtual array is about the same length as the Rayleigh distance. Thus, there is some difference between the model and the conditions of the experiment. Since it is impossible, with the tank available, to measure the beam pattern at a range long enough to be in far field (about ten times the length of the virtual array distance is required), the numerical method proposed by Mellen and Moffett<sup>20</sup> was

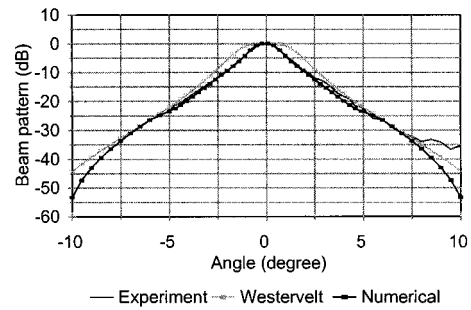


FIG. 5. Measured and predicted secondary beam patterns at 270 kHz.

used to calculate the beam pattern under the experimental conditions. It can be seen that a much better agreement between the measured beam pattern and that predicted by the numerical model is obtained. However, the rapid fall of the secondary signal level away from the acoustic axis resulted in a very low signal-to-noise ratio at angles greater than  $7.5^\circ$ , thus the measured beam pattern at large angles was distorted due to the background noise as shown in the figure.

In order to directly observe the effects caused by the surface reflection, the beam patterns of the transducer in the vertical plane were measured with the hydrophone at a distance 5 m away from the transducer and a depth of 0.75 m below the surface. The initial angle of the transducer was  $-14^\circ$ . The minus sign indicates the transducer was pointing toward the surface. The primary and secondary fields were measured as the transducer tilted downward step-by-step to  $17^\circ$ . There should be two main peaks at each frequency corresponding to the direct and reflected paths as the transmit transducer is rotated in the vertical plane. The acoustical axis is  $7.97^\circ$  for the direct path, and  $-9.09^\circ$  for the surface reflected path with this geometry. The path length before reflection is about 0.316 m, which is much shorter than the length of the virtual array (2.153 m). The pulse duration is about 400 cycles at  $f_1$  and 450 cycles at  $f_2$  with a repetition period of 20 ms.

The measured results, as displayed in Fig. 6, indicate that the peaks of the beam patterns are at  $7.9^\circ$  and  $-8.9^\circ$ , respectively. The primary signals were used in the measurement of the orientation of the transmit transducer, in this case because of their narrow beam widths. These values are in very good agreement with the predictions made from the geometry of the problem. It is obvious that the amplitude of the surface reflected secondary signal was about 2 dB less than that of the direct path. It is shown that the total mea-

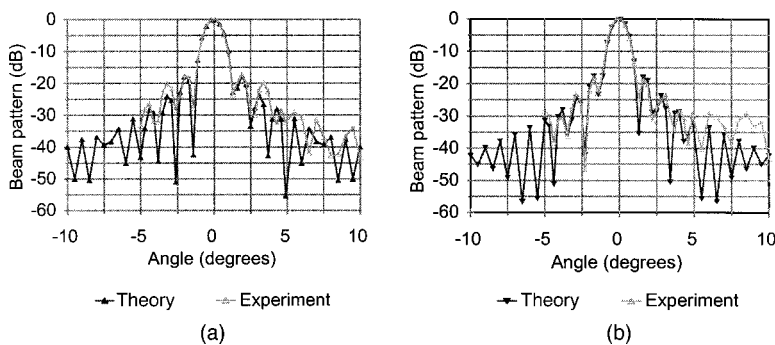


FIG. 4. Measured and predicted primary beam patterns at 2.44 and 2.71 MHz.

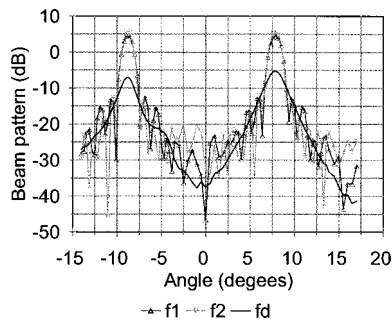


FIG. 6. Measured primary and secondary beam patterns as functions of angle at a fixed hydrophone depth 0.75 m.

sured primary signal level for the surface reflected path was just 0.3 dB less than that of the direct path. Apart from the slightly increased spreading loss associated with the surface reflection path, the small differences in the amplitudes of the two primary signals from the surface may also be due to the angular response of the hydrophone and small misalignment between the transducer and hydrophone. However, the difference between the direct and surface reflected secondary signals is greater than the spreading loss and the possible errors that the angular response and misalignment could introduce. This demonstrates that the interaction of the secondary signal with the surface resulted in a reduction of the signal level. In addition to this decrease in amplitude, there is also a ‘hump’ in the surface reflection path in comparison with the steady reduction in the direct-path signal with changing angle.

### III. COMPARISONS OF RESULTS

To measure the sound fields as a function of angle with the parametric source interacting with the surface, the transducer was fixed at grazing angles of  $5.4^\circ$  and  $7.7^\circ$ , respectively. The hydrophone was lowered down step-by-step from the surface to a depth of 1 m, at a horizontal distance of 5 m away from the transducer. The measured primary and secondary fields are plotted in Figs. 7 and 8 for the two transmitting angles. The values were normalized with respect to the maximum value of the direct path. In order to make comparisons, the secondary field radiated directly from the image source position in a free field condition is also shown in the figures. The measured results demonstrate that there are two main features in the secondary field after reflection. The first one is the reduction of the signal level on the axis

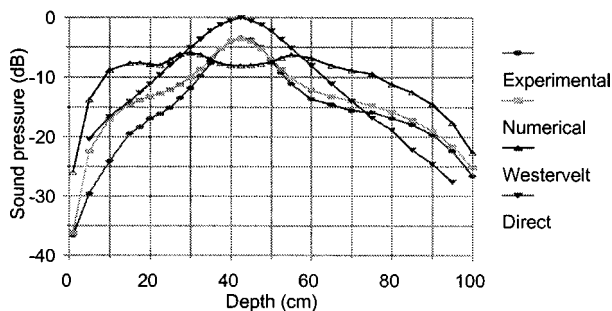


FIG. 7. Measured primary and secondary field as functions of depth with the transducer at  $5.4^\circ$ .

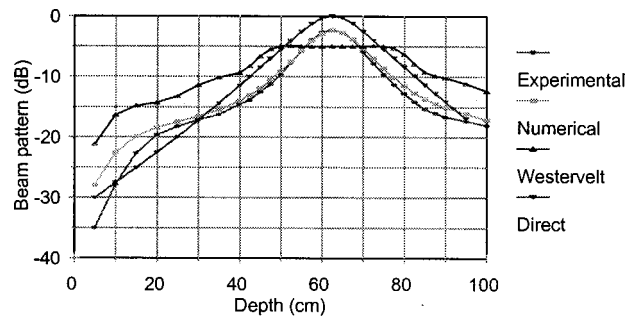


FIG. 8. Measured primary and secondary field as functions of depth with the transducer at  $7.7^\circ$ .

and its vicinity. The second one is the relative increasing skirt of the secondary field in comparison with that in a free field. This is due to the destructive summation of the virtual array of the parametric source introduced by the surface reflection which reduces the overall effective length of the array in its end-fire direction, resulting in a slow roll-off in the beam pattern. It is also seen that the beam pattern of the secondary field is not symmetrical any more. There is a very rapid decrease in the sound level close to the surface away from the beam center in contrast to a more gentle roll-off at the opposite side. This is because of the Lloyds mirror effect of the surface, where the direct path from the real source interacts with the reflected path from the virtual source. This interference is more significant at the shallow end due to the narrow beam width of the parametric source.

The theoretical predictions for the corresponding measurements are also shown in Figs. 7 and 8. Compared with the experimental results, it can be seen that the Westervelt model predicts a larger signal loss than it is observed at and near the center of the beam. Also, the predicted peaks of the secondary field and the skirts of the beam are considerably higher than the experimental results show. In contrast, the predictions given by the numerical model show much closer agreement between theory and experiment. In fact, there is exact agreement in the signal levels at the center of the beam, and the beam patterns on the deeper-water side of the beam center follow each other closely. However, there are some discrepancies between the beam patterns on the shallow-water side of the beam center. In summary these comparisons demonstrate that the Westervelt model can only qualitatively predict the reduction of the secondary field by the water surface for the particular experiments reported here,<sup>22</sup> whereas the numerical model is able accurately to quantify the reduction.

It is clear that the reflection effects depends on the grazing angle of the waves incident upon the surface. Figure 9 shows the measured sound pressure on the acoustic axis of the virtual source as a function of the grazing angle. The source transducer was fixed at a depth of 5 cm and it was tilted to change the grazing angle. The depth of the hydrophone was varied so as to remain on the acoustic axis. Thus for large grazing angles the hydrophone was positioned deeper than for small angles. The result is normalized with respect to the measured secondary signal via the direct path in a free field condition. The sound field produced by a conventional source, with the same aperture as the parametric



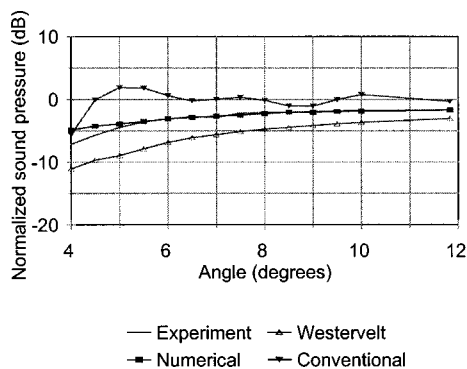


FIG. 9. Normalized sound pressure at the secondary frequency with parametric and conventional sources.

source, which is radiating directly at the frequency of 270 kHz, is also plotted on the same figure in a line with solid triangle marks. It can be seen that there are some ripples in the sound level due to the Lloyds mirror effect of the surface for the conventional source, whereas the secondary signal level increases monotonically with the grazing angle in this region. It approached a constant level at higher grazing angles, as shown in the figure. The constant level is determined by the depth of the transducer with the acoustic axis of the transducer perpendicular to the water surface. There are two extreme cases in which no cancellation loss is introduced by the surface; one is the transducer just touches the surface (only possible in theory) and the other one is the transducer far away (at least 10 times greater than that of the length of the virtual array) from the surface. The only difference between the two cases is the phases of the secondary wave are opposite.

The predicted result with Westervelt model is shown in Fig. 9. It can be seen that the result follows the same trend as the experiment. However, the predicted loss is higher than that of the measured. The secondary field calculated by the numerical method is also shown in Fig. 9. The experimental and numerical results agree well with each other seen at the grazing angle greater than 5 degrees.

#### IV. SUMMARY AND CONCLUSIONS

Theoretical and experimental studies have been carried out to examine the effects of a flat water surface on the signal level at the secondary frequency of a nonsaturation limited parametric source. The results presented here show that the interaction of the virtual array of a parametric source with a water surface introduces reduction in the secondary signal level on the axis and its vicinity and distortion in the beam pattern of the parametric source. The effects depend on the conditions relevant to the generation of the secondary wave. The predicted results with the Westervelt model demonstrate the same trend as that of experimental results, but they overestimate the signal loss for the cases concerned in this study, whereas the numerical method provides good prediction on the secondary field for surface interaction with high grazing angles.

#### ACKNOWLEDGMENTS

The authors would like to thank Ming Zheng for his assistance in some of the experiments. This work is supported by the European Union Marine Science and Technology Directorate under Contract No. MAS2-CT91005: project "PARACOM."

- <sup>1</sup>P. J. Westervelt, "Parametric acoustic array," *J. Acoust. Soc. Am.* **35**, 535–537 (1963).
- <sup>2</sup>T. G. Muir, "A survey of several nonlinear acoustics experiments on traveling wave fields," in *Proceedings of the 1973 Symposium on Finite Amplitude Wave Effects in Fluids*, Copenhagen, edited by L. Bjono (I.P.C. Science and Technology, Guilford, 1974).
- <sup>3</sup>J. N. Tjøtta and S. Tjøtta, "Reflection and refraction of parametrically generated sound at a water sediment interface," in *Proceedings of the NATO Conference on Bottom Interacting Ocean Acoustics, Saclant ASW Research Centre, Italy, 9–13 June 1980*, edited by W. A. Kuperman and F. B. Jensen (Plenum, New York, 1980).
- <sup>4</sup>N. G. Pace and R. V. Ceen, "Time domain study of the terminated transient parametric array," *J. Acoust. Soc. Am.* **73**, 1972–1978 (1983).
- <sup>5</sup>D. J. Wingham, "A theoretical study of the penetration of a water-sediment interface by parametric array," *J. Acoust. Soc. Am.* **76**, 1192–1200 (1984).
- <sup>6</sup>D. J. Wingham, N. G. Pace, and R. V. Ceen, "An experimental study of the penetration of a water-sediment interface by a parametric beam," *J. Acoust. Soc. Am.* **79**, 363–374 (1986).
- <sup>7</sup>W. L. Konrad, "Applications of the parametric source," *Proceedings of the Institute of Acoustic Conference, Bath University*, paper 6.1, September 1979.
- <sup>8</sup>A. H. Quazi and W. L. Konrad, "Underwater acoustic communications," *IEEE Commun. Mag.* **20**, 24–30, (1982).
- <sup>9</sup>T. Hyodo, "Underwater parametric acoustic source," *Defence Technology Journal* **3**, 1–18 (1983).
- <sup>10</sup>R. Coates and L. Kopp, "The use of parametric transduction for underwater acoustic communication: Project PARACOM," in *Proceedings of the 1st European Conference on Underwater Acoustics* (Elsevier Applied Science, New York, 1992), pp. 96–99.
- <sup>11</sup>L. L. Mellenbruch and T. G. Muir, "Experiments on phase-reversed shock propagation and parametric generation," *J. Acoust. Soc. Am.* **55**, 429 (1974).
- <sup>12</sup>J. C. Lock, "Theoretical performance of a parametric array with phase-reversed primary radiation," *J. Acoust. Soc. Am.* **55**, 429 (1974).
- <sup>13</sup>R. C. Khosla and Anthony I Eller, "Surface reflection of an underwater acoustic parametric source," *J. Acoust. Soc. Am.* **55**, S23 (1974).
- <sup>14</sup>T. G. Muir, L. L. Mellenbruch, and J. C. Lockwood, "Reflection of finite-amplitude waves in a parametric array," *J. Acoust. Soc. Am.* **62**, 271–276 (1977).
- <sup>15</sup>B. V. Smith, "Conditions for Distortionless underwater communication using parametric array," *Acoustica* **81**, 161–166 (1995).
- <sup>16</sup>B. V. Smith, "Introduction to nonlinear acoustics and the parametric array," *Proc. Of Institute of Acoustics* **14**, 1–25 (1992).
- <sup>17</sup>H. O. Berktaay and D. J. Leahy, "Far-field performance of parametric transmitters," *J. Acoust. Soc. Am.* **55**, 539–546 (1974).
- <sup>18</sup>M. B. Moffett and R. H. Mellen, "Model for parametric acoustic sources," *J. Acoust. Soc. Am.* **61**, 325–337 (1977).
- <sup>19</sup>M. B. Moffett, R. H. Mellen, and W. L. Konrad, "Parametric acoustic sources of rectangular aperture," *J. Acoust. Soc. Am.* **63**, 1326–1331 (1978).
- <sup>20</sup>R. H. Mellen and M. B. Moffett, "A numerical method for calculating the near-field of a parametric acoustic source," *J. Acoust. Soc. Am.* **63**, 1622–1624 (1978).
- <sup>21</sup>M. B. Moffett and R. H. Mellen, "On parametric source aperture factors," *J. Acoust. Soc. Am.* **60**, 581–583 (1976).
- <sup>22</sup>L. Wang, B. Smith, and R. Coates, "Effects of Water Surface on the Field of a Parametric Source," in *Proceedings of the IEEE Oceans'96 Conference* (Florida, 1996), pp. 383–389.

# Sound scattering cross-section in a stratified moving atmosphere

V. E. Ostashev, G. H. Goedecke, and R. Wood

*Department of Physics, New Mexico State University, Las Cruces, New Mexico 88003-8001*

H. Auvermann

*U.S. Army Research Laboratory, 2800 Powder Mill Road, Adelphi, Maryland 20783-1155*

S. F. Clifford

*NOAA/ERL/ETL, 325 Broadway, Boulder, Colorado 80303*

(Received 30 April 1998; revised 4 March 1999; accepted 8 March 1999)

For a realistic model of a stratified moving atmosphere with arbitrary vertical profiles of the adiabatic sound speed and wind velocity vector, an equation is derived for the sound scattering cross-section per unit volume,  $\sigma$ , as a function of apparent scattering angle  $\Theta_0$ . The effects of these profiles on  $\sigma$  are studied numerically. It is shown that if the wind velocity is zero, but the adiabatic sound speed varies with height, then  $\sigma$  can be affected significantly for  $\Theta_0 < 110^\circ$ . Furthermore, if the wind velocity varies with height, but the adiabatic sound speed does not, then  $\sigma$  can be affected significantly for  $\Theta_0 < 110^\circ$  as well as  $\Theta_0$  near  $180^\circ$ . These and other numerical calculations have shown that in many cases acoustic remote sensing of the structure parameters of temperature and wind velocity fluctuations in the atmosphere should be based on the derived equation for  $\sigma$  rather than on that used in the literature. The derived equation for  $\sigma$  is also compared to those obtained by Clifford and Brown [J. Acoust. Soc. Am. **55**, 929–933 (1974)] and by Ye [J. Acoust. Soc. Am. **102**, 754–758 (1997)] for a model of a stratified moving atmosphere as two homogeneous layers in relative motion. It is explained why the predictions by Clifford and Brown are different from those by Ye. © 1999 Acoustical Society of America. [S0001-4966(99)03206-3]

PACS numbers: 43.28.Fp, 43.20.Fn [LCS]

## INTRODUCTION

The sound scattering cross-section per unit volume,  $\sigma(\Theta)$ , is one of the most important statistical characteristics of a sound wave propagating in a turbulent atmosphere. Here,  $\Theta$  is the “true” scattering angle, i.e., the angle between the propagation directions of the incident and scattered waves at the scattering volume, as defined by Clifford and Brown.<sup>1</sup> An equation for  $\sigma$  describes some of the main features of the interaction of sound with turbulence. Furthermore, the mean intensity of the scattered wave is proportional to  $\sigma$ , which is important for many practical concerns, e.g., sound scattering into a refractive shadow zone. Finally, an equation for  $\sigma$  is a theoretical basis for acoustic remote sensing of the atmosphere by sodars. Nowadays sodars are produced commercially and are widely used for acoustic sounding of the atmosphere, air pollution control, weather monitoring near airports, etc.<sup>2</sup>

The classical equation for the sound scattering cross-section was derived in 1961 by Monin<sup>3</sup> for the case of Kolmogorov spectra of temperature and wind velocity fluctuations:

$$\sigma(\Theta) = \frac{5}{2^{2/3} 288 \Gamma(1/3)} \frac{k^{1/3} \cos^2 \Theta}{(\sin(\Theta/2))^{11/3}} \times \left[ \frac{C_T^2}{T_0^2} + \frac{22}{3} \frac{\cos^2(\Theta/2) C_v^2}{c_0^2} \right]. \quad (1)$$

Here,  $C_T^2$  and  $C_v^2$  are the structure parameters of temperature and wind velocity fluctuations,  $T_0$  and  $c_0$  are mean values of

the temperature  $T$  and adiabatic sound speed  $c$ ,  $\Gamma(1/3)$  is the gamma-function, and  $k$  is the sound wave number. Subsequently Eq. (1) has been used in hundreds of papers dealing with the aforementioned problems. (Some references on these papers can be found in Refs. 2 and 4.) Use of this equation and sodar measurements of  $\sigma(\Theta)$  allow retrieval of  $C_T^2$  and  $C_v^2$ , which are important parameters of the atmospheric boundary layer. Indeed, according to Eq. (1), the backscattering cross-section  $\sigma(\Theta=180^\circ)$  is directly proportional to  $C_T^2$ :

$$\sigma(180^\circ) = 4.08 \times 10^{-3} k^{1/3} \frac{C_T^2}{T_0^2}. \quad (2)$$

Measurements of  $\sigma(180^\circ)$  by a monostatic sodar allow retrieval of  $C_T^2$ . Furthermore, in the convective boundary layer of the atmosphere  $C_v^2/c_0^2$  is usually greater than  $C_T^2/T_0^2$ . For example, at several hundred meters above the ground,  $C_v^2/c_0^2$  can be greater than  $C_T^2/T_0^2$  by two orders of magnitude.<sup>5</sup> Therefore, if the angle  $\Theta$  is not very close to  $180^\circ$ , the scattering cross-section is approximately proportional to  $C_v^2$ :

$$\sigma(\Theta \neq 180^\circ) \approx 0.03 \frac{k^{1/3} \cos^2 \Theta \cos^2(\Theta/2)}{(\sin(\Theta/2))^{11/3}} \frac{C_v^2}{c_0^2}. \quad (3)$$

Bistatic sounding of the atmosphere by sodars allows measurement of  $\sigma(\Theta \neq 180^\circ)$  and, hence, retrieval of  $C_v^2$ .

Thus, Eq. (1) is indeed one of the most important theoretical results of sound propagation in a turbulent atmosphere. However, this equation was derived for the case when the mean value of the adiabatic sound speed  $c$  is a

constant and the mean wind velocity vector  $\mathbf{v}=0$ . This never happens in the real atmosphere, where  $c(z)$  and  $\mathbf{v}(z)$  always depend on the height  $z$  above the ground. Therefore, there has always been concern about whether it is sufficiently accurate to use Eqs. (1)–(3) for describing sound scattering in the atmosphere, since stratification will cause refraction and thus the “apparent” scattering angle  $\Theta_0$  (as defined in Ref. 1) between emitted and received waves will differ from the true scattering angle  $\Theta$ .

This issue was addressed by Clifford and Brown<sup>1</sup> in 1974. They gave qualitative arguments that the difficult problem of calculating  $\sigma(\Theta_0)$  in a stratified moving atmosphere with arbitrary profiles of  $c(z)$  and  $\mathbf{v}(z)$  can be approximately (to the order  $v/c$ ) replaced by the simpler problem of calculating  $\sigma(\Theta_0)$  for a model of an atmosphere as two homogeneous layers in relative motion. In the lower layer, the mean wind velocity is zero, and in the upper layer, where a scattering volume is located,  $\mathbf{v}=\text{const}$ ; furthermore,  $c=c_0$  in both layers. Clifford and Brown derived an equation for  $\sigma$  for this model atmosphere and showed that Eq. (2) still holds for the case  $\mathbf{v}\neq 0$ . The derivation of this equation in Ref. 1 essentially uses the refraction law for the normal to a wavefront of a sound wave at the horizontal interface between the upper and lower layers; this interface is tacitly assumed to be at rest with respect to the ground. The results obtained by Clifford and Brown have been used in many subsequent papers.

However, the recent paper by Ye<sup>6</sup> revises the refraction law for the normal to a wavefront derived by Clifford and Brown<sup>1</sup> and, hence, their equation for  $\sigma$ , for the case when the vertical component of wind velocity,  $w$ , is not zero. Ye<sup>6</sup> also obtained the interesting result that Eq. (2) does not hold anymore if  $\mathbf{v}\neq 0$ . Furthermore, he pointed out that Eq. (25) of Ref. 1 for the cosine of the “true” scattering angle does not satisfy nonreciprocity for backscattering.

In the present paper, for the first time we derive an equation for the scattering cross-section  $\sigma$  in the turbulent atmosphere with arbitrary profiles of  $c(z)$  and  $\mathbf{v}(z)$ . The derived equation for  $\sigma$  depends on the values of  $c(z)$  and  $\mathbf{v}(z)$  only at the source-receiver height and the height of the scattering volume. Using this equation, we numerically study the effects of  $c$  and  $\mathbf{v}$  on  $\sigma$ . Furthermore, we discuss whether Eqs. (2) and (3) can be used as a theoretical basis for acoustic remote sensing of  $C_T^2$  and  $C_v^2$ .

Since we were able to rigorously derive the equation for  $\sigma$  for arbitrary profiles of  $c(z)$  and  $\mathbf{v}(z)$ , the model of an atmosphere as two homogeneous layers in relative motion, which is used in Refs. 1 and 6 to simplify this derivation and obtain an approximate equation for  $\sigma$ , is not needed anymore. However, our analysis would not be complete if we did not compare the derived equation for  $\sigma$  with those obtained in Refs. 1 and 6 and explain why some of the results obtained in these two papers are different. To complete this task, we derive an equation for  $\sigma$  for the model of an atmosphere as two homogeneous layers in relative motion by an approach different from those used in both Refs. 1 and 6. This equation is compared with the equation for  $\sigma$  for arbitrary profiles of  $c(z)$  and  $\mathbf{v}(z)$  and with those obtained in Refs. 1 and 6. We show that the refraction law for the normal

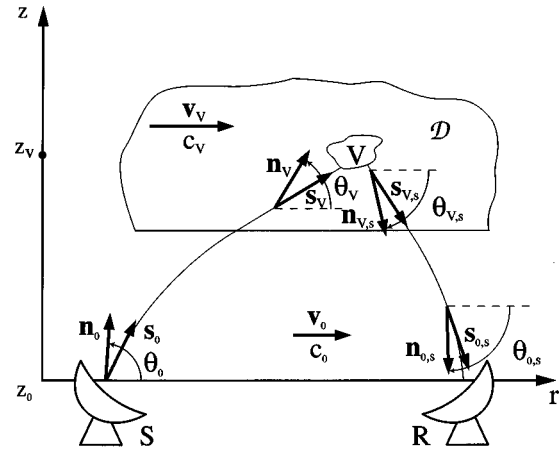


FIG. 1. Geometry of the problem for a stratified moving atmosphere. S is source; R is receiver;  $\mathbf{s}_0$ ,  $\mathbf{s}_v$ ,  $\mathbf{s}_{v,s}$  and  $\mathbf{s}_{0,s}$  are the unit vectors tangential to the trajectory of sound propagation; and  $\mathbf{n}_0$ ,  $\mathbf{n}_v$ ,  $\mathbf{n}_{v,s}$  and  $\mathbf{n}_{0,s}$  are the unit vectors normal to the wavefronts.

to a wavefront derived by Clifford and Brown<sup>1</sup> is correct if the interface is at rest, while that derived in Ref. 6 is correct only if the interface is moving with vertical velocity  $w$ . We also show that the difference in predicting sound backscattering in Refs. 1 and 6 is due to the difference in accuracy to which  $\sigma$  is calculated in these papers; and that the formula for the cosine of the true scattering angle [Eq. (25) of Ref. 1] should not satisfy nonreciprocity for backscattering, to the order of  $v/c_0$  to which calculations are done in Ref. 1.

In Sec. I, the equation for  $\sigma$  is derived for a realistic model of a stratified moving atmosphere with arbitrary profiles of  $c(z)$  and  $\mathbf{v}(z)$ . The effects of the adiabatic sound speed and wind velocity stratification on  $\sigma$  are studied numerically in Sec. II. In Sec. III, the equation for  $\sigma$  is derived for the model of an atmosphere as two homogeneous layers in relative motion. It also deals with comparison of this equation and its consequences with those presented in Refs. 1 and 6. The results obtained are summarized in Sec. IV.

## I. STRATIFIED MOVING ATMOSPHERE

### A. Geometry of the problem

The geometry of the problem is shown in Fig. 1. This geometry is considered in the ground-fixed coordinate system  $K$ . Source  $S$  and receiver  $R$  are located at the height  $z=z_0$  above the ground ( $z_0$  might be zero). The source emits a sound wave which is scattered by atmospheric turbulence located in the scattering volume  $V$  at the height  $z_v$ . A scattered wave is recorded by the receiver. The adiabatic sound speed  $c(z)$  and wind velocity vector  $\mathbf{v}(z)=(\mathbf{u}(z), w(z))$  are functions of the height  $z$ , where  $\mathbf{u}$  and  $w$  are the horizontal and vertical components of  $\mathbf{v}$ . Nontrivial  $w$  can be caused by surface topography, upward movement of warm air in thermal plumes, downward movement of cold air, and some other mechanisms. The absolute value of  $w$  is usually less than 2m/s. We also denote  $c_0=c(z_0)$  and  $\mathbf{u}_0=\mathbf{u}(z_0)$ , and assume for simplicity that  $w(z_0)=0$ . (The source and receiver are usually located close to the ground where  $w$  vanishes.)

Note that a sound wave propagates in the direction of the unit vector  $\mathbf{s}$  tangential to the group velocity of this wave. In a moving medium,  $\mathbf{s}$  is given by the formula  $\mathbf{s}=(\mathbf{n}+\mathbf{v}/c)/|\mathbf{n}+\mathbf{v}/c|$  and is different from the unit vector  $\mathbf{n}$  normal to the wavefront of the sound wave (e.g., Refs. 4 and 7).

The directions of propagation of the emitted wave near the source and near the scattering volume  $V$  are denoted by the unit vectors  $\mathbf{s}_0$  and  $\mathbf{s}_V$ , respectively (see Fig. 1). In the same regions, the unit vectors normal to the wavefront of the emitted wave,  $\mathbf{n}_0$  and  $\mathbf{n}_V$ , do not coincide with  $\mathbf{s}_0$  and  $\mathbf{s}_V$ . Similarly,  $\mathbf{s}_{V,s}$  and  $\mathbf{s}_{0,s}$  are the unit vectors in the directions of propagation of the scattered wave near the scattering volume  $V$  and receiver, and  $\mathbf{n}_{V,s}$  and  $\mathbf{n}_{0,s}$  are the unit vectors for the normal to the wavefront in these regions. The scattering volume  $V$  is formed by an intersection of cones around  $\mathbf{s}_V$  and  $\mathbf{s}_{V,s}$ , corresponding to radiation patterns of source and receiver.

## B. Sound scattering cross-section

We shall assume that the vertical scale  $l$  of variations in  $c(z)$  and  $\mathbf{v}(z)$  is much greater than the vertical scale  $l_V$  of the scattering volume  $V$ . In this case, it is possible to introduce a small region  $D$  which encloses the scattering volume  $V$  and where the adiabatic sound speed and wind velocity can be considered constant (see Fig. 1). We denote the values of  $c$ ,  $T$  and  $\mathbf{v}$  in  $D$  by  $c_V$ ,  $T_V$  and  $\mathbf{v}_V=(\mathbf{u}_V, w_V)$ .

Let us consider a process of sound scattering in the coordinate system  $K_1$  moving with the velocity  $\mathbf{v}_V$ . In this coordinate system and in the region  $D$ , the wind velocity is zero, the adiabatic sound speed is constant and equal to  $c_V$ , and the vectors  $\mathbf{n}_V$  and  $\mathbf{s}_V$  coincide, as do the vectors  $\mathbf{n}_{V,s}$  and  $\mathbf{s}_{V,s}$ . These are conditions under which Eq. (1) was derived. Therefore, in the coordinate system  $K_1$  and in the region  $D$ , the sound scattering cross-section  $\sigma$  is given by Eq. (1), where  $\Theta$  is the angle between the vectors  $\mathbf{n}_V$  and  $\mathbf{n}_{V,s}$ , and  $k$  should be replaced by its value  $k_V=k(z_V)$  at the height of the scattering volume. The angle  $\Theta$  is the true scattering angle. Noting that  $\cos \Theta=\mathbf{n}_V \cdot \mathbf{n}_{V,s}$ , the formula for  $\sigma$  can be written as

$$\sigma=\sigma_T+\sigma_v=\frac{5}{2^{5/6}72\Gamma(1/3)}\frac{k_V^{1/3}(\mathbf{n}_V \cdot \mathbf{n}_{V,s})^2}{(1-\mathbf{n}_V \cdot \mathbf{n}_{V,s})^{11/6}} \times \left[ \frac{C_T^2}{T_V^2} + \frac{11C_v^2(1+\mathbf{n}_V \cdot \mathbf{n}_{V,s})}{3c_V^2} \right], \quad (4)$$

where  $\sigma_T$  and  $\sigma_v$  are the temperature and wind velocity contributions to  $\sigma$ , corresponding to the first and second terms in the square brackets, respectively. Formula (4) for  $\sigma$  is valid in the coordinate system  $K_1$  moving with the velocity  $\mathbf{v}_V$ . It has exactly the same form in the ground-fixed coordinate system  $K$  since the unit vectors  $\mathbf{n}_V$  and  $\mathbf{n}_{V,s}$  normal to the wavefront and the wavenumber  $k_V$  are invariant under the Galilean transformation (e.g., section 5.4.1 from Ref. 4). Thus, a formula for the sound scattering cross-section  $\sigma$  in the ground-fixed coordinate system  $K$  is given by Eq. (4).

The next step in deriving a final equation for  $\sigma$  is to express  $\mathbf{n}_V \cdot \mathbf{n}_{V,s}$  and  $k_V$  in terms of  $\mathbf{n}_0$ ,  $\mathbf{n}_{0,s}$  and  $k_0=k(z_0)$ , which can be measured experimentally. It is convenient

to express the unit vectors normal to the wavefronts of the emitted and scattered waves in the forms:  $\mathbf{n}_0=(\mathbf{e}_0 \cos \theta_0, \sin \theta_0)$ ,  $\mathbf{n}_V=(\mathbf{e}_V \cos \theta_V, \sin \theta_V)$ ,  $\mathbf{n}_{V,s}=(\mathbf{e}_{V,s} \times \cos \theta_{V,s}, -\sin \theta_{V,s})$  and  $\mathbf{n}_{0,s}=(\mathbf{e}_{0,s} \cos \theta_{0,s}, -\sin \theta_{0,s})$ . Here,  $\theta_0$ ,  $\theta_V$ ,  $\theta_{V,s}$  and  $\theta_{0,s}$  are the grazing angles of the vectors  $\mathbf{n}_0$ ,  $\mathbf{n}_V$ ,  $\mathbf{n}_{V,s}$  and  $\mathbf{n}_{0,s}$  (see Fig. 1), while  $\mathbf{e}_0$ ,  $\mathbf{e}_V$ ,  $\mathbf{e}_{V,s}$  and  $\mathbf{e}_{0,s}$  are the unit vectors in the azimuthal directions of these vectors. All grazing angles are in the range from 0 to  $\pi/2$ .

## C. Refraction law for the normal to the wavefront

A sodar's wavelength  $\lambda$  is of the order of dozens of centimeters and is much less than the scale  $l$  of variations in the adiabatic sound speed and wind velocity. Therefore, geometrical acoustics can be used to obtain relationships between vectors  $\mathbf{n}_0$  and  $\mathbf{n}_V$ , and vectors  $\mathbf{n}_{V,s}$  and  $\mathbf{n}_{0,s}$ . The refraction law for the normal  $\mathbf{n}=(\cos \theta, \sin \theta)$  to the wavefront of a sound wave propagating in a stratified moving medium, derived by using the geometrical acoustics, is well known in the literature (e.g., Refs. 4 and 7–9):

$$\frac{c(z)}{\cos \theta(z)}+\mathbf{e} \cdot \mathbf{u}(z)+w(z) \tan \theta(z)=\text{const}, \quad \mathbf{e}=\text{const}. \quad (5)$$

Here,  $\theta$  is the grazing angle of the normal  $\mathbf{n}$ , and  $\mathbf{e}$  is the unit vector in the azimuthal direction of  $\mathbf{n}$ . According to Eq. (5), the azimuthal direction of the vector  $\mathbf{n}$  remains constant as a wave propagates in a medium while its grazing angle changes. Applying Eq. (5) for the emitted wave first at  $z=z_0$  and then at  $z=z_V$  and eliminating the constant between the equations obtained yields

$$\frac{c_V}{\cos \theta_V}+\mathbf{e}_V \cdot \mathbf{u}_V+w_V \tan \theta_V=\frac{c_0}{\cos \theta_0}+\mathbf{e}_0 \cdot \mathbf{u}_0, \quad \mathbf{e}_V=\mathbf{e}_0. \quad (6)$$

A relationship between the grazing angles  $\theta_{V,s}$  and  $\theta_{0,s}$  of the scattered wave can be found analogously:

$$\frac{c_V}{\cos \theta_{V,s}}+\mathbf{e}_{V,s} \cdot \mathbf{u}_V-w_V \tan \theta_{V,s}=\frac{c_0}{\cos \theta_{0,s}}+\mathbf{e}_{0,s} \cdot \mathbf{u}_0, \quad \mathbf{e}_{V,s}=\mathbf{e}_{0,s}. \quad (7)$$

Solving Eqs. (6) and (7) for  $\theta_V$  and  $\theta_{V,s}$  yields

$$\cos \theta_V=\frac{\cos \theta_0}{N} \frac{1-M+N \beta \sqrt{(1-M)^2-\cos^2 \theta_0(N^2-\beta^2)}}{(1-M)^2+\beta^2 \cos^2 \theta_0}, \quad (8)$$

$$\cos \theta_{V,s}=\frac{\cos \theta_{0,s}}{N} \frac{1-M_s-N \beta \sqrt{(1-M_s)^2-\cos^2 \theta_{0,s}(N^2-\beta^2)}}{(1-M_s)^2+\beta^2 \cos^2 \theta_{0,s}}. \quad (9)$$

Here,  $N=c_0/c_V$  is a value of the refractive index in a motionless atmosphere at the height  $z_V$ ,  $M=\cos \theta_0 \mathbf{e}_0 \cdot (\mathbf{u}_V-\mathbf{u}_0)/c_0$ ,  $M_s=\cos \theta_{0,s} \mathbf{e}_{0,s} \cdot (\mathbf{u}_V-\mathbf{u}_0)/c_0$ , and  $\beta=w_V/c_0$ . We assume that there are no critical angles for the geometry shown in Fig. 1 so the square roots in Eqs. (8) and (9) are always real.

The dot product  $\mathbf{n}_V \cdot \mathbf{n}_{V,s}$  in Eq. (4) can be written as

$$\mathbf{n}_V \cdot \mathbf{n}_{V,s} = \mathbf{e}_0 \cdot \mathbf{e}_{0,s} \cos \theta_V \cos \theta_{V,s} - \sin \theta_V \sin \theta_{V,s}. \quad (10)$$

Here, we take into account that  $\mathbf{e}_0 = \mathbf{e}_V$ , and  $\mathbf{e}_{0,s} = \mathbf{e}_{V,s}$ . Combining Eqs. (8)–(10) allows us to express  $\mathbf{n}_V \cdot \mathbf{n}_{V,s}$  in terms of  $\theta_0$ ,  $\theta_{0,s}$ ,  $\mathbf{e}_0$  and  $\mathbf{e}_{0,s}$ , which can be measured experimentally.

$$\cos \Theta_0 = \mathbf{s}_0 \cdot \mathbf{s}_{0,s} = \frac{\mathbf{n}_0 \cdot \mathbf{n}_{0,s} + (\mathbf{e}_0 \cos \theta_0 + \mathbf{e}_{0,s} \cos \theta_{0,s}) \cdot \mathbf{u}_0 / c_0 + u_0^2 / c_0^2}{[(1 + 2 \cos \theta_0 \mathbf{e}_0 \cdot \mathbf{u}_0 / c_0 + u_0^2 / c_0^2)(1 + 2 \cos \theta_{0,s} \mathbf{e}_{0,s} \cdot \mathbf{u}_0 / c_0 + u_0^2 / c_0^2)]^{1/2}}, \quad (11)$$

where  $\mathbf{n}_0 \cdot \mathbf{n}_{0,s} = \mathbf{e}_0 \cdot \mathbf{e}_{0,s} \cos \theta_0 \cos \theta_{0,s} - \sin \theta_0 \sin \theta_{0,s}$ .

To find a value of  $k_V$  in Eq. (4), we use a dispersion equation which relates the frequency  $\omega$  and the wavevector  $\mathbf{k}$  of a plane wave in a moving medium (e.g., Refs. 4 and 7):

$$\omega = kc + \mathbf{k} \cdot \mathbf{v}. \quad (12)$$

According to this equation, the frequency of the emitted wave near the source-receiver height is given by  $\omega_0 = k_0 c_0 + \mathbf{k}_0 \cdot \mathbf{u}_0$  while that near the scattering volume  $\omega_V = k_V c_V + \mathbf{k}_V \cdot \mathbf{v}_V$ . Here,  $\mathbf{k}_0 = k_0 \mathbf{n}_0$  and  $\mathbf{k}_V = k_V \mathbf{n}_V$  are the wavevectors in these regions. Taking into account that the frequency of a sound wave is not changed for stationary stratification and eliminating  $\omega_0 = \omega_V$  between these equations yields

$$k_V c_V + \mathbf{k}_V \cdot \mathbf{v}_V = k_0 c_0 + \mathbf{k}_0 \cdot \mathbf{u}_0. \quad (13)$$

Using this equation and the formulas  $\mathbf{k}_V = k_V \mathbf{n}_V$  and  $\mathbf{e}_V = \mathbf{e}_0$ , one obtains the following formula for  $k_V$ :

$$k_V = k_0 \frac{c_0 + \mathbf{e}_0 \cdot \mathbf{u}_0 \cos \theta_0}{c_V + \mathbf{e}_0 \cdot \mathbf{u}_V \cos \theta_V + w_V \sin \theta_V}. \quad (14)$$

This formula and Eq. (8) express  $k$  in terms of  $\theta_0$  and  $\mathbf{e}_0$ .

Equations (4), (8)–(11) and (14) allow one to calculate the sound scattering cross-section  $\sigma$  in a stratified moving atmosphere with arbitrary profiles of  $c(z)$  and  $\mathbf{v}(z)$  as a function of the apparent scattering angle  $\Theta_0$ . Indeed, for given values of  $\theta_0$ ,  $\theta_{0,s}$ ,  $\mathbf{e}_0$  and  $\mathbf{e}_{0,s}$ , the angle  $\Theta_0$  is calculated from Eq. (11). Then,  $k_V$  is calculated using Eq. (14), and  $\mathbf{n}_V \cdot \mathbf{n}_{V,s}$  is calculated using Eqs. (8)–(10). Finally, substitution of  $k_V$  and  $\mathbf{n}_V \cdot \mathbf{n}_{V,s}$  into Eq. (4) yields  $\sigma(\Theta_0)$ . It follows from Eqs. (4), (8)–(11) and (14) that  $\sigma$  depends on values of the adiabatic sound speed and wind velocity vector at the source-receiver height and the height of the scattering volume, i.e., on  $c_0$ ,  $\mathbf{u}_0$ ,  $c_V$  and  $\mathbf{v}_V$ , but does not depend on how  $c$  and  $\mathbf{v}$  are changing between these heights. However, the position of the scattering volume  $V$  is affected by the vertical profiles of  $c(z)$  and  $\mathbf{v}(z)$ .

## D. Particular case of a stratified moving atmosphere

Equations (4), (8)–(11) and (14) which allow calculation of  $\sigma(\Theta_0)$  are rather involved. Therefore, it is worthwhile to consider a particular case of atmospheric stratification when these equations can be simplified significantly and can be studied analytically. We chose the following stratification:  $c(z) = c_0$ ,  $\mathbf{u}_0 = 0$ , and  $\mathbf{v} = \mathbf{v}(z)$ . Namely this atmospheric

Note that for an observer at the source-receiver height, the apparent scattering angle  $\Theta_0$  is the angle between the vectors  $\mathbf{s}_0$  and  $\mathbf{s}_{0,s}$ . Using the formula  $\mathbf{s} = (c\mathbf{n} + \mathbf{v}) / |c\mathbf{n} + \mathbf{v}|$ , the angle  $\Theta_0$  can be expressed in terms of  $\theta_0$ ,  $\theta_{0,s}$ ,  $\mathbf{e}_0$  and  $\mathbf{e}_{0,s}$ :

stratification is approximated by the model of an atmosphere used in Refs. 1 and 6 as two homogeneous layers in relative motion.

For the considered stratification,  $\mathbf{n}_0 = \mathbf{s}_0$  and  $\mathbf{n}_{0,s} = \mathbf{s}_{0,s}$  so the apparent scattering angle  $\Theta_0$  coincides with the angle between vectors  $\mathbf{n}_0$  and  $\mathbf{n}_{0,s}$ , and Eq. (11) simplifies significantly:  $\cos \Theta_0 = \mathbf{n}_0 \cdot \mathbf{n}_{0,s}$ . Furthermore, Eqs. (8), (9) and (14) take the form

$$\cos \theta_V = \cos \theta_0 \frac{1 - M + \beta \sqrt{(1 - M)^2 - \cos^2 \theta_0 (1 - \beta^2)}}{(1 - M)^2 + \beta^2 \cos^2 \theta_0}, \quad (15)$$

$$\cos \theta_{V,s} = \cos \theta_{0,s} \frac{1 - M_s - \beta \sqrt{(1 - M_s)^2 - \cos^2 \theta_{0,s} (1 - \beta^2)}}{(1 - M_s)^2 + \beta^2 \cos^2 \theta_{0,s}}, \quad (16)$$

$$k_V = k_0 \frac{1}{1 + (\mathbf{e}_0 \cdot \mathbf{u}_V / c_0) \cos \theta_V + \beta \sin \theta_V}. \quad (17)$$

Note that to the order  $v_V / c_0$  formula (17) reads

$$k_V = k_0 (1 - \mathbf{n}_0 \cdot \mathbf{v}_V / c_0). \quad (18)$$

If  $w_V = 0$ , Eqs. (17) and (18) are equivalent to any order in  $v_V / c_0$  since in this case  $1 - \mathbf{n}_0 \cdot \mathbf{v}_V / c_0 = (1 + \mathbf{n}_V \cdot \mathbf{v}_V / c_0)^{-1}$ . This formula follows from Eq. (13) where  $c_V = c_0$  and  $\mathbf{u}_0 = 0$  and the formula  $\mathbf{k}_V \cdot \mathbf{v}_V = \mathbf{k}_0 \cdot \mathbf{v}_V$ , which is valid for the considered case  $w_V = 0$ .

Substituting the values of  $\theta_V$  and  $\theta_{V,s}$  from Eqs. (15) and (16) into Eq. (10) and retaining only terms of order  $v_V / c_0$ , the formula for  $\mathbf{n}_V \cdot \mathbf{n}_{V,s}$  can be simplified significantly:

$$\mathbf{n}_V \cdot \mathbf{n}_{V,s} = \left( 1 + \frac{\mathbf{n}_0 \cdot \mathbf{v}_V}{c_0} + \frac{\mathbf{n}_{0,s} \cdot \mathbf{v}_V}{c_0} \right) \cos \Theta_0 + \frac{\sin \theta_{0,s}}{\sin \theta_0} \frac{\mathbf{n}_0 \cdot \mathbf{v}_V}{c_0} + \frac{\sin \theta_0}{\sin \theta_{0,s}} \frac{\mathbf{n}_{0,s} \cdot \mathbf{v}_V}{c_0}. \quad (19)$$

Using Eqs. (10), (15) and (16), consider the important case of backscattering for an observer in the ground-fixed coordinate system  $K$  when  $\theta_0 = \theta_{0,s}$ ,  $\mathbf{e}_0 = -\mathbf{e}_{0,s}$ , and  $\Theta_0 = \pi$ . In accordance with the nonreciprocity of sound propagation in a moving medium, which is well known in the literature, the true scattering angle  $\Theta$  between the vectors  $\mathbf{n}_V$  and  $\mathbf{n}_{V,s}$  should be less than  $\pi$ . This can also be shown from Eqs. (15) and (16). Indeed, it follows from these equations that  $\cos \theta_V - \cos \theta_{V,s} = 2M \cos \theta_0$  to the order of  $v_V / c_0$ , so

$\Delta\theta = \theta_V - \theta_{V,s} = 2M \cot \theta_0$ . [Here, we assume for simplicity that  $\beta=0$  and all vectors in Eqs. (10), (15) and (16) are located in one vertical plane.] Therefore, the true scattering angle  $\Theta = \pi - \Delta\theta = \pi - 2M \cot \theta_0$  is less than  $\pi$ , in accordance with nonreciprocity. Nevertheless, the dot product  $\mathbf{n}_V \cdot \mathbf{n}_{V,s} = \cos \Theta = -1 + 2M^2 \cot^2 \theta_0$  and hence  $\mathbf{n}_V \cdot \mathbf{n}_{V,s} = -1$  to the order of  $v_V/c_0$ , just as if  $\Theta = \pi$  and reciprocity held. Since the formula for  $\sigma$  contains  $\mathbf{n}_V \cdot \mathbf{n}_{V,s}$ , it also does not reveal nonreciprocity of sound propagation for sound backscattering to order  $v_V/c_0$ . Note that Eq. (19) obtained to this order also predicts that  $\mathbf{n}_V \cdot \mathbf{n}_{V,s} = -1$  for the considered case of backscattering.

## II. NUMERICAL RESULTS

In this section, the dependence of  $\sigma$  on  $\Theta_0$  is studied numerically using Eqs. (4), (8)–(11) and (14). For simplicity, we assume that  $\mathbf{u}_0 = 0$ . As shown above, in this case the apparent scattering angle  $\Theta_0$  is the angle between the vectors  $\mathbf{n}_0$  and  $\mathbf{n}_{0,s}$ . Furthermore, we assume that the vectors  $\mathbf{u}_V$ ,  $\mathbf{e}_0$  and  $\mathbf{e}_{0,s}$  are located in one vertical plane, and  $c_0 = 340$  m/s. It is also convenient to denote  $\Delta c = c_V - c_0$ , and introduce the normalized temperature contribution to the sound scattering cross-section  $\hat{\sigma}_T(\Theta_0) = [2^{5/6} 72 \Gamma(1/3) T_V^2 / 5 k_0^{1/3} C_T^2] \sigma_T(\Theta_0)$  and the normalized wind velocity contribution  $\hat{\sigma}_v(\Theta_0) = [2^{5/6} 216 \Gamma(1/3) c_V^2 / 55 k_0^{1/3} C_v^2] \sigma_v(\Theta_0)$ .

Let us first study the case  $\mathbf{u}_V = 0$ . For this case, the normalized temperature  $\hat{\sigma}_T(\Theta_0)$  and wind velocity  $\hat{\sigma}_v(\Theta_0)$  contributions to  $\sigma$  are plotted in Fig. 2(a) and (b), respectively, for  $w_V = 0$  and  $\theta_0 = 35^\circ$ . Solid lines correspond to the case  $\Delta c = 0$ . In this case,  $\hat{\sigma}_T(\Theta_0)$  and  $\hat{\sigma}_v(\Theta_0)$  coincide with those calculated from Eq. (1). Dashed and dash-dotted lines correspond to  $\Delta c = 6$  m/s and  $\Delta c = -6$  m/s, respectively. It is seen from Fig. 2(a) and (b) that the solid, dashed and dash-dotted lines are significantly different for  $\Theta_0 < 110^\circ$ , but have nearly the same shape. A positive (negative) value of  $\Delta c$  causes a shift of the minima in  $\sigma_T$  and  $\sigma_v$ , which are located at  $\Theta_0 = 90^\circ$  for  $\Delta c = 0$ , towards larger (smaller) values of  $\Theta_0$ . According to Fig. 2(b), for  $50^\circ < \Theta_0 < 70^\circ$  the value of  $\hat{\sigma}_v(\Theta_0)$  for  $\Delta c = 0$  differs from those for  $\Delta c = \pm 6$  m/s by factors in the range from 0.5 to 1.7. The same factors would be the difference between  $C_v^2$  retrieved from measurements of  $\sigma$  by bistatic acoustic sounding with the use of Eq. (3) and its value obtained with the use of correct equations (4), (8)–(11) and (14). Thus, even for moderate values of  $\Delta c = \pm 6$  m/s, in the range of scattering angles  $50^\circ < \Theta_0 < 70^\circ$ , existing techniques of remote sensing of  $C_v^2$  should use the correct equations (4), (8)–(11) and (14) for  $\sigma$ , rather than Eq. (3).

For the considered case  $\mathbf{u}_V = 0$ , we have also calculated  $\hat{\sigma}_T(\Theta_0)$  and  $\hat{\sigma}_v(\Theta_0)$  for  $w_V = \pm 2$  m/s and other parameters as in Fig. 2. In this case,  $\hat{\sigma}_T(\Theta_0)$  and  $\hat{\sigma}_v(\Theta_0)$  are given practically by the same curves as in those figures.

Let us now study the case  $\Delta c = 0$ . The normalized temperature  $\hat{\sigma}_T(\Theta_0)$  and wind velocity  $\hat{\sigma}_v(\Theta_0)$  contributions to  $\sigma$  are plotted in Fig. 3(a) and (b), respectively, for  $w_V = 0$  and  $\theta_0 = 35^\circ$ . Solid, dashed, and dash-dotted lines correspond to  $u_V = 0$ ,  $u_V = 12$  m/s and  $u_V = -12$  m/s, respectively.

These lines are significantly different for  $\Theta_0 < 110^\circ$  in Fig. 3(a), and for  $\Theta_0 < 110^\circ$  and  $\Theta_0$  near  $180^\circ$  in Fig. 3(b). Furthermore, finite values of  $u_V$  cause shifts in the minima of  $\sigma_T$  and  $\sigma_v$  located at  $\Theta_0 = 90^\circ$  for  $u_V = 0$ . It follows from Fig. 3(b), that for  $55^\circ < \Theta_0 < 70^\circ$  the value of  $\hat{\sigma}_v(\Theta_0)$  for  $u_V = 0$  differs from those for  $u_V = \pm 12$  m/s by factors in the range from 0.3 to 2.3. This result is similar to that discussed above in connection with Fig. 2(b). Therefore, we can make a similar conclusion: Even for moderate values of  $u_V = \pm 12$  m/s and at least in the range of scattering angles  $55^\circ < \Theta_0 < 70^\circ$ , existing bistatic techniques for remote sensing of  $C_v^2$  should use correct equations (4), (8)–(11) and (14) for  $\sigma$  rather than Eq. (3).

For the considered case  $\Delta c = 0$ , we have also calculated  $\hat{\sigma}_T(\Theta_0)$  and  $\hat{\sigma}_v(\Theta_0)$  for  $w_V$  ranging from  $-2$  to  $2$  m/s and other parameters as in Fig. 3. We have found that  $\hat{\sigma}_T(\Theta_0)$  and  $\hat{\sigma}_v(\Theta_0)$  are given by practically the same curves as in those figures.

Finally, Fig. 4(a) and (b) shows the normalized scattering cross-section  $\hat{\sigma}(\Theta_0) = \sigma(\Theta_0) [2^{5/6} 72 \Gamma(1/3) T_V^2 / 5 k_0^{1/3} C_T^2]$  for the case when the grazing angle  $\theta_0$  of the emitted wave is  $35^\circ$ . In Fig. 4,  $\Delta c = 0$ ,  $w_V = 0$ , and solid lines correspond to  $u_V = 0$ , while other lines correspond to  $u_V = \pm 7.5$  and  $\pm 15$  m/s. Furthermore,  $C_v^2/c_V^2 = 10 C_T^2/T_V^2$  [Fig. 4(a)], and  $C_v^2/c_V^2 = 100 C_T^2/T_V^2$  [Fig. 4(b)]. It follows from Fig. 4(a), that for  $C_v^2/c_V^2 = 10 C_T^2/T_V^2$  the values of  $\sigma(180^\circ)$  for  $u_V = \pm 7.5$  m/s and  $\pm 15$  m/s are different from that for  $u_V = 0$  by factors 1.05 and 1.2, respectively. This difference can probably be ignored in remote sensing of  $C_T^2$ . However, for  $C_v^2/c_V^2 = 100 C_T^2/T_V^2$  [see Fig. 4(b)] the values of  $\sigma(180^\circ)$  for  $u_V = \pm 7.5$  m/s and  $\pm 15$  m/s are different from that for  $u_V = 0$  by factors 1.5 and 3, respectively. This difference is significant and should be taken into account. Thus, for relatively large values of  $C_v^2$ , wind velocity fluctuations contribute significantly to  $\sigma(180^\circ)$  and Eq. (2) cannot be used anymore to retrieve  $C_T^2$ . In this case, a signal measured by a receiver is proportional to both  $C_T^2$  and  $C_v^2$ . If  $C_v^2$  is known (e.g., as a result of bistatic acoustic sounding), then the measured signal and Eqs. (4), (8)–(11) and (14) allow retrieval of  $C_T^2$ .

Note that the values of  $\hat{\sigma}(180^\circ)$  from Fig. 4 are practically not changed for  $\Delta c$  in the range from  $-6$  to  $6$  m/s and  $w_V$  in the range from  $-2$  to  $2$  m/s if other parameters remain unchanged.

## III. TWO HOMOGENEOUS LAYERS IN RELATIVE MOTION

### A. Geometry of the problem

Sections I and II gave a rigorous derivation and detailed numerical analysis of the equation for  $\sigma$  in the atmosphere with arbitrary profiles of  $c(z)$  and  $\mathbf{v}(z)$ . This accomplishes the main goal of the paper. In the present section we compare the derived equation with those in Refs. 1 and 6 that were obtained for the model of an atmosphere as two homogeneous layers in relative motion, and explain why Clifford and Brown's predictions are different from Ye's.

The geometry of the problem is shown in Fig. 5. This

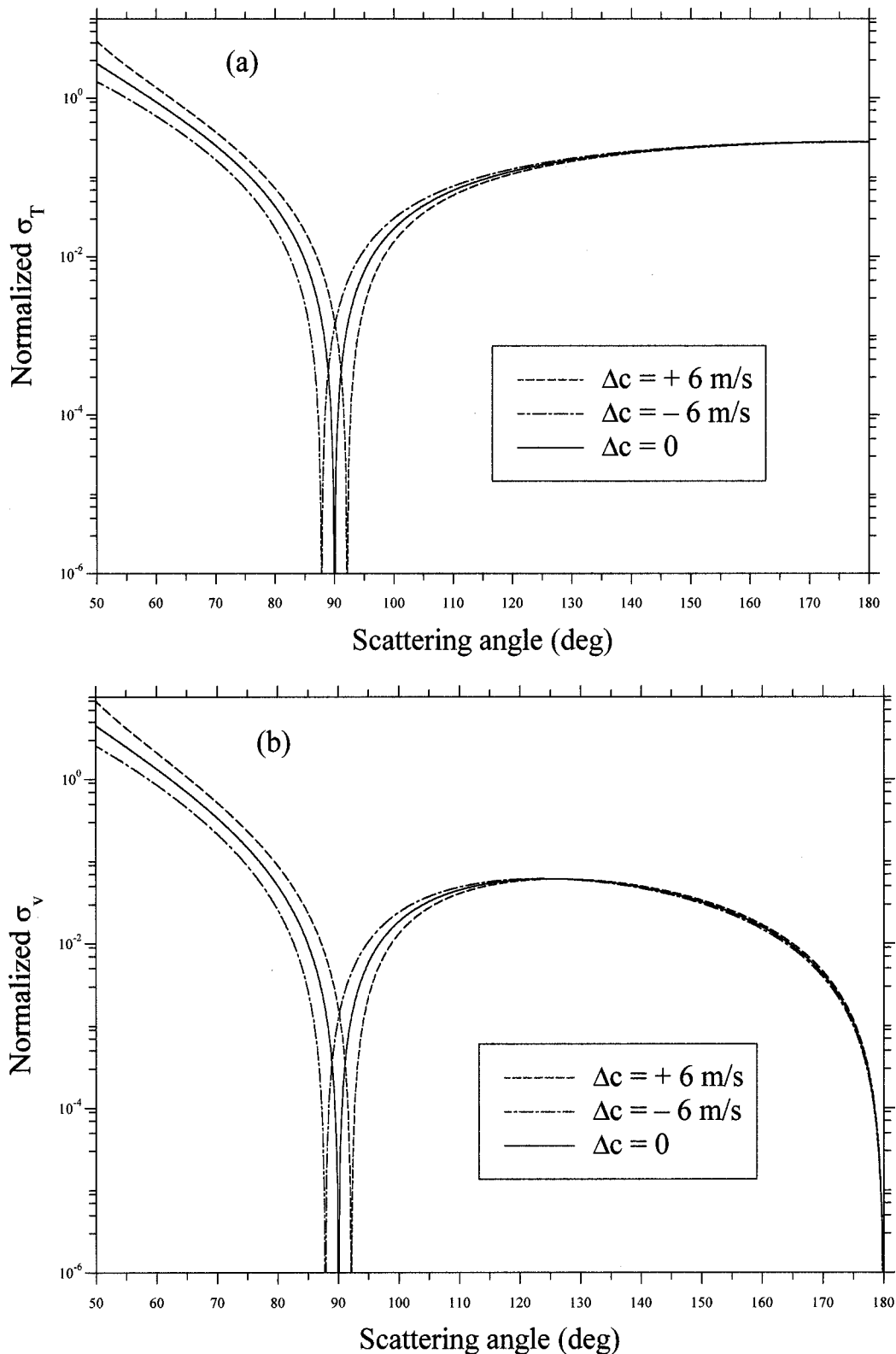


FIG. 2. The normalized temperature  $\hat{\sigma}_T$  (a) and wind velocity  $\hat{\sigma}_v$  (b) contributions to the sound scattering cross-section versus the apparent scattering angle  $\Theta_0$  for  $u_V=0$ ,  $\theta_0=35^\circ$  and different values of  $\Delta c$ .

geometry is the same as that in Ref. 1, and differs from that in Fig. 1 only by the vertical profiles of  $c$  and  $\mathbf{v}$ . The mean wind velocity  $\mathbf{v}=0$  if  $z < z_i$ , and  $\mathbf{v}=(\mathbf{u}_V, w_V)$  is constant for  $z \geq z_i$ . The mean value of the adiabatic sound speed  $c_0$  is a constant for  $z \geq z_0$ .

The model of the moving medium shown in Fig. 5 can-

not actually exist because of instability of flow at the interface, etc. In fact,  $\mathbf{v}$  changes smoothly from the ground to the height where the scattering occurs, as is in the model of the atmosphere in Fig. 1. One of the ideas of Ref. 1 was that to the order  $v_V/c_0$  this smooth change in  $\mathbf{v}$  can approximately be replaced by an abrupt change at the interface  $z=z_i$  shown

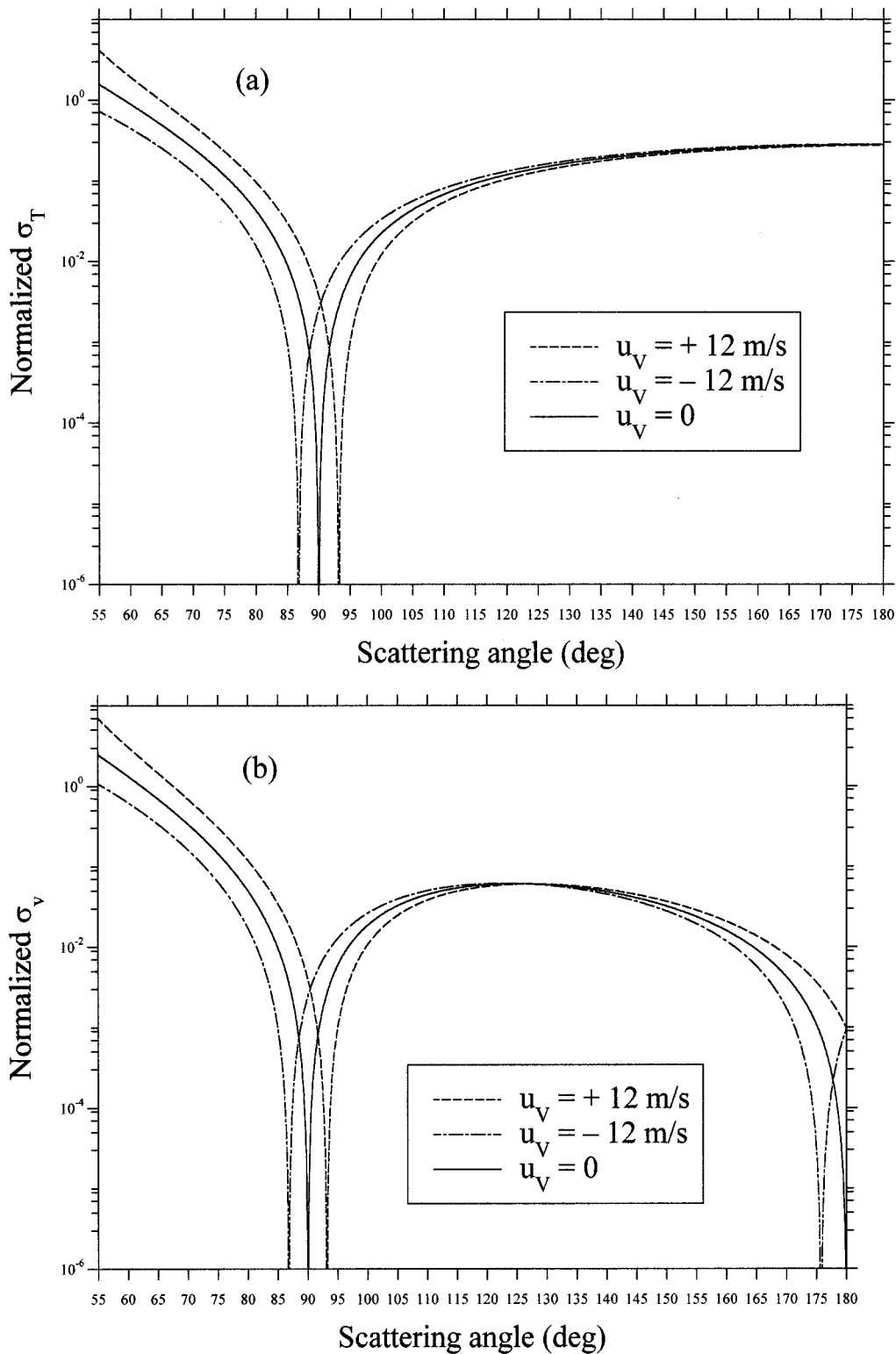


FIG. 3. The normalized temperature  $\hat{\sigma}_T$  (a) and wind velocity  $\hat{\sigma}_v$  (b) contributions to the sound scattering cross-section versus the apparent scattering angle  $\Theta_0$  for  $\Delta c=0$ ,  $\theta_0=35^\circ$  and different values of  $u_v$ .

in Fig. 5. This idea implies that the interface  $z=z_i$  is at rest with respect to the ground. Indeed, if the interface were to move vertically with the velocity  $w_v$ , the wind velocity  $\mathbf{v}$  would depend not only on  $z$  but also on the time  $t$  which is not the case considered in Ref. 1. Because of this, Clifford and Brown<sup>1</sup> did not specifically state that the interface is at rest with respect to the ground. But they did equate the fre-

quencies of the sound waves below and above the interface [see the text from Ref. 1 between Eqs. (13) and (14)]. These frequencies are the same only if the interface is at rest. In what follows the interface is assumed to be at rest with respect to the ground if it is not specifically stated.

It is stated several times in Ref. 6 that this paper treats the same problem as Clifford and Brown<sup>3</sup> did. Note that



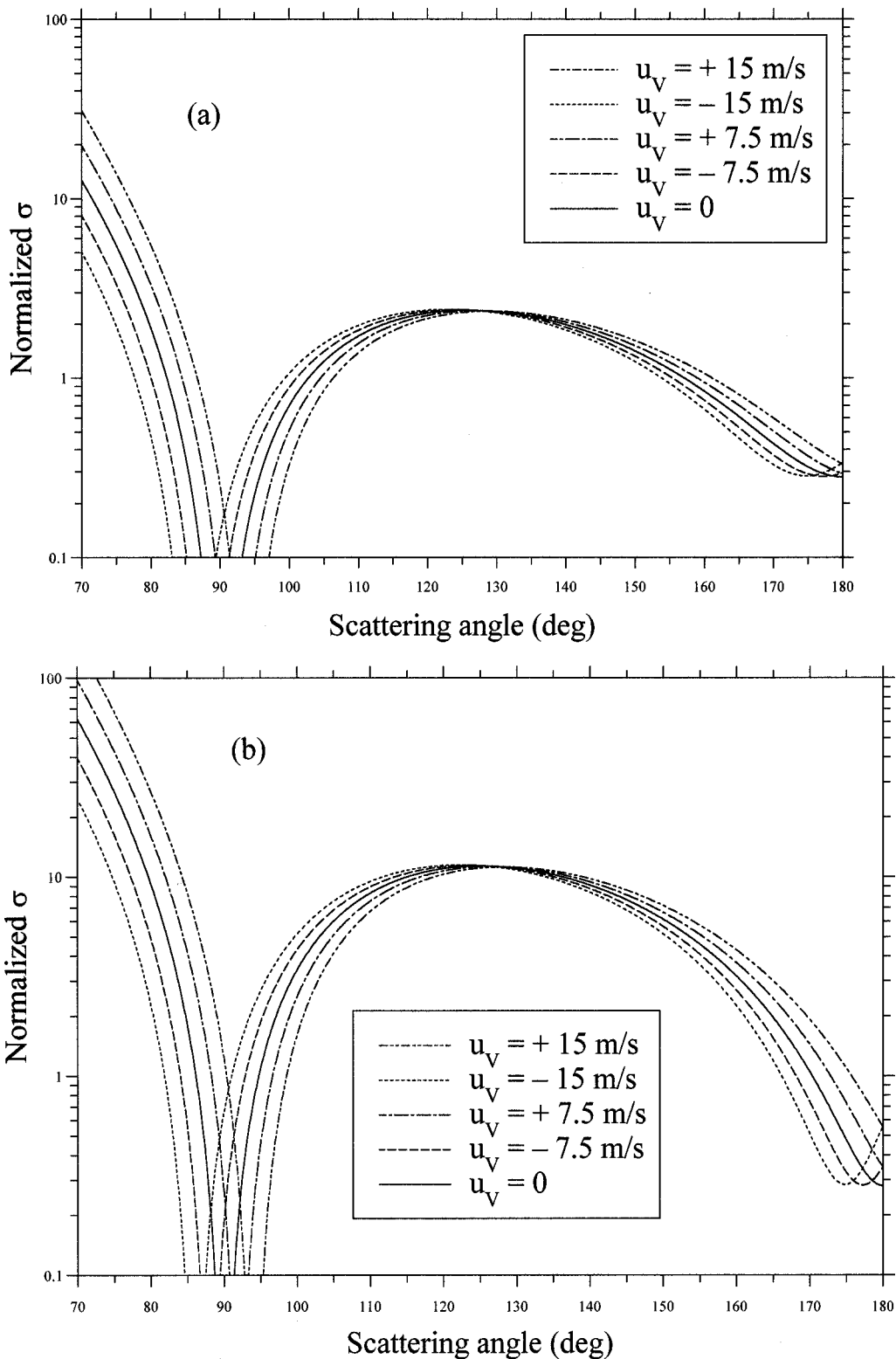


FIG. 4. The normalized sound scattering cross-section  $\hat{\sigma}$  versus the apparent scattering angle  $\Theta_0$  for  $\Delta c=0$ ,  $\theta_0=35^\circ$ , and different values of  $u_V$ ;  $C_v^2/c_v^2 = 10C_T^2/T_V^2$  (a), and  $C_v^2/c_v^2 = 100C_T^2/T_V^2$  (b).

there is no direct statement in Ref. 6 about whether the interface is at rest. However, it was assumed that the frequencies of the waves as observed in the coordinate system  $K_1$  moving with velocity  $\mathbf{v}_V$  are the same on both sides of the interface [see the text from Ref. 6 between Eqs. (9) and (10)]. This assumption is correct only if the interface is at rest in the coordinate system  $K_1$  and is therefore moving

with the vertical velocity  $w_V$  with respect to the ground-fixed coordinate system  $K$ . This is one of the differences between Refs. 6 and 1.

The region of space  $z > z_i$  in Fig. 5 is similar to the region  $D$  in Fig. 1: In both of these regions  $c$  and  $\mathbf{v}$  are constant. Therefore, the sound scattering cross-section  $\sigma$  will still be given by Eq. (4), where  $\mathbf{n}_V$  and  $\mathbf{n}_{V,s}$  are the unit

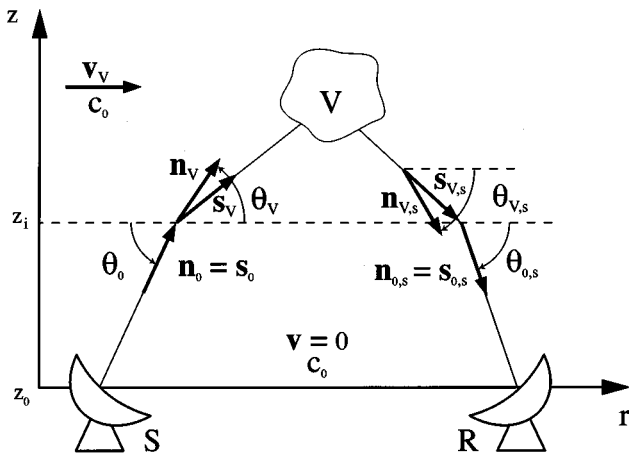


FIG. 5. Geometry of the problem for a model of an atmosphere as two homogeneous layers in relative motion. S is source; R is receiver;  $\mathbf{n}_0$  and  $\mathbf{n}_V$  are the unit vectors normal to the wavefronts of the emitted wave;  $\mathbf{n}_{V,s}$  and  $\mathbf{n}_{0,s}$  are the unit vectors normal to the wavefronts of the scattered wave; and  $\theta_0$ ,  $\theta_V$ ,  $\theta_{V,s}$  and  $\theta_{0,s}$  are the grazing angles of these vectors.

vectors normal to the wavefront of the emitted and scattered waves for  $z > z_i$ , and  $\mathbf{k}_V$  is the wavevector of the emitted wave in this region. In Eq. (4), the values of  $\mathbf{n}_V \cdot \mathbf{n}_{V,s}$  and  $k_V$  need to be determined in terms of the unit vectors in the directions of the emitted wave near source  $\mathbf{n}_0$  and the scattered wave near receiver  $\mathbf{n}_{0,s}$ . To do this, a refraction law for the normal to the wavefront at the interface  $z = z_i$  is used in Refs. 1 and 6.

## B. Derivation of the refraction law

We will derive, by a rigorous approach, the refraction law for the normal to the wavefront for the geometry of the problem shown in Fig. 5. This derivation will be done in the ground-fixed coordinate system  $K$ . First consider a plane wave emitted by a source. If  $z < z_i$ , the sound pressure  $p$  (in arbitrary units) is given by

$$p = \exp(i\mathbf{a}_0 \cdot \mathbf{r} + iq_0 z - i\omega_0 t) + p_r. \quad (20)$$

Here, the first term on the right hand side describes the emitted plane wave, and  $p_r$  is the wave reflected at the interface  $z = z_i$ . The reflected wave does not need to be considered in detail when deriving a refraction law. Furthermore, in Eq. (20)  $\mathbf{r} = (x, y)$  are the horizontal Cartesian coordinates, and  $\mathbf{a}_0$  and  $q_0$  are the horizontal and vertical components of the wavevector  $\mathbf{k}_0 = (\mathbf{a}_0, q_0)$  of the emitted wave. For  $z > z_i$ , the sound pressure of the refracted wave is given by

$$p = A \exp(i\mathbf{a}_V \cdot \mathbf{r} + iq_V z - i\omega_V t), \quad (21)$$

where  $\mathbf{a}_V$  and  $q_V$  are the horizontal and vertical components of the wavevector  $\mathbf{k}_V = (\mathbf{a}_V, q_V)$  of the refracted wave, and  $A$  is the transmission coefficient across the interface.

As in Sec. I, it is convenient to express  $\mathbf{n}_0$  and  $\mathbf{n}_V$  in the form  $\mathbf{n}_0 = (\mathbf{e}_0 \cos \theta_0, \sin \theta_0)$  and  $\mathbf{n}_V = (\mathbf{e}_V \cos \theta_V, \sin \theta_V)$ . A refraction law for the normal to the wavefront relates the angles  $\theta_0$  and  $\theta_V$ , and the vectors  $\mathbf{e}_0$  and  $\mathbf{e}_V$ . To find this law, we equate the sound pressure  $p$  on both sides of the interface  $z = z_i$ :

$$\exp(i\mathbf{a}_0 \cdot \mathbf{r} + iq_0 z_i - i\omega_0 t) + p_r = A \exp(i\mathbf{a}_V \cdot \mathbf{r} + iq_V z_i - i\omega_V t). \quad (22)$$

Equating fields on both sides of an interface is a rigorous approach for deriving refraction laws in many fields of physics. Equation (22) can be satisfied at all  $\mathbf{r}$  and  $t$  only if

$$\mathbf{a}_0 = \mathbf{a}_V, \quad \mathbf{e}_0 = \mathbf{e}_V, \quad \omega_0 = \omega_V. \quad (23)$$

Thus, the azimuthal directions of the normal to the wavefronts of the emitted and refracted waves coincide. The same result has been obtained in Refs. 1 and 6.

The dispersion equation (12) is also valid for the model of an atmosphere shown in Fig. 5. So do Eq. (13) (where  $\mathbf{u}_0$  should be zero) and Eq. (17) for  $k_V$ , which are both obtained from Eq. (12). For the considered model, Eq. (13) can be written as  $k_V + \mathbf{k}_V \cdot \mathbf{v}_V / c_0 = k_0$ . Replacing  $\mathbf{k}_V \cdot \mathbf{v}_V$  in this equation by  $\mathbf{a}_V \cdot \mathbf{u}_V + q_V w_V$ , dividing both sides of the resulting equation by  $a_V = a_0$  and taking into account that  $\cos \theta_0 = a_0 / k_0$ ,  $\cos \theta_V = a_V / k_V$  and  $q_V / a_V = \tan \theta_V$ , we obtain

$$\frac{c_0}{\cos \theta_V} + \mathbf{e}_V \cdot \mathbf{u}_V + w_V \tan \theta_V = \frac{c_0}{\cos \theta_0}. \quad (24)$$

This expresses the refraction law for the normal to the wavefront at the fixed interface  $z = z_i$  of two homogeneous layers in relative motion. Equation (24) coincides with the refraction law for the normal to the wavefront obtained by Clifford and Brown<sup>1</sup> [see Eq. (14) from Ref. 1], which has been derived by a different approach. However, Eq. (24) does not coincide with the refraction law from Ref. 6 [Eq. (15) from Ref. 6], which in our notation takes the form

$$\frac{c_0}{\cos \theta_V} + \mathbf{e}_V \cdot \mathbf{u}_V + w_V \tan \theta_0 = \frac{c_0}{\cos \theta_0}. \quad (25)$$

Equations (24) and (25) coincide if  $w_V = 0$  and are different for  $w_V \neq 0$ . This difference is explained by the fact that in Ref. 6 the frequency of a sound wave, as observed in the comoving coordinate system  $K_1$ , is assumed to be the same on both sides of the interface. As noted above, this is correct only if the interface itself is moving vertically with the velocity  $w_V$  with respect to the ground-fixed coordinate system  $K$ . In this case, the sound field of the emitted wave, Eq. (20), should be equated to the sound field of the refracted wave, Eq. (21), at the moving interface  $z = z_i + w_V t$ . This results in

$$\begin{aligned} \exp(i\mathbf{a}_0 \cdot \mathbf{r} + iq_0(z_i + w_V t) - i\omega_0 t) + p_r \\ = A \exp(i\mathbf{a}_V \cdot \mathbf{r} + iq_V(z_i + w_V t) - i\omega_V t). \end{aligned} \quad (26)$$

Using this equation and the same approach that follows Eq. (22), one obtains the refraction law of Eq. (25).

We have also derived the refraction law for the normal to the wavefront using quantities expressed in the coordinate system  $K_1$  moving with the velocity  $\mathbf{v}_V$ . This was the method used in Ref. 6. Our calculations show that the refraction law is still given by Eq. (24) if the interface is at rest in the ground-fixed coordinate system  $K$ , and by Eq. (25) if the interface is moving with the velocity  $w_V$ .

Consider now the scattered wave. As in Sec. I, let  $\mathbf{n}_{V,s} = (\mathbf{e}_{V,s} \cos \theta_{V,s}, -\sin \theta_{V,s})$  and  $\mathbf{n}_{0,s} = (\mathbf{e}_{0,s} \cos \theta_{0,s}, -\sin \theta_{0,s})$ . Then it can be shown by the approach which has been used to derive Eq. (24) that

$$\frac{c_0}{\cos \theta_{V,s}} + \mathbf{e}_{V,s} \cdot \mathbf{u}_V - w_V \tan \theta_{V,s} = \frac{c_0}{\cos \theta_{0,s}}, \quad \mathbf{e}_{V,s} = \mathbf{e}_{0,s}. \quad (27)$$

This is the refraction law for the normal to the wavefront of the scattered wave.

The refraction laws (24) and (27) allow us to express  $\mathbf{n}_V \cdot \mathbf{n}_{V,s}$  in terms of  $\mathbf{n}_0$  and  $\mathbf{n}_{0,s}$ . Indeed,  $\mathbf{n}_V \cdot \mathbf{n}_{V,s}$  is given by formula (10). The grazing angles  $\theta_V$  and  $\theta_{V,s}$  in this formula can be found as solutions of Eqs. (24) and (27). The results exactly coincide with Eqs. (15) and (16) obtained for a particular case of a stratified moving atmosphere.

Therefore, for the model of an atmosphere as two homogeneous layers in relative motion,  $\sigma$  is given by Eq. (4), where  $k_V$  is given by Eq. (17) and  $\mathbf{n}_V \cdot \mathbf{n}_{V,s}$  is given by Eqs. (10), (15) and (16). Furthermore, to order  $v_V/c_0$ , the formula for  $\mathbf{n}_V \cdot \mathbf{n}_{V,s}$  is given by Eq. (19), and consideration of the important case of backscattering, following Eq. (19), still holds. Thus, we have rigorously shown that for the model of an atmosphere as two homogeneous layers in relative motion, the sound scattering cross-section  $\sigma$  exactly coincides with  $\sigma$  for a particular case of a stratified moving atmosphere considered in Sec. ID.

### C. Comparisons among results

Let us now compare Eqs. (4), (10), and (15)–(19) for  $\sigma$  with analogous equations from Refs. 1 and 6.

In Ref. 1, the equation for  $\sigma$  coincides with Eq. (4) while  $k_V$  and  $\mathbf{n}_V \cdot \mathbf{n}_{V,s}$  are calculated to order  $v_V/c_0$ . The corresponding equations coincide with Eqs. (18) and (19). Thus, the present analysis has shown that, under the assumptions and approximations made in Ref. 1, all calculations are correct. Reference 1 shows no change in backscattering due to wind velocity  $\mathbf{v}_V$ . This is in complete agreement with the discussion following Eq. (19).

In Ref. 6,  $\mathbf{n}_V \cdot \mathbf{n}_{V,s}$  and  $k_V$  are calculated to any order in  $v_V/c_0$ . Equation (4) for  $\sigma$  coincides with Eq. (25) from Ref. 6 with one difference. In Ref. 6,  $k_V$  is given by Eq. (18), rather than by Eq. (17). The reason for this is the assumption made in Ref. 6 that the sound frequency as observed in the comoving coordinate system  $K_1$  remains unchanged across the interface. This is correct only if the interface is moving with the velocity  $w_V$  in the ground-fixed coordinate system  $K$  and is not correct if the interface is at rest. If  $w_V=0$ , Eqs. (17) and (18) are equivalent.

The equation for  $\mathbf{n}_V \cdot \mathbf{n}_{V,s}$  in Ref. 6 [Eq. (21) from Ref. 6] coincides with Eqs. (10), (15) and (16) if  $w_V=0$  and differs from them if  $w_V \neq 0$  for the same reason. Figures 3 and 4 from Ref. 6 show the angular dependence of  $\sigma_v$  and  $\sigma_T$ . Since these figures are obtained for the case  $w_V=0$ , they would agree with those obtained from Eqs. (4), (10) and (15)–(17). So it is a valid and interesting consequence of these equations that there is a change in apparent backscattering in a moving atmosphere in comparison to the case  $\mathbf{v}_V=0$ . This effect is of order  $v_V^2/c_0^2$ .

It should be noted that the formula for  $\mathbf{n}_V \cdot \mathbf{n}_{V,s}$  obtained in Ref. 6 [see Eq. (21) from Ref. 6] was not compared with

that obtained in Ref. 1 [see Eq. (25) from Ref. 1]. Comparison reveals that, to order  $v_V/c_0$ , Eq. (21) from Ref. 6 exactly coincides with Eq. (25) from Ref. 1 if  $w_V=0$ . Therefore, in this case, neither Eq. (21) from Ref. 6 nor Eq. (25) from Ref. 1 satisfy nonreciprocity for backscattering.

## IV. CONCLUSIONS

We have derived the equation for the sound scattering cross-section  $\sigma$  in a stratified moving atmosphere with arbitrary profiles of the adiabatic sound speed  $c(z)$  and the wind velocity vector  $\mathbf{v}(z)$ . It is shown that  $\sigma$  depends on the values of  $c$  and  $\mathbf{v}$  at the source-receiver height and the height of the scattering volume, and does not depend on  $c$  and  $\mathbf{v}$  in regions between these heights.

Using the derived equation for  $\sigma$ , we have numerically studied the effects of atmospheric stratification on the sound scattering cross-section. We have shown that the difference  $\Delta c$  in the adiabatic sound speed between the source-receiver height and the height of the scattering volume does not practically affect  $\sigma_T$  and  $\sigma_v$  for the apparent scattering angles  $\Theta_0$  close to  $180^\circ$ , but can affect  $\sigma_T$  and  $\sigma_v$  for  $\Theta_0 < 110^\circ$ . Furthermore, we have shown that a finite value  $u_V$  of the horizontal component of the wind velocity vector can significantly affect  $\sigma_T$  for scattering angles  $\Theta_0 < 110^\circ$ , and  $\sigma_v$  for  $\Theta_0 < 110^\circ$  and  $\Theta_0 \approx 180^\circ$ . We have also shown that for moderate horizontal wind and relatively large values of  $C_v^2$ , wind velocity fluctuations can contribute significantly to the backscattering cross-section.

This numerical analysis allows us to conclude that, for many geometries and even moderate values of  $\Delta c$  and  $u_V$ , the existing monostatic and bistatic techniques for acoustic remote sensing of  $C_T^2$  and  $C_v^2$  should use correct Eqs. (4), (8)–(11) and (14) for  $\sigma$  rather than Eqs. (1)–(3) used in the literature.

Finally, using a rigorous approach, different from those in Refs. 1 and 6, we have derived the equation for the sound scattering cross-section  $\sigma$  and the formulas for  $\mathbf{n}_V \cdot \mathbf{n}_{V,s}$  and  $k_V$  for a model of a stratified moving atmosphere as two homogeneous layers in relative motion. Our approach has allowed us to check results obtained in Refs. 1 and 6 and to compare these results.

All calculations and conclusions made in Ref. 1 are correct to the order of  $v_V/c_0$  assumed in the paper. All results in Ref. 6 are also correct if  $w_V=0$ . However, for  $w_V \neq 0$ , equations for  $\mathbf{n}_V \cdot \mathbf{n}_{V,s}$  and  $k_V$  in Ref. 6 and hence the resulting equation for  $\sigma$  are correct only if the interface between the two homogeneous layers is moving vertically with the velocity  $w_V$  and are not correct if the interface is at rest. The difference in the predicted sound backscattering in Refs. 1 and 6 is explained by the difference in accuracy in  $v_V/c_0$  to which the dot product  $\mathbf{n}_V \cdot \mathbf{n}_{V,s}$  is calculated in those papers. Formula (10) for  $\mathbf{n}_V \cdot \mathbf{n}_{V,s} = \cos \Theta$  reveals nonreciprocity of sound propagation in a moving medium, well known in the literature. However, to the order of  $v_V/c_0$  and for sound backscattering, this formula reduces to Eq. (19) which does not satisfy nonreciprocity. The same result follows from Ref. 1.

## ACKNOWLEDGMENTS

This material is partly based upon work done by V. E. Ostashev, G. H. Goedecke and R. Wood that was supported by the U.S. Army Research Office under Contract Nos. DAAH04-95-1-0593, DAAG55-97-1-0178 and DAAG55-98-1-0463. We also thank Z. Ye for his cooperation in clarifying some issues concerning Ref. 6 and the Associate Editor Louis Sutherland for his suggestions in writing the present paper.

<sup>1</sup>S. F. Clifford and E. H. Brown, "Acoustic scattering from a moving turbulent medium," *J. Acoust. Soc. Am.* **55**, 929–933 (1974).

<sup>2</sup>S. P. Singal, "Acoustic sounding stability studies," in *Encyclopedia of*

*Environment Control Technology. Vol. 2: Air Pollution Control* (Gulf Publishing, Houston, 1989), pp. 1003–1061.

<sup>3</sup>A. S. Monin, "Characteristics of the scattering of sound in a turbulent atmosphere," *Sov. Phys. Acoust.* **7**, 370–373 (1962).

<sup>4</sup>V. E. Ostashev, *Acoustics in Moving Inhomogeneous Media* (E & FN SPON, An Imprint of Thompson Professional, London, 1997).

<sup>5</sup>E. H. Brown and F. F. Hall, "Advances in atmospheric acoustics," *Rev. Geophys. Space Phys.* **16**, 47–110 (1978).

<sup>6</sup>Z. Ye, "Sound scattering from a moving turbulent medium," *J. Acoust. Soc. Am.* **102**, 752–758 (1997).

<sup>7</sup>V. E. Ostashev, "Theory of sound propagation in an inhomogeneous moving medium (Review)," *Izv., Acad. Sci., USSR, Atmos. Oceanic Phys.* **21**, 274–285 (1985).

<sup>8</sup>G. V. Groves, "Geometrical theory of sound propagation in the atmosphere," *J. Atmos. Terr. Phys.* **7**, 113–127 (1955).

<sup>9</sup>E. T. Kornhauser, "Ray theory for moving fluids," *J. Acoust. Soc. Am.* **25**, 945–949 (1953).

# Continuous monitoring of fish in a shallow channel using a fixed horizontal sonar

Bjarke Pedersen

LIC Engineering A/S, Ehlersvej 24, DK-2900 Hellerup, Denmark

Mark V. Trevorrow

Acoustical Oceanography, Institute of Ocean Sciences, 9860 West Saanich Road, Sidney, British Columbia V8L 4B2, Canada

(Received 10 February 1998; accepted for publication 1 March 1999)

An acoustic monitoring program of herring migration in Drogden channel, near Copenhagen, Denmark was conducted from June 1996 until the end of May 1997. Fixed 100-kHz side-looking sonars provided nearly continuous surveillance in a 1-km-wide by 12-m-deep navigation channel. Water temperature, salinity, and current profiles were simultaneously monitored at this site. The sonars were positioned to insonify regions near the seabed at ranges up to 800 m, such that the typical reverberation was due to low-grazing angle seabed backscatter. It was found that under normal, weakly stratified flow conditions, fish schools attributable to herring (*Clupea harengus*) were observed from the 50- to 500-m range. This could be done despite interference from the dense vessel traffic, specifically direct echoes from hulls, propeller cavitation noise, and bubbly wakes. At close ranges (<150 m) hyperbolic trajectories attributable to individual herring were observed, with horizontal advection speeds in close agreement with measured current magnitudes. It was further observed that occasional intrusions of saline bottom waters created strong upward-refracting conditions that significantly limited the range for fish school detection. Ray-tracing analysis is used to define the insonified areas and describe the backscattered reverberation under normal and stratified flow conditions. It is shown using simulations of fish-school echoes that seabed-reflected multipaths can create an upward bias in fish-school densities calculated using echo-integration techniques. © 1999 Acoustical Society of America. [S0001-4966(99)01806-8]

PACS numbers: 43.30.Gv, 43.30.Sf [DLB]

## INTRODUCTION

Side-looking sonars offer many advantages for long-term monitoring of fish populations in shallow estuaries and channels. In shallow waters (depths ~10 m), conventional vessel-based echo-sounding operations are generally limited due to the small sampling volumes beneath the ship, vessel avoidance behavior by the fish, and operational difficulties in maintaining continuous coverage, especially under adverse weather conditions. In contrast, fixed horizontal sonar installations have the capability for fish detection at ranges many times that of the water depth, in essence creating an *acoustic net* through which migrating fish must pass. With appropriate instrument maintenance and methods for handling the large quantities of data created by such systems, continuous surveillance can be sustained. This paper describes the implementation of and examines some results from a year-long monitoring program using 100-kHz sidescan sonars in a shallow, 1-km-wide navigation channel near Copenhagen, Denmark.

Beginning in the 1960's, work reported by Weston, Revie, Andrews, and others (1971, 1989, 1990) demonstrated long-range (up to 65 km) monitoring of fish populations using a high-power military sonar installation near Perranporth (England). Since then, several vessel-based surveys in the open ocean have demonstrated fish detection and biomass estimation at intermediate ranges (Rusby *et al.*, 1973; Hewitt *et al.*, 1976; Misund, 1993; Misund *et al.*, 1995). These long-

and intermediate-range studies, while demonstrating the sensitivity of fish detection to the acoustic propagation conditions, were conducted in deep waters and thus were relatively immune to bottom reverberation effects. More recently, the use of short-range horizontal sonars for fish detection in rivers and lakes has been gaining acceptance (e.g., Gaudet, 1990; Kubecka *et al.*, 1994). However, in many of these riverine or lake applications, reverberation from surface and bottom boundaries has been specifically avoided by careful sonar-beam steering and restricting the measurement range (typically <20 m). Problems with this approach are that the fish targets are commonly found very close to the boundaries (especially the bottom) and a greater sampling range is desirable. Extending the insonified range and probing close to the boundaries necessarily invites backscattered interference and reflection focusing effects. However, recent work (Trevorrow, 1997; Trevorrow and Claytor, 1998) has shown that under typical shallow-water conditions, fish echoes can be resolved relative to the background reverberation, with some limitations imposed under breaking wave conditions and by vessel wakes.

The purpose of this project was to monitor herring (*Clupea harengus*) migration through the Drogden channel during the construction of a bridge and tunnel connecting Copenhagen, Denmark, and Malmö, Sweden [see Fig. 1(a)]. In particular, it was desired to investigate any impacts on fish migration behavior caused by dredging operations. This channel lies at the southern end of a body of water known as

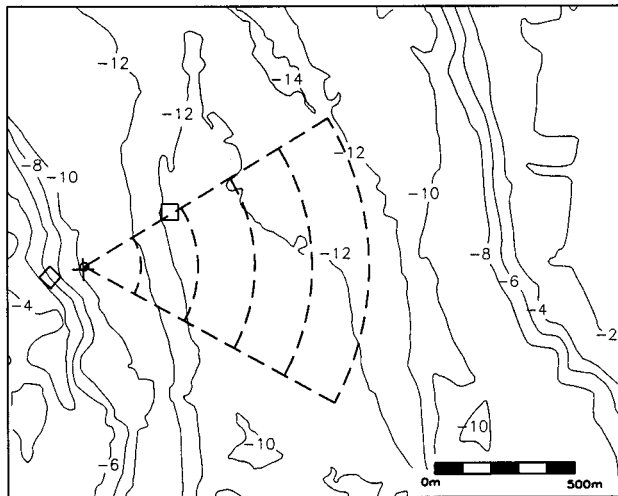
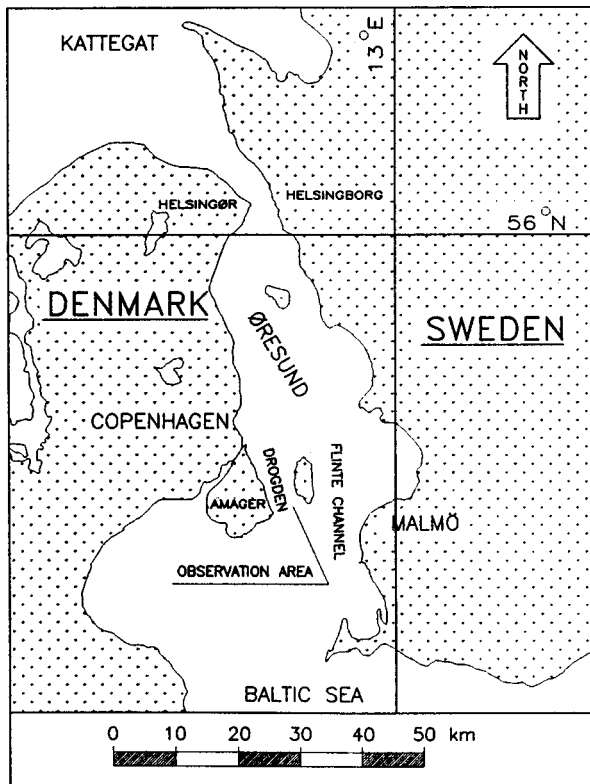


FIG. 1. (a) Map of Denmark and Sweden with location of the Øresund. (b) Close-up on Drogden Channel showing locations of sonar (circle), ADCP (square), and thermistor-conductivity sensor (cross) installations, and approximate dimensions of the sonar beam. Diamond symbol is Nordre-Røse lighthouse.

the Øresund, connecting the Kattegat with the Baltic Sea. Herring are known to migrate into the Baltic for spawning during both the spring and autumn periods, returning to the Kattegat and North Sea for feeding and over-wintering (Nielsen, 1996); however, the relative numbers of herring choosing Drogden rather than the shallower Flinte channel on the Swedish side are unknown. The data presented here were taken as part of an ongoing acoustic monitoring program started in the summer of 1996, which is only one part of an extensive environmental control program conducted during the bridge and tunnel construction by the Danish and Swed-

ish authorities. After executing a pilot project during the winter of 1995, this acoustic monitoring program was started in June of 1996. A unique feature of this project was the continuous monitoring of water temperature, salinity, and current profiles alongside the sonar installations. This enabled a more complete investigation of the environmental influences on both acoustic propagation and herring migration behavior. In this paper, we will focus on results collected during the spring herring migration season in March and April of 1997. Examples of herring schools detected above the background seabed reverberation will be shown. Also, examples of abrupt changes in acoustic propagation conditions induced by occasional saline bottom-water intrusions will be investigated. Finally, some acoustic modeling of reverberation and fish detection will be performed to explain the observations.

## I. DESCRIPTION OF INSTRUMENTATION AND MONITORING PROGRAM

The herring monitoring program was conducted across the Drogden navigation channel between Amager and Saltholm [see Fig. 1(a),(b)]. The sonar was installed on the western edge of Drogden channel, approximately 150 m east of the Nordre-Røse lighthouse. This location was near the narrowest part of the main channel, which was approximately 1 km wide by 11 to 14 m deep. There were 500-m-wide shallows (2 to 5 m deep) on either side of this main navigation channel. The sonar installations were focused on the deeper, main channel because the herring were assumed to avoid the shallows. Two 100-kHz sidescan transducers (separate transmit and receive) were mounted on a tripod 1.2 m above the seabed in an overall water depth of 10.2 m. The sidescan transducers (EDO/Western model 6400) had fan-shaped beams 3 by 60 deg (total angle to  $-3$  dB). The sonars were mounted vertically, such that their wide-beam axes were horizontal, and were pointed eastward, perpendicular to the navigation channel. The transducers were attached to a leveling mechanism on the tripod, which was operated by scuba divers using a spirit level during the installation. A 180-m-long underwater cable connected the sonars to the data acquisition system inside the lighthouse.

The data acquisition system was housed inside the Nordre-Røse lighthouse. It was comprised of a BioSonics model 101 sonar transceiver, two networked personal computers (PCs), and a radio-modem for telemetering data back to the onshore base station. One of the PCs was used for data acquisition and processing, while the other acted as a server and controller for the radio-modem. The sonar receiver provided a  $20 \log[r]$  time-variable gain (TVG) compensation followed by mix-down to an 8-kHz carrier and amplitude detection. The amplitude-detected signal was sampled at 1000 samples per second (0.74-m spatial sampling resolution) using a PC-based analog to digital converter with 16-bit resolution. A pulse length of 2 ms was used, yielding an acoustic resolution of 1.45 m, transmitted at a rate of 1 ping every 2 s. The data-acquisition PC processed the digital data in real time, generating images of backscattered intensity versus range and time. Images composed of 800 pings covering 26.7 min were then telemetered to the base station. The

raw data was normally purged a few h after processing. However, during March and April of 1997 most of the raw digital data were stored to enable later detailed analysis. The images were reviewed on a weekly basis together with the environmental data. Finally, the results of the review and interpretation of the echograms were distributed to the environmental monitoring agencies.

Prior to deployment, the complete sidescan system was calibrated utilizing as reference the backscatter from Tungsten-Carbide target spheres (following Vagle *et al.*, 1996). Separate calibration runs with three different sphere radii (19.05, 20.00, and 21.50 mm) were performed. The target spheres were suspended at several ranges from 5 to 15 m along the main axis of the transducer beam. Echo intensities were averaged over 200 transmissions. By comparing the calibration results from the three spheres at multiple ranges, the accuracy in target strength measurements was estimated as  $\pm 0.5$  dB. The transducer beam pattern was provided by the manufacturer and additionally verified with *in situ* tests. Then, including the calibration results, the estimated acoustic target strength (ETS, in dB) within 0.74-m range cells centered at range,  $r$ , can be given by the well-known sonar equation (e.g., Clay and Medwin, 1977)

$$\text{ETS} = K + 20 \cdot \log_{10}[A(r)] - \text{TVG}(r) + 40 \cdot \log_{10}[r] + 2 \cdot \alpha \cdot r, \quad (1)$$

where  $K$  is the calibration coefficient (which includes transmit power, transducer sensitivity, waveform detection, and A/D conversion factors),  $A(r)$  is the echo amplitude in digital counts (sampled at 1 kHz and using a 2-ms-long pulse),  $\text{TVG}(r)$  is the receiver time-varying gain in dB, and  $\alpha$  is the acoustic absorption in seawater ( $0.02 \text{ dB} \cdot \text{m}^{-1}$  at 100 kHz under these water conditions).

Another important characteristic of the sonar system is the limitations imposed by systemic noise. As will be shown later, this noise imposes some limits on detection of schools at longer range. There are two types of noise voltages: (i) transducer preamplifier and cable noise (i.e., pre-TVG), with amplitude  $V_0$ , and (ii) digitization noise with amplitude  $V_1$ . An overall noise threshold can then be taken as the incoherent sum of the time-averaged noise root-mean-square voltages, i.e., as

$$V_N^2(r) = \langle V_0^2 \rangle \cdot 10^{[\text{TVG}(r)/10]} + \langle V_1^2 \rangle, \quad (2)$$

where  $\langle \rangle$  denotes time-averaging. This can be easily converted to digital counts,  $A(r)$ , and substituted back into Eq. (1) to calculate a noise threshold in ETS space. From *in situ* measurements, the values of

$$\sqrt{\langle V_0^2 \rangle}$$

and

$$\sqrt{\langle V_1^2 \rangle}$$

were determined to be 0.85 and 10 mV, respectively.

A series of thermistor and conductivity (T/C) sensors and an acoustic Doppler current profiler (ADCP) were installed near the lighthouse. The data from this measuring system were made available through the general environ-

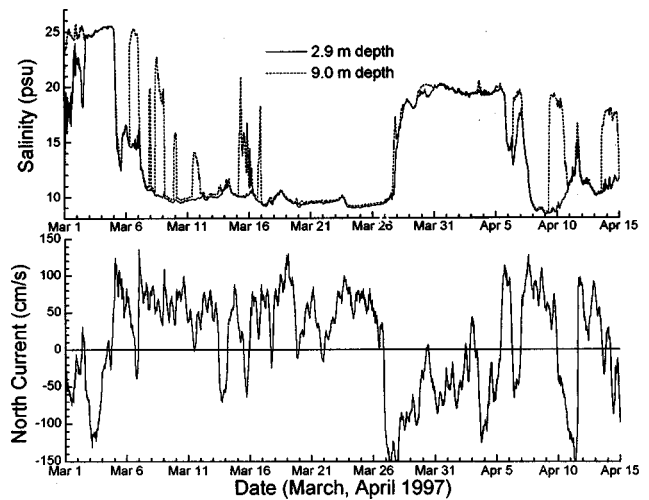


FIG. 2. Surface and bottom water salinity and along-channel current versus time from Drogden channel 1 March to 15 April 1997.

mental monitoring program. The position of the T/C chain and the ADCP are shown in Fig. 1(b), and the T/C mooring is visible in the sonar data as a discrete target near 25 m range. On the T/C mooring, temperature and conductivity were measured at five depths: 2.9, 4.4, 5.6, 7.4, and 9.0 m, respectively. On the ADCP mooring, there was an additional T/C measuring point at 10.6 m depth. These sensors were sampled continuously at 30-min intervals.

The Drogden channel waters were characterized by either a northward flow of relatively fresh Baltic Sea water or a southward flow of more saline water from the northern Øresund and the Kattegat. These flow regimes were not dominantly tidal, but driven by wind forcing and seasonal fluctuations. Figure 2 presents surface and bottom salinity and along-channel (north–south) current data for March and early April, 1997. Although the temperature of both water types was similar (between  $2.5^\circ\text{C}$  and  $4.0^\circ\text{C}$  during March), the typical salinity of the Kattegat water was 25 psu as compared to the outflowing Baltic Sea water near 10 psu. Under normal flow conditions, the waters were relatively homogeneous in both temperature and salinity. However, during transition periods between the flow regimes or during very weak Baltic outflows, a stratified two-phase flow regime sometimes occurred. These events are clearly evident in Fig. 2 as large differences between surface and bottom salinity, for example on 1, 6, and 8 March and 9 and 14 April. The acoustic implications of these changes in flow regimes will be discussed in the next two sections.

## II. EXAMPLE DATA

The most straightforward (although time-consuming) method for processing the many months of continuous sonar data is visual inspection. Human pattern-recognition skills are superior to any computer-processing system in existence today. Interpretation of the echograms is basically an experience-based process, whereby prototypical patterns and signatures are recognized against the nominal background reverberation. What follows are examples and descriptions of common target patterns.

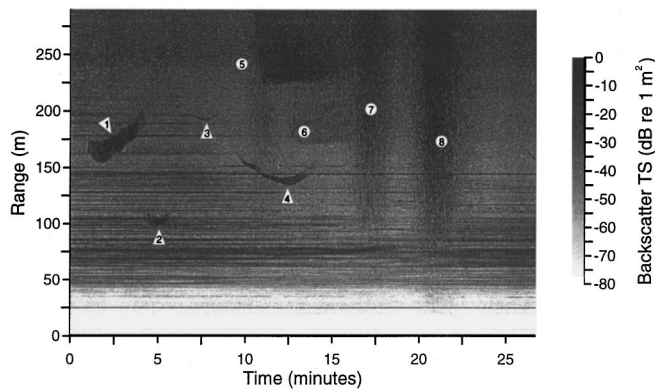


FIG. 3. Estimated target strength versus range and time starting 0853UT, 5 April 1997, showing fish schools (triangles 1–4) and vessel wakes (circles 5–8).

Figure 3 shows a typical echogram that contains many of the important observational features. The raw echo intensities were converted to target strength (TS) values using Eq. (1). Several fish schools and the acoustic signatures from four ships are shown, all of which are recognized by their transient behavior relative to the nominally time-invariant seabed backscattering. The constant-range bands from 40 to 220 m are due to backscatter from the seabed sediments, and these were moderately time-invariant under typical flow conditions. The disappearance of seabed reverberation near 220 m range marks the edge of a roughly 100-m-wide by 3-m-deep gully in the channel bottom. The thin, high-intensity line at 25 m is the echo from the T/C mooring and its sub-surface float.

Figure 3 shows several fish schools fish traversing the acoustic beam, with the most prominent near 160 m range. Gill net surveys conducted in this region during March and April 1997 (and in previous years, Stæhr, 1997) confirmed that schools of this size migrating at this time of year were herring. Note that these herring schools were clearly detectable above the background reverberation from the seabed, with echo intensity up to 20 dB above the background. The largest school is roughly 20 m wide with duration in the beam of roughly 180 s. If this school is assumed to be approximately circular in plan form, the duration-in-beam at this range suggests an advection speed past the sonar of  $1.1 \text{ m} \cdot \text{s}^{-1}$ . At this time the current was  $< 10 \text{ cm} \cdot \text{s}^{-1}$  in a transition from southward to northward flows; thus, the fish were swimming rapidly. Assuming a typical herring-school density of 5 fish per  $\text{m}^3$  (Misund, 1993), a cylindrical school of diameter 20 m and height 8 m (two-thirds the water depth) would be composed of roughly 12 600 individuals. Note that with the horizontally wide beam it is impossible to extract the true school shape, but this beam geometry has certain advantages for echo integration (discussed below). Similar-sized herring schools were observed from 50 to 500 m range, generally during daylight hours, throughout March and April.

Major interference from ship wakes was regularly present, as Drogden channel is a busy traffic lane, averaging approximately 1 ship every 20 min both day and night. The traffic management scheme placed southbound vessels

nearer the sonar (250 to 400 m range) while northbound vessels traveled the far side of the channel. These ship tracks had several common interference characteristics. First, for larger, deep-draft vessels there was a direct echo from the hull and a resultant shadowing of any fish or bathymetric targets behind them. As a result of the ship traversing the horizontally wide beam nearly perpendicular to the main axis, these direct reflections appeared as hyperbolic trajectories on the echograms. At a given range, the time for the ship to cross the sonar beam is inversely proportional to its speed. Also commonly observed as the ship's propeller crossed the beam axis was an intense noise event assumed to be caused by cavitation. Following the passage of the ship, intense, constant-range backscattering regions were observed due to the injection of small air bubbles within the vessel wakes. These wake bubbles were presumably created by breaking of the ship's bow, quarter, and stern waves. These wakes dissipated as the air bubbles either rose to the surface or dissolved, a process which typically required 6 to 10 min (see, Trevorrow *et al.*, 1994). Finally, newly created wakes shadowed acoustic backscatter from behind them due to the strong acoustic absorption within the wake bubble clouds. All these features are seen in Figure 3 for the ship trajectory and wake centered at 11 min, 225 m. Smaller, slower vessels with minimal wakes (yachts) were commonly observed on weekends.

For individual herring, the target strength (TS, defined as 10 times the common logarithm of the backscatter cross section) at dorsal incidence is a well-known function of length, i.e., using an empirical relation as  $\text{TS}(\text{dB}) = 20 \log(L) - 71.9$  (Foote, 1987) for 38-kHz echo-sounders. Assuming a geometrical scattering regime (implicit in this empirical relation), the herring TS at 100 kHz is the same. This frequency-independent TS relation is consistent with results found by Degnbol *et al.* (1985) for herring and sprat in the Kattegat and Baltic. Results from gill net surveys conducted in the Øresund during March and April, 1997 found a mean herring length of 25.9 cm (Stæhr, 1997), which corresponds to a  $\text{TS} = -43.6 \text{ dB}$  at dorsal incidence. In this situation, the herring can be assumed swimming along the channel perpendicular to the sonar beam; thus, they were insonified at lateral incidence. Unfortunately, the acoustic target strength for herring at other incidence directions has not been reported. However, empirical results from a large number of fish species (including clupeoids) compiled by R. Love (1977) revealed that the lateral incidence TS was 1.5 dB higher than the dorsal incidence value. Thus, the expected mean TS for these Øresund herring at lateral incidence was  $-42.1 \text{ dB}$ .

The ETS of the herring schools shown in Fig. 3 was significantly higher than the seabed reverberation levels, as shown in Fig. 4 by a comparison between a short-duration profile through herring school #1 (from Fig. 3) and the background reverberation. Over most of the insonified range, the short duration profile is nearly identical to the background reverberation. However, in the range from 153 to 173 m, the herring school appears as an intensity anomaly up to 22 dB above the background reverberation. The peak ETS was near  $-13.8 \text{ dB}$ . This herring-school profile was taken at the time



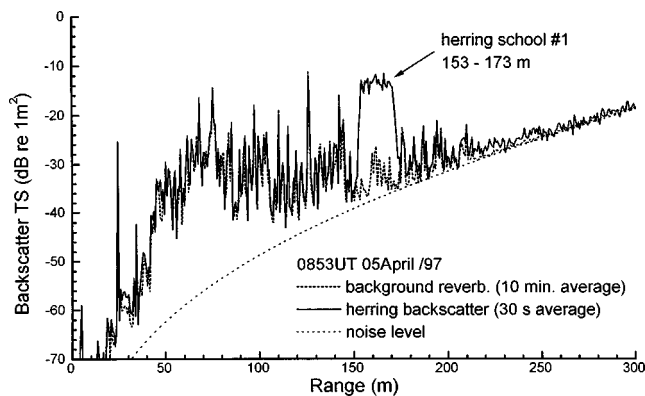


FIG. 4. Profile of estimated target strength versus range through the herring school #1 in Fig. 3, compared with the background reverberation and noise level. Profile through the herring school is averaged over 15 transmissions (30 s) while background reverberation is averaged over 300 transmissions (10 min). Noise level calculated from Eqs. (1) and (2).

of closest approach, so that the school lies near the main beam axis. Given the beam dimensions at this range (approximately 180 m wide by 8 m high) it can be reasonably assumed that the entire herring school lies near the main axis of the sonar beam, and certainly within the  $-3$ -dB points. Note that this herring school would be less visible if located at either closer range (e.g., near 80 m) where the seabed reverberation was stronger, or at greater range ( $>200$  m) where the background systemic noise level was higher. The background reverberation curve is important as it defines a range-dependent threshold for echo-integration analysis. Specifically, an integration region for the herring school can be defined where the ETS is more than 3.0 dB above the background reverberation (i.e., in this example, from 153 to 173 m range). Thus, the total number of herring in the school,  $N$ , can be calculated using

$$N = \int_{\text{school}} 10^{0.1 \cdot (\text{ETS} - \text{TS})} \cdot dr, \quad (3)$$

where  $dr$  is the range increment (in this case 74 cm), and the integral is evaluated only when the TS threshold is exceeded. Note that this calculation ignores acoustic absorption by the fish within the school and assumes that the entire school is located within the horizontally wide beam at the time of the profile. Thus, in this example the integrated number of herring contained in this averaged profile is 13 600, which is in agreement with the earlier rough estimate based on assuming the school to be cylindrical with a packing density of 5 per  $\text{m}^3$ .

At close range ( $<150$  m), individual trajectories attributable to non-schooling fish were observed, as shown in Fig. 5. Swarms of such traces were generally seen in conjunction with saline intrusion events (as in this present example) when the seabed reverberation was minimal or absent due to upward refracting conditions. Similarly to the ship trajectories, these fish traces exhibited a distinctive hyperbolic shape on this range vs time display, i.e., as

$$r^2 = r_0^2 + u^2(t - t_0)^2,$$

where  $r$  is the range from sonar,  $u$  is the swim speed perpendicular to the sonar axis, and  $t$  is time, with  $r_0$  and  $t_0$  the

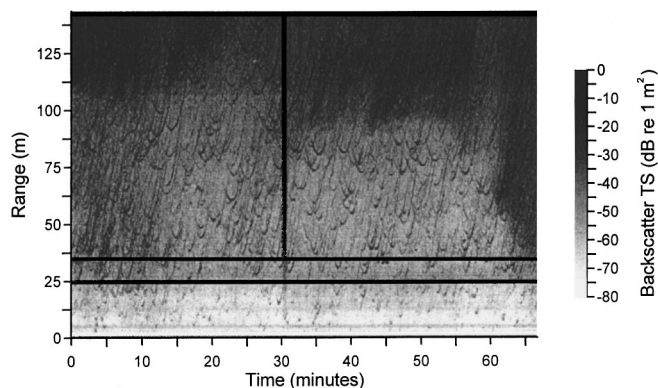


FIG. 5. Estimated target strength versus range and time starting 1457UT, 15 March 1997 showing hyperbolic trajectories due to individual fish drifting through beam.

point of closest approach. By fitting this functional form to the observed trajectories, the along-channel swimming speeds relative to the fixed sonar installation could be estimated. This was accomplished through computer-aided visual identification (using a computer mouse) of the target trajectories on range versus time echograms. Note that this method alone can not distinguish the *direction* of fish motion. During the time period of Fig. 5, there were 101 clearly identifiable trajectories observed from 20 to 100 m range, with a mean advection speed of  $0.49 \pm 0.04 \text{ m} \cdot \text{s}^{-1}$ . At this time the current was  $50 \text{ cm} \cdot \text{s}^{-1}$  southward, so the fish were most probably drifting southward with the current. This conclusion was strengthened by a close agreement of the fish advection speed and ADCP current magnitude over a 48-h period on 15 and 16 March (see Fig. 6). Note that contrary to expectations, the individual fish targets were observed during both daylight and night periods and during flows in both directions along the channel. Furthermore, the echo ETS at the point of closest approach (generally the maximum) cor-

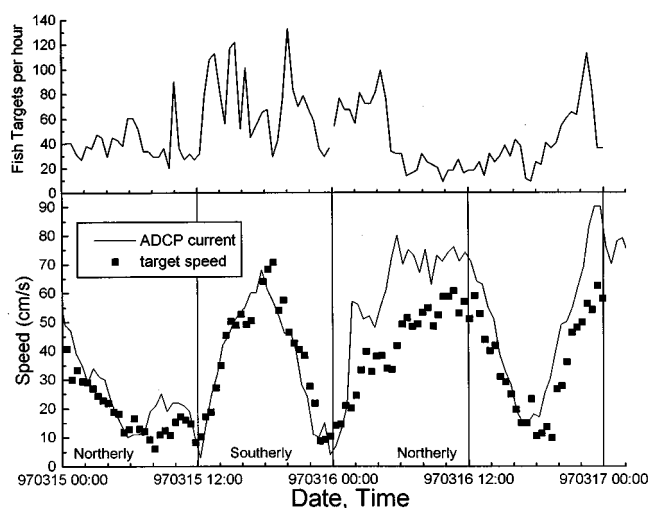


FIG. 6. Comparison of along-channel current magnitude (from ADCP at 6 m depth) with fish target advection speeds over a 48-h interval starting 15 March 1997. Fish speeds estimated from hyperbolic trajectories averaged over 26.7-min blocks and ranges from 20 to 100 m. Top panel shows number of fish trajectories from 20 to 100 m range, normalized to hourly intervals. Current direction labeled along bottom axis.

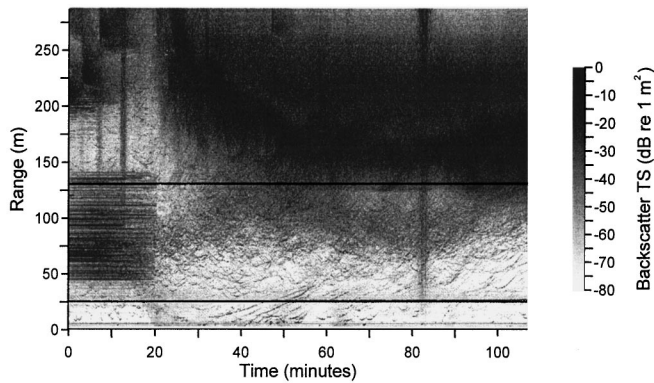


FIG. 7. Estimated target strength versus range and time starting 0403UT, 15 March 1997, showing transition from normal to stratified flow regimes.

responded to the fish located horizontally near the main axis of the sonar beam, where the horizontal beam deviation loss was negligible. For over 600 targets examined on 15 March, the closest-approach ETS was approximately log-normal distributed with mean of  $-43.2$  dB and standard deviation of  $5.6$  dB. Note the potential for some of these targets to be located on the fringes of the vertically narrow sonar beam, both above and beneath, especially at close range. Thus, this closest-approach ETS was an underestimate of the true fish TS, which can only be extracted through the use of some kind of beam-pattern compensation technique. However, the magnitude of this uncompensated ETS strongly suggests that these targets were individual herring (or herring-sized fish) in dispersed mode.

Abrupt changes in acoustic propagation conditions were observed during occasional stratified flow conditions in March and April, 1997 (see Fig. 2). The normal, relatively unstratified conditions were characterized by slightly downward-refracting sound-speed profiles induced by weak temperature gradients. Thus, under normal conditions the sonar insonified a confined region near the bottom of the channel, such that the acoustic background was dominated by seabed backscattering. However, the transition flows were characterized by strongly upward-refracting sound-speed profiles, with sound-speed differences sometimes exceeding  $10 \text{ m} \cdot \text{s}^{-1}$  between bottom and surface. This allowed strong surface backscattering regions to appear at 150 to 300 m range, and created acoustic shadow zones where fish detection was not possible. The transition between unstratified and stratified flow conditions occurred rapidly, typically requiring less than 15 min.

Figure 7 shows a typical echogram during one of these transition events. Before the transition occurred, the acoustic reverberation from 60 to 240 m range was dominated by seabed backscattering. This bottom scattering was nominally time invariant. After the transition to upward-refracting, the dominant backscattering was now from the water surface, with the first convergence region roughly 40 m wide but varying temporally at ranges from 150 to 300 m. Note further examples of the individual fish trajectories discussed above. Figure 8 compares averaged background reverberation curves for the two cases. At ranges up to 120 m, the saline intrusion case gave reduced background scattering lev-

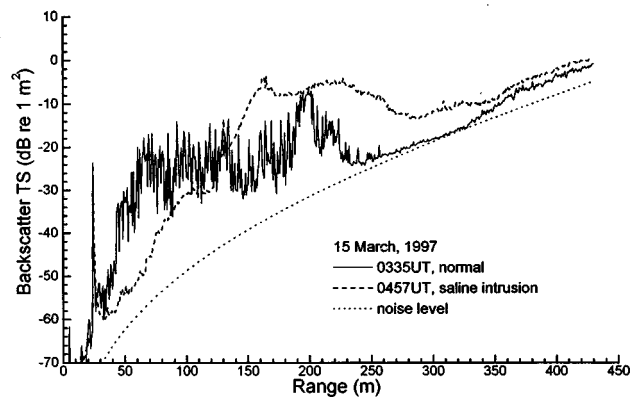


FIG. 8. Comparison of time-averaged estimated target strength versus range for normal and stratified flow regimes. Data taken starting 0325UT and 0525UT on 15 March 1997. Each curve averaged over 800 transmissions (26.7 min). Noise level calculated from Eqs. (1) and (2).

els because of the lack of any seabed interaction. Beyond 150 m range the sea-surface backscattering in the saline intrusion case became dominant. By comparison with the fish school ETS signature in Fig. 4, it is clear that the level of this sea-surface backscatter beyond 150 m range was sufficient to obscure the echoes from any fish schools present.

### III. DISCUSSION OF PROPAGATION AND REVERBERATION

At this location, the potential for surface and seabed reflections and acoustic refraction due to water stratification adds important acoustic propagation effects that must be included in any quantitative estimates of herring biomass. The bottom and surface boundaries create a kind of *reflective acoustic waveguide* which has the potential for increasing the backscattered amplitude and increasing the *effective* pulse length of the echo relative to an unbounded medium. In this particular situation, the use of a relatively long transmitted pulse and the fact that herring schools are composed of multiple closely spaced targets implies that the direct and reflected multipaths from each fish target are generally not separable. This complicates the use of integrated echo-intensity information for quantitatively estimating numbers of herring in each school. Furthermore, as has been shown above, certain flow conditions create acoustic refraction effects that significantly limit fish detectability at longer ranges.

Owing to the sharp contrast in acoustic impedance between air and water, the specular acoustic reflections from the sea surface will suffer only small losses at all angles of incidence. Similarly, for the relatively hard, limestone sediments found in Drogden channel, low-grazing angle ( $<10$  deg) reflections will have similarly low losses. This follows from the fact that for sediments with a bulk sound speed greater than seawater, as is the case here, there exists a critical grazing angle for total reflection, given by  $\arccos[c_w \cdot c_s^{-1}]$ , where  $c_w$  is the sound speed in water (nominally  $1435 \text{ m} \cdot \text{s}^{-1}$ ), and  $c_s$  is the sound speed in the sediments. At this location, the nominal sediment sound speed lies near  $1800 \text{ m} \cdot \text{s}^{-1}$ , yielding total reflection for grazing angles  $<37$

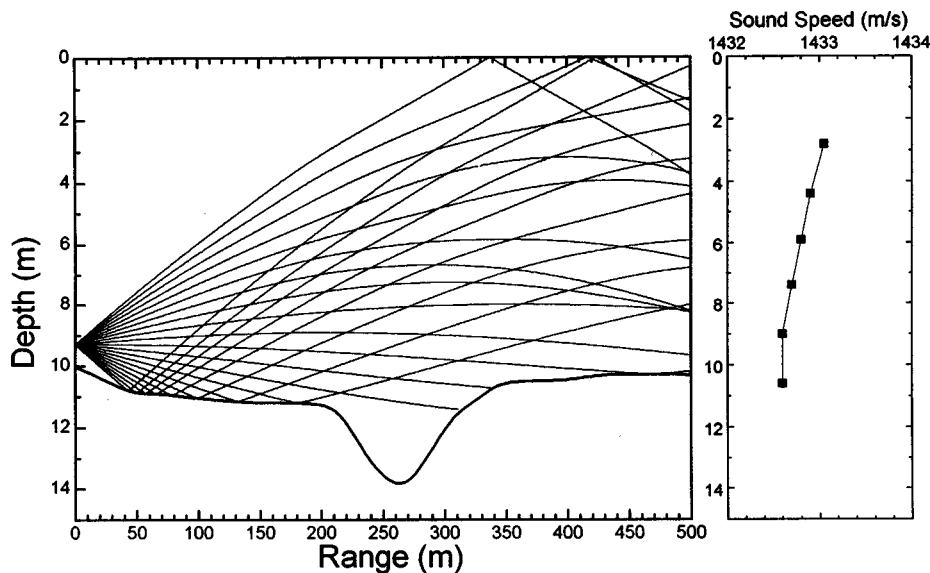


FIG. 9. Ray-tracing simulation with sound-speed profile for normal (unstratified) flow regime. Paths for 21 rays covering  $\pm 3.0$  deg from horizontal are shown.

deg. Clearly, there will be small bottom reflection losses ( $< 3$  dB) due to sediment roughness effects not included in the simple two-layer result used above.

For 100-kHz sonars, it is appropriate to model acoustic propagation using ray-tracing analysis. In the simplest form, ray tracing can graphically display the propagation and boundary interactions for a *fan* of beams launched from the transducer. More quantitatively, given the source and target locations, bathymetry versus range along the beam axis, and a sound-speed profile, a ray-tracing program can calculate the ray trajectory, times of arrival, launch and arrival angles, and amplitudes for all the possible *eigenrays* connecting the source and target. Note that for the case of targets lying outside the beam and/or with a nonuniform sound-speed profile, it is possible to have no valid eigenrays. Although more sophisticated models exist, for this first attempt at applying acoustic-propagation modeling to fishery problems we chose a simplified model. Thus, we assumed no range dependence other than bathymetry, zero surface reflection loss, and bottom reflection losses given by simple two-layer reflection theory, with a 2-dB loss at grazing angles below critical.

Figure 9 shows a ray-tracing calculation for the normal, moderately homogeneous flow regime, which can be directly compared to the background reverberation curves shown in Figs. 4 and 8. In this situation, the vertically narrow acoustic beam is confined near the seabed by mild downward-refracting conditions. This allows low grazing-angle backscatter from seabed sediments from approximately 50 to 200 m. The seabed backscatter then abruptly stops at the edge of a deeper subchannel roughly 100 m wide by 3 m deep. The seabed backscattering then mildly increases again after 300 m as the rays begin to graze the seabed again on the far side of the sub-channel. However, the density of rays hitting the seabed beyond 300 m range is significantly less than for the 50 to 200 m region; thus, the levels of seabed reverberation should be less. All these reverberation features qualitatively found in the ray trace can be observed in the actual reverberation curve shown in Fig. 8. It is clear that fish targets swimming between the surface and 2 to 6 m depth, especially nearer to the transducer location, will not be insonified

and thus not detected. Also, fish within the deeper subchannel will similarly be undetectable. However, this ray-tracing analysis shows that a substantial proportion of the channel cross section is insonified at ranges up to 500 m and beyond.

Figure 10 shows another ray-tracing simulation for the stratified, upward-refracting flow regime. Again, this ray tracing can be directly compared with the observed background reverberation curve for the saline intrusion flow conditions displayed in Fig. 8. In this case, the vertically narrow beam is refracted upwards, intersecting the surface from 130 to 230 m range. This surface intersection creates relatively strong backscatter reverberation from surface roughness and potentially from near-surface plumes or layers of air bubbles created by breaking waves (if present). Beyond 200-m range, the acoustic beam is then reflected off the surface and down into the stratified flow region again, with a second surface convergence region produced beyond 500-m range. It is clear that under these flow conditions, extensive areas of the channel cross section are not covered, especially near the bottom where the fish are expected to migrate.

The quantitative effects of this shallow-water reflective waveguide can be extracted from the ray-tracing results by generating synthetic target echoes. This was accomplished by propagating 2 ms by 100-kHz pulse replicas out and back along all possible combinations of the eigenrays. The pulse amplitudes along each eigenray, which include the effects of refraction and bottom loss, were further corrected for transmit and receive transducer beam-pattern effects, with added acoustic absorption losses proportional to the path length. The 100-kHz sidescan transducer-beam pattern closely approximated a rectangular source (see, e.g., Clay and Medwin, 1977 for evaluation), such that rays launched or arriving at angles outside of  $\pm 3$  deg were significantly attenuated and thus ignored. The multiple subechoes from a single target were then *coherently* added together, taking into account their differing arrival times, to generate a single-target waveform. To extend this simulation to a herring school, the single-target waveforms for multiple targets in a closely spaced cluster were *coherently* added together. Specifically, a coherent-cluster waveform was synthesized for 24 targets

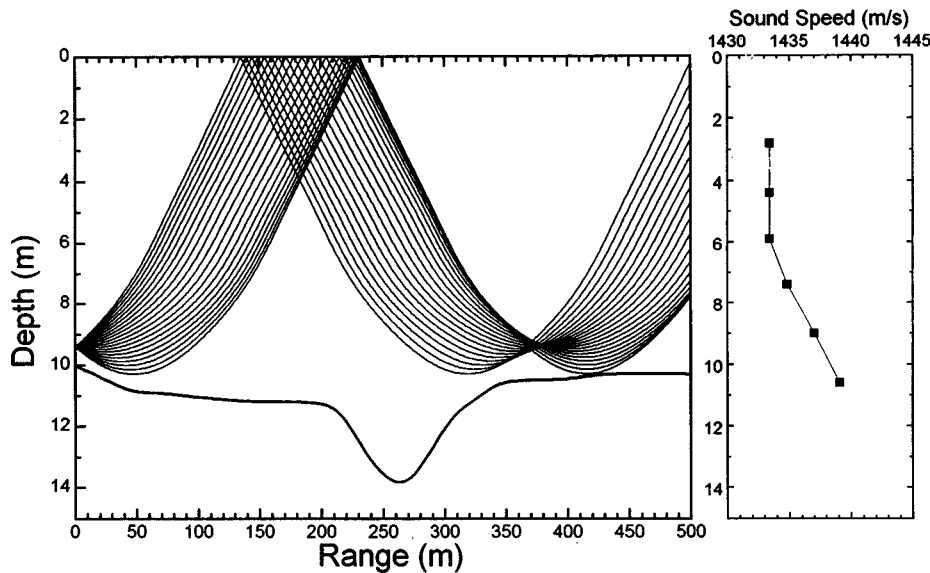


FIG. 10. Ray-tracing simulation with sound-speed profile for stratified flow regime. Paths for 21 rays covering  $\pm 3.0$  deg from horizontal are shown.

spaced 20 cm vertically from 6.2 to 10.8 m depth near 160 m range under the normal flow conditions (i.e., simulating the herring school examined in Fig. 4). Note that, for the saline intrusion case, there were no valid eigenrays for targets at this location. In this example, there were two eigenrays connecting the source and the target cluster, a direct path and a seabed-reflected path. Thus, there were four possible two-way combinations between the source and each target. The effect of adding this seabed-reflected eigenray was to increase the maximum echo amplitude by roughly 50% and to increase the overall echo length by roughly 50  $\mu$ s. Note that in estimating the averaged echo from a real herring school, the echoes from individual fish can be assumed to add *incoherently*. Thus, to correctly model incoherent echo summation, the root-mean-square signals from 200 separate coherent-cluster simulations were averaged together. For each cluster simulation, the spatial phase of each of the 24 targets was randomized by  $2\pi$ . Mimicking the sidescan data acquisition process, the root-mean-square sampling interval was 1 ms.

The net quantitative effect of this multipath enhancement can be found through echo integration [as in Eq. (3)] of the resultant root-mean-square synthetic waveforms. Specifically, the ratio of integrated number of targets estimated for the case of both direct and bottom-reflected paths relative to the direct path only was calculated over the range interval 153 to 173 m. The averaged multipath enhancement ratio was then  $1.4 \pm 0.1$  (mean  $\pm$  standard deviation). As a check, the resulting averaged target density using only the direct path simulations was  $24. \pm 2.6$  targets. Thus, these simulations suggest that for a herring school located near 160 m range under these flow conditions, the total number of herring in the school will be overestimated by 43%. For the herring school discussed in Fig. 4, this reduces the apparent total number of fish from 13 600 to 9500.

#### IV. DISCUSSIONS AND CONCLUSIONS

Review of the many months of 100-kHz sidescan data from Drogden channel uncovered many clear examples of

herring schools, mostly observed during the fall migration season in September and October, and again in the spring from February to April. These herring schools were recognized by a combination of their greater echo strength and transient behavior relative to the nominally invariant background seabed reverberation. Quantitative analysis demonstrated a 15 to 20-dB (typical) detection signal to reverberation ratio under normal flow conditions. Herring schools were observed at horizontal ranges up to 500 m. The long-range limit was imposed by increasing levels of systemic noise, which can presumably be improved through the use of more sophisticated sonar electronics. Additionally, interference from vessel traffic in the channel was common.

Perhaps the major advantage of this semiautomated system was the potential for sustained, 24-hour-a-day surveillance covering a significant portion of this herring migration channel. Due to a combination of equipment breakdowns and power supply problems at the lighthouse, the sonar system was only operational 60% of the time between June 1996 and February 1997. However, between 21 February and 30 May 1997 the sonar system was fully operational, covering the spring herring migration and a variety of flow conditions in the channel. Another useful property of this horizontally wide sonar beam was that, under typical flow conditions and at ranges greater than roughly 100 m, an entire herring school could be localized near the main axis of the beam. This allowed the direct use of echo-integration techniques to estimate total school biomass, without the need for beam-pattern compensation. A down side of this horizontally wide sonar beam was that vertical and along-channel herring-school dimensions could not be determined. This could be partially alleviated by the simultaneous use of a second narrow-beam sonar (e.g., a 3 deg conical beam).

Abrupt changes in acoustic propagation conditions were observed during transitions between northward and southward flow regimes in Drogden channel. These transitions occurred on the order of a few times per month, being forced by meteorological conditions rather than tidal action. These transition events were characterized by intrusions of more

saline bottom water, creating a strongly upward-refracting sound-speed profile. This changed the reverberation regime from seabed to surface backscattering. During these upward-refracting conditions, fish detection was not possible beyond the starting range of the first surface convergence, or roughly 150 m. These transitions between normal and stratified flow were quite rapid, taking place over periods of 10 to 30 min. The impact on the acoustic performance of the system created by these environmental changes highlights the *necessity* for simultaneous monitoring of water property and current profiles. This is in contrast to vertical echo-sounding surveys, where the water properties need only be approximately known and refraction effects are negligible.

On occasion, individual fish tracks attributable to adult herring were observed from 20 to 150 m range, generally during flow transition periods when the seabed reverberation was minimal or absent. These fish trajectories exhibited a hyperbolic shape on range vs time echograms, such that their perpendicular (i.e., along-channel) advection speed could be extracted from curvature of the trajectory. The mean target advection speed was found to closely follow the measured along-channel current speed, strongly suggesting that the fish were drifting with the current. Additionally, the mean ETS (at point of closest approach but uncorrected for vertical beam-deviation loss) of these fish targets was  $-43.2$  dB, suggesting that these targets were adult herring in dispersed mode, which had an expected lateral incidence TS of  $-42.1$  dB. Clearly, the true TS of these targets would be better measured through the use of a more sophisticated dual or split-beam sonar. Additionally, a split-beam sonar could provide directly the target advection speed and *direction*. One possibility for further work is to adapt beam-pattern deconvolution methods for this case of a horizontally wide, boundary-grazing beam, including the effects of refraction and boundary reflection. More importantly, this observation of individual fish tracks confirmed the existence of some proportion of herring (or herring-like fish) in *dispersed mode* within Drogden channel. It had been assumed *a priori* that herring would most likely exhibit schooling behavior within this shallow channel. This complicates attempts to estimate total migrating herring biomass, as this horizontal sonar system did not (except under special circumstances and to limited range) have sufficient signal-to-noise capability to resolve individual fish.

Through the use of ray-tracing analysis, it was possible to model the effects of boundary reflection focusing and acoustic refraction. Ray tracing is perhaps the least sophisticated form of acoustic-propagation modeling, yet it provides an intuitively satisfying picture of the acoustic geometry. In particular, the ray-tracing plots were useful for identifying regions where significant boundary backscattering could occur. Also, areas could be identified where fish targets were not detectable due to shadowing and refraction effects. More quantitatively, the ray-tracing results could be used to synthesize backscattered echoes from fish targets, including the effects of boundary reflected multipaths. For an example cluster of 24 targets between 6 and 11 m depth at 160 m range, the boundary reflections were shown to increase the maximum echo amplitude and effective pulse length. Quan-

titatively, this resulted in an overestimation of the number of targets by a factor of 1.4 when using echo-integration analysis. Acoustic modeling of this kind thus provided a means to compensate for these boundary reflection biases, although this technique needs further study and *in situ* verification, for example through a combined sonar and net trawl survey. More generally, we believe that by performing this analysis under a variety of flow conditions for target clusters covering the entire channel cross section, contours of expected estimation bias vs flow conditions could be generated. Similarly by comparing the synthesized echo strengths from fish clusters with modeled, or measured, boundary reverberation levels, contours of detection signal-to-noise ratio could be generated. These two types of information (estimation bias and target detectability limits) could then be incorporated within the acoustic monitoring system to improve the real-time quantitative reporting of herring school abundance.

A critical requirement for estimating school numbers is a knowledge of the acoustic target strength of the individual fish, including the effects of horizontal orientation. This could be measured directly through acoustic surveys on the herring while they are in dispersed mode (as done in Degnbol *et al.*, 1985). Alternately, we used the mean length estimated for the population as a whole from a regional gill net survey conducted during the spring migration (Stæhr, 1997). Then we used a well-established 38-kHz, dorsal-incidence target strength vs length relationship. Either way, it is imperative that some kind of trawl or acoustic survey be performed in conjunction with this horizontal sonar operation to establish the target identity and mean size. Additionally, there is a need to measure TS for herring at other frequencies and incidence angles. This is crucial for sidescan sonar applications because the horizontal incidence TS of an individual herring (or herring-like fish) can vary by roughly 12 to 14 dB, depending on orientation (i.e., head or tail vs lateral incidence, Love, 1977). In this case, there was reason to expect fish to be swimming along channel, i.e., at lateral incidence to the sonar. However, this assumption has the potential for underestimating the school numbers if the herring are all swimming at some *nonlateral* orientation relative to the sonar beam. The extreme case is tail incidence, with a corresponding target strength 14.7 dB lower than the lateral incidence value (Love, 1977). At this orientation, the total school numbers estimated in discussions surrounding Fig. 4 would change from 9500 at lateral incidence to 280 000 at tail incidence. Clearly, it is important to be on guard for situations where the herring might be swimming in some nonlateral direction, and adjust the mean fish TS accordingly.

## ACKNOWLEDGMENTS

The Danish Environmental Protection Agency, Kontroll- och Styrgruppen för Öresundsforbindelsen (KSO Sweden) and Öresundskonsortiet A/S are acknowledged for permission to publish the present work. Additional support for M.V.T. was provided by the National HydroAcoustics (New Technologies) program of the Canadian Dept. of Fisheries and Oceans. The authors are indebted to Dr. Niels-Erik Ot-

tesen Hansen and Dr. David M. Farmer, of LIC Engineering and the Institute of Ocean Sciences respectively, for their leadership in this collaborative effort.

- Clay, C., and Medwin, H. (1997). *Acoustical Oceanography: Principles and Applications* (Wiley-Interscience, New York).
- Degnbol, P., Lassen, H., and Stæhr, K.-J. (1985). "In situ determination of target strength of herring and sprat at 38 and 120 kHz," *Dana* **5**, 45–54.
- Foote, K. (1987). "Fish target strengths for use in echo integrator surveys," *J. Acoust. Soc. Am.* **82**, 981–987.
- Gaudet, D. (1990). "Enumeration of migrating salmon populations using fixed-location sonar counters," *Rapp. P.-V. Reun.-Cons. Int. Explor. Mer.* **189**, 197–209.
- Hewitt, R., Smith, P., and Brown, J. (1976). "Development and use of sonar mapping for pelagic stock assessment in the California Current area," *Fish. Bull.* **74**, 281–300.
- Kubecka, J., Duncan, A., Duncan, W., Sinclair, D., and Butterworth, A. (1994). "Brown trout populations of three Scottish lochs estimated by horizontal sonar and multimesh gill nets," *Fish. Res.* **20**, 29–48.
- Love, R. (1977). "Target strength of an individual fish at any aspect," *J. Acoust. Soc. Am.* **62**, 1397–1403.
- Misund, O. (1993). "Abundance estimation of fish schools based on a relationship between school area and school biomass," *Aquat. Living Resour.* **6**, 235–241.
- Misund, O., Aglen, A., and Fronaes, E. (1995). "Mapping the shape, size, and density of fish schools by echo integration and a high-resolution sonar," *ICES J. Mar. Sci.* **52**, 11–20.
- Nielsen, J. R. (1996). "Acoustic monitoring of herring," Technical Report, Danish Institute for Fisheries Research, Dept. of Fish Biology, North Sea Center, DK.
- Revie, J., Weston, D., Harden-Jones, F., and Fox, G. (1990). "Identification of fish echoes located at 65 km range by shore-based sonar," *J. Cons. Int. Explor. Mer.* **46**, 313–324.
- Rusby, J., Somers, M., Revie, J., McCartney, B., and Stubbs, A. (1973). "An experimental survey of a herring fishery by long-range sonar," *Mar. Biol.* **22**, 271–292.
- Stæhr, K.-J. (1997). "Short summary of environmental impact monitoring on herring in the sound, Autumn 1995 to Spring 1997," Technical Report, Danish Institute for Fisheries Research, Copenhagen, DK.
- Trevorrow, M., Vagle, S., and Farmer, D. (1994). "Acoustic measurements of microbubbles within ship wakes," *J. Acoust. Soc. Am.* **95**, 1922–1930.
- Trevorrow, M. (1997). "Detection of migrating salmon in the Fraser River using 100 kHz sidescan sonars," *Can. J. Fish. Aquat. Sci.* **54**, 1619–1629.
- Trevorrow, M., and Claytor, R. (1998). "Detection of Atlantic herring (*Clupea harengus*) schools in shallow waters using high-frequency sidescan sonars," *Can. J. Fish. Aquat. Sci.* **55**, 1419–1429.
- Vagle, S., Foote, K., Trevorrow, M., and Farmer, D. (1996). "Absolute calibrations of monostatic echosounder systems for bubble counting," *IEEE J. Ocean Eng.* **21**, 298–305.
- Weston, D., and Revie, J. (1971). "Fish echoes on a long-range sonar display," *J. Sound Vib.* **17**, 105–112.
- Weston, D., and Revie, J. (1989). "A 5-day long-range-sonar record of an extensive concentration of fish," *J. Acoust. Soc. Am.* **86**, 1608–1611.
- Weston, D., and Andrews, H. (1990). "Monthly estimates of fish numbers using a long-range sonar," *J. Cons. Int. Explor. Mer.* **47**, 104–111.

# Time-evolution modeling of seafloor scatter. I. Concept

E. Pouliquen, O. Bergem, and N. G. Pace

*SACLANT Undersea Research Centre, Viale S. Bartolomeo 400, I-19138 La Spezia, Italy*

(Received 12 June 1997; accepted for publication 22 February 1999)

A time-series model for acoustic seafloor backscattering is described. The method analytically expresses the elementary time-backscattered response of every seafloor surface and every seafloor volume infinitesimal element. For chosen geometric, acoustical, and acquisition parameters, they are summed to produce in the time domain a realization of the reverberation time-pressure field received at the source. The approach is based on the Kirchhoff approximation for the seafloor interface backscattering and on the Small Perturbation theory for the seafloor volume. It only accounts for single backscattering mechanisms of the compressional wave with the seafloor. The model is implemented using calculated height fields for the water-sediment interface and distributed seafloor volume inhomogeneities. The analytical description of the model and its limitations is described in this paper. © 1999 Acoustical Society of America. [S0001-4966(99)01906-2]

PACS numbers: 43.30.Gv, 43.30.Hw, 43.30.Ft, 43.30.Pc [DLB]

## INTRODUCTION

Growing interest in seafloor characteristics in shallow water is accompanied by an expectation for improved accuracy in the estimation of seafloor parameters. In order to achieve satisfactory characterization, the physics of the scattering mechanisms on and in the seafloor have to be understood. It is particularly important to evaluate the importance of each contribution from the seafloor surface and from the seafloor volume. The sole consideration of the total backscattering strength does not allow a detailed study of the scattering mechanisms. One way to achieve that is to observe the whole time series of the backscattered signal. Each time series depends on geo-acoustic properties such as local density, local sound speed, local attenuation, roughness spectrum, correlation lengths, and on the geometry of measurement such as the water height of the source from the seafloor surface, the orientation of the source, and on the source characteristics (pulse characteristics, directivity pattern). When an acoustic pulse is interacting with a seafloor, all of these parameters affect the backscattering process, resulting in a complex time-backscattered signal. Intuitively, the seafloor surface contribution should be at the beginning of the returned echo and the seafloor volume contribution should affect the tail of the echo more. But it is impossible to separate these two contributions with a time cutoff approach. Intuitively again, a hard seafloor would have a stronger seafloor surface return than a soft bottom. But this is also contingent on seafloor surface roughness or seafloor volume inhomogeneities. The geometry of measurement also has an important effect on the backscattered time signal. One should expect an extension of the time signal duration as the water height between the source and the seafloor surface increases. A simple calculation shows that the extension is not linear. Another element which strongly affects the signal is the directivity pattern and orientation of the source. For each ping, it not only conditions the insonified area but it also weights acoustic contributions from each direction. Side lobes, for example, are often the origin of significant acoustic backscattering. If one requires a close-to-reality representation of

the scattering phenomena, it is not possible to approximate the directivity pattern with a cone, or with an exponential function. Other factors are the pulse characteristics (pulse shape, spectrum, and duration) which affect the type of scattering mechanisms, e.g., geometric or Rayleigh, the expected amount of coherent and incoherent signal, and average shape of the signal.

These factors demonstrate the need to model the backscattered pressure field in the time domain for various geometric, physical, and acoustical configurations in order to: (1) understand the processes occurring during echo sounding; (2) define the limits of a possible quantitative and qualitative characterization method; and (3) articulate a possible methodology for a reliable and efficient seabed classification.

## I. MODEL APPROACH

The modeling of a backscattered time signal is a hybrid process, using a classical approach for the physical interaction of acoustic wave and seafloor on a small scale. On a larger scale, it combines a full description of the source characteristics and of the geometry of the measurement with the use of an artificially generated seafloor, the main geo-acoustic parameters of which are completely controlled. The combination of these two different approaches allows synthesis of most of the physical phenomena, which determine the time-backscattered signal. More specifically, the model combines analytical expressions of the elementary seafloor surface and elementary seafloor volume time-pressure field response, with a stochastic description of the seafloor. The expression of the time-backscattered pressure from an elementary seafloor surface is based on the Kirchhoff approximation (described in Sec. II), and the expression for the time-backscattered pressure from an elementary volume on Small Perturbation theory (described in Sec. III). Extension to larger seafloor surfaces and volumes is effected by stochastic realizations of the seafloor surface and volume features. Each stochastic sample of the bottom is realized for a given set of statistical properties of the seafloor surface and the volume: The seafloor surface is characterized principally by rms

height (roughness) and by parameters describing the horizontal features (correlation lengths for example). The volume is characterized by a rms variation and vertical and horizontal correlation lengths of the density and sound speed variations around mean values. This approach is new, in the sense that the physics of the interaction of sound waves with the seafloor is included with a description of the environment, in order to compute a time evolution of the seafloor backscattered signal.

Some constraints and limitations have been set in order to reduce the dimension of the region of validity and to allow some approximations. The model has to at least be valid in the angular region, close to normal incidence, but extension to higher angles of incidence could be envisaged in later studies. The frequency of the transmitted signal has to be high enough and the seafloor surface smooth enough to allow use of the Kirchhoff approximation.<sup>1</sup> Equally, the characteristics of the inhomogeneities should vary slightly around their mean values to allow the use of Small Perturbation theory (Secs. II and III). As the above conditions are not too restrictive, the model is valid for a wide range of acoustic instruments including echo sounders, swath (at least for the most vertical beams), and parametric systems.

## II. SEAFLOOR SURFACE SCATTERING

The main condition for the validity of the Kirchhoff approximation is the following: If the roughness of the surface is sufficiently smooth in the horizontal dimension in order to avoid shadowing effects and multiple reflection on the surface, one can assume that the sound reflection on each point of the surface is equivalent to a reflection by a tangent plane located at this considered point,  $\mathbf{R}$ . This condition is expressed by Brekhovskikh<sup>1</sup> in terms of the necessity of having a minimal interface radius of curvature  $r_c$  at a given incidence  $\theta_{\text{inc}}$  and wave number  $k_0$ :

$$2k_0 r_c \cos(\theta_{\text{inc}}) \gg 1. \quad (1)$$

Even close to normal incidence, this condition might seem too strict for some applications described later. Actually, the reason for this condition is to avoid a large amount of sharp edges or sharp points having a small radius of curvature creating edge or shadow effects. In the case of our study, we will assume the validity of the tangent plane approximation even for small wave numbers providing that the seafloor surface correlation length  $l$  is greater than the wavelength ( $l \gg \lambda$  condition from Thorsos<sup>2</sup>). This hypothesis is debatable but makes sense in the case of smooth bottoms, i.e., from mud to sand at normal incidence, and the experimental agreement displayed later on seems to justify this approach. In addition, at the computation stage, it is possible to check and reject the acoustic rays which are affected by a possible shadow effect or that are impinging on the bottom with a local angle of incidence that is beyond critical (see Fig. 1).

Assuming that these conditions are verified, the boundary conditions at point  $\mathbf{R}$  on the interface are the following (Brekhovskikh,<sup>1</sup> Clay and Medwin<sup>3</sup>):

$$p_r(\mathbf{R}) = \mathcal{R}_{01}(\mathbf{R}_0, \mathbf{n}) p_i(\mathbf{R}), \quad (2)$$

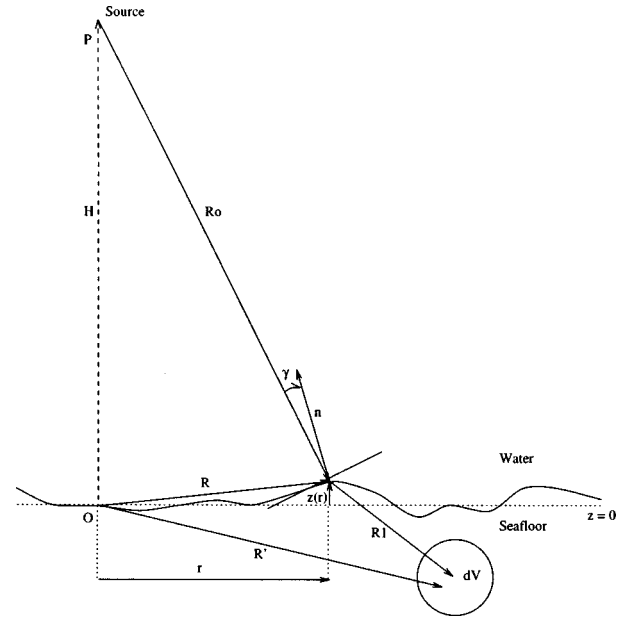


FIG. 1. Geometry of measurement.

$$\frac{\partial}{\partial n} p_r(\mathbf{R}) = -\mathcal{R}_{01}(\mathbf{R}_0, \mathbf{n}) \frac{\partial}{\partial n} p_i(\mathbf{R}). \quad (3)$$

Vector  $\mathbf{n}$  is the vector normal to the surface at  $\mathbf{R}$ . The operator  $\partial/\partial n$  is the normal derivative operator.  $\mathcal{R}_{01}(\mathbf{R}_0, \mathbf{n})$  is the local water-sediment reflection coefficient at point  $\mathbf{R}$ . We consider a source located at a height  $H$  ( $\mathbf{P} = \{0, 0, H\}$ ). This source is at first emitting a cw wave of frequency  $f$  (i. e., with an implicit  $e^{2\pi jft}$  time-harmonic dependence), at a source level  $p_0$  with a far field transmitted directivity  $D_i(\mathbf{R}_0)$  given by the direction of  $\mathbf{R}_0 = |\mathbf{R} - \mathbf{P}|$  (Fig. 1). The water column above the seabed interface is assumed to be homogeneous with a constant sound velocity profile  $c_0$ . In this case, following the same demonstration as Chotiros in Ref. 4, the incident pressure field received at point  $\mathbf{R}$  can be approximated in the far field region by the following expression:

$$p_i(\mathbf{R}) = p_0 G_i(\mathbf{R}_0) D_i(\mathbf{R}_0), \quad (4)$$

where  $G_i(\mathbf{R}_0)$  is the Green's function for the incident pressure field:

$$G_i(\mathbf{R}_0) = \frac{e^{-jk_0 R_0}}{R_0}, \quad (5)$$

with  $R_0 = |\mathbf{R}_0|$ . Using Eq. (2) and Eq. (4), we can express the reflected pressure  $p_r(\mathbf{R})$  just above the surface:

$$p_r(\mathbf{R}) = p_0 G_i(\mathbf{R}_0) D_i(\mathbf{R}_0) \mathcal{R}_{01}(\mathbf{R}_0, \mathbf{n}). \quad (6)$$

Also, the expression of the normal displacement velocity of the reflected field  $v_r$  at point  $\mathbf{R}$  can be expressed by<sup>3</sup>

$$\frac{\partial}{\partial n} p_r(\mathbf{R}) = -\rho_0 \frac{\partial}{\partial t} v_r(\mathbf{R}). \quad (7)$$

$\rho_0$  is the water mean density. Using Eq. (3) and Eq. (4), and assuming that the directivity term  $D_i(\mathbf{R}_0)$  is slowly varying, the left hand side term of Eq. (7) gives



$$\frac{\partial}{\partial n} p_r(\mathbf{R}) = -p_0 D_i(\mathbf{R}_0) \mathcal{R}_{01}(\mathbf{R}_0, \mathbf{n}) \frac{\partial}{\partial n} (G_i(\mathbf{R}_0)). \quad (8)$$

A closer look at  $\partial/\partial n(G_i(\mathbf{R}_0))$  shows that

$$\frac{\partial}{\partial n} (G_i(\mathbf{R}_0)) = \nabla[G_i(\mathbf{R}_0)] \cdot \mathbf{n} = \frac{\mathbf{R}_0 \cdot \mathbf{n}}{R_0} \frac{\partial}{\partial R_0} \left( \frac{e^{-jk_0 R_0}}{R_0} \right). \quad (9)$$

Since  $k_0 R_0 \gg 1$  in our case (interface in the far field region), we can approximate Eq. (9) by

$$\begin{aligned} \frac{\partial}{\partial n} (G_i(\mathbf{R}_0)) &= - \left( \frac{\mathbf{R}_0 \cdot \mathbf{n}}{R_0} \right) j k_0 G_i(\mathbf{R}_0) \\ &= j \cos(\gamma(\mathbf{R}_0, \mathbf{n})) k_0 G_i(\mathbf{R}_0), \end{aligned} \quad (10)$$

with

$$\left( \frac{\mathbf{R}_0 \cdot \mathbf{n}}{R_0} \right) = -\cos(\gamma(\mathbf{R}_0, \mathbf{n})). \quad (11)$$

$\gamma(\mathbf{R}_0, \mathbf{n})$  is the angle made between the incident direction  $\mathbf{R}_0$  and the normal vector  $\mathbf{n}$  to the surface at  $\mathbf{R}$ . Using Eq. (8), the first term of Eq. (7) equals

$$\frac{\partial}{\partial n} p_r(\mathbf{R}) = -j \cos(\gamma(\mathbf{R}_0, \mathbf{n})) k_0 p_r(\mathbf{R}). \quad (12)$$

The right hand side term of Eq. (7) gives

$$-\rho_0 \frac{\partial}{\partial t} v_r(\mathbf{R}) = -j \rho_0 c_0 k_0 v_r(\mathbf{R}). \quad (13)$$

Combining Eq. (7), Eq. (12), and Eq. (13), we obtain an expression of the local displacement  $v_r$  at  $\mathbf{R}$ :

$$v_r(\mathbf{R}) = \frac{\cos(\gamma(\mathbf{R}_0, \mathbf{n}))}{\rho_0 c_0} p_r(\mathbf{R}). \quad (14)$$

Morse<sup>5</sup> gives an expression of the radiated pressure at point  $P$  due to the element  $d\mathbf{S}_R$  located at  $\mathbf{R}$ :

$$dp_s(\mathbf{P}) = j \frac{\rho_0 c_0 k_0}{2\pi} v_r(\mathbf{R}) G_r(\mathbf{R}_0) D_r(\mathbf{R}_0) d\mathbf{S}_R, \quad (15)$$

where  $G_r(\mathbf{R}_0)$  is the Green's function of the reflected pressure field and  $D_r(\mathbf{R}_0)$  is the receiving directivity of the receiver collocated at the source at position  $\mathbf{P}$ . Equation (15) can be expanded using Eq. (6) and Eq. (14):

$$\begin{aligned} dp_s(\mathbf{P}) &= j \frac{k_0 \cos(\gamma(\mathbf{R}_0, \mathbf{n}))}{2\pi} p_0 \\ &\quad \times G_i(\mathbf{R}_0) G_r(\mathbf{R}_0) D_i(\mathbf{R}_0) D_r(\mathbf{R}_0) \mathcal{R}_{01}(\mathbf{R}_0, \mathbf{n}) d\mathbf{S}_R. \end{aligned} \quad (16)$$

The product  $D_i(\mathbf{R}_0) D_r(\mathbf{R}_0)$  is the gain of the emitter-receiver in the direction defined by  $\mathbf{R}$ . Rigorously,  $D_i(\mathbf{R}_0) D_r(\mathbf{R}_0)$  should be considered as frequency dependent but the directivity pattern will be assumed to be constant over the signal bandwidth. The product  $G_i(\mathbf{R}_0) G_r(\mathbf{R}_0)$  contains the propagation loss in the water column. It also contains the phase variation during the propagation. Since we are simply considering a monostatic case (source and receiver at the same location),  $G_i(\mathbf{R}_0) = G_r(\mathbf{R}_0)$  and this product can be replaced by  $G(\mathbf{R}_0)^2$ . It is also assumed that the

reflection coefficient  $\mathcal{R}_{01}(\mathbf{R}_0, \mathbf{n})$  is frequency invariant over the bandwidth of the transmitted pulse. Then, separating the frequency dependent terms of Eq. (16) and putting them into curly brackets, we obtain

$$\begin{aligned} dp_s(\mathbf{P}, f) &= j \frac{\cos(\gamma(\mathbf{R}_0, \mathbf{n}))}{c_0} p_0 \\ &\quad \times D_i(\mathbf{R}_0) D_r(\mathbf{R}_0) \mathcal{R}_{01}(\mathbf{R}_0, \mathbf{n}) \\ &\quad \times \{f G(\mathbf{R}_0, f)^2\} d\mathbf{S}_R. \end{aligned} \quad (17)$$

In the case of a non-cw signal,  $E(f)$  being the spectrum of the transmitted pulse  $e(t)$ , one can write

$$\begin{aligned} dp_s(\mathbf{P}, f) &= j \frac{\cos(\gamma(\mathbf{R}_0, \mathbf{n}))}{c_0} p_0 \\ &\quad \times D_i(\mathbf{R}_0) D_r(\mathbf{R}_0) \mathcal{R}_{01}(\mathbf{R}_0, \mathbf{n}) \\ &\quad \times \{f G(\mathbf{R}_0, f)^2 E(f)\} d\mathbf{S}_R. \end{aligned} \quad (18)$$

Until now, the contribution of an elementary surface  $d\mathbf{S}_R$  is expressed at a given frequency  $f$ . This expression is not directly generalizable to a larger scale by integrating each elementary pressure field over the whole insonified sea-floor surface ( $S$ ) before switching to the time domain. Switching into the time domain can be done at this point using the inverse Fourier transform. Analytically,  $dp_s(\mathbf{P}, t)$  is first expressed thus:

$$\begin{aligned} dp_s(\mathbf{P}, t) &= \mathcal{F}^{-1}\{dp_s(\mathbf{P}, f)\} \\ &= j \frac{\cos(\gamma(\mathbf{R}_0, \mathbf{n}))}{c_0} p_0 D_i(\mathbf{R}_0) D_r(\mathbf{R}_0) \mathcal{R}_{01}(\mathbf{R}_0, \mathbf{n}) \\ &\quad \times \mathcal{F}^{-1}\{f G(\mathbf{R}_0, f)^2 E(f)\} d\mathbf{S}_R, \end{aligned} \quad (19)$$

where  $\mathcal{F}^{-1}$  represents the inverse Fourier transform operator. Let us define  $\mathcal{K}_s(\mathbf{R}_0, t)$  as being equal to  $\mathcal{F}^{-1}\{f G(\mathbf{R}_0, f)^2 E(f)\}$  and simplify it:

$$\mathcal{K}_s(\mathbf{R}_0, t) = \int_{-\infty}^{+\infty} f G(\mathbf{R}_0, f)^2 E(f) e^{2\pi j f t} df. \quad (20)$$

Then, using Eq. (5), we find

$$\begin{aligned} \mathcal{K}_s(\mathbf{R}_0, t) &= \frac{1}{R_0^2} \int_{-\infty}^{+\infty} f E(f) e^{2\pi j f (t - 2R_0/c_0)} df \\ &= \frac{1}{R_0^2} \mathcal{F}^{-1}\{f E(f) e^{-2\pi j f (2R_0/c_0)}\}. \end{aligned} \quad (21)$$

So,

$$\mathcal{K}_s\left(\mathbf{R}_0, t + \frac{2R_0}{c_0}\right) = \frac{1}{R_0^2} \mathcal{F}^{-1}\{f E(f)\}. \quad (22)$$

Knowing that for a continuous and differentiable function  $e$  we have for the  $n$ th derivative  $e^{(n)}$

$$\frac{e^{(n)}(\mathbf{R}, t)}{(2\pi)^n} = \mathcal{F}^{-1}\{j^n f^n E(f)\}. \quad (23)$$

We obtain in this case

$$\mathcal{K}_s(\mathbf{R}_0, t) = \frac{1}{2\pi j R_0^2} e' \left( t - \frac{2R_0}{c_0} \right), \quad (24)$$

where  $e'(t)$  stands for the first time derivative of the transmitted pulse  $e(t)$ . Then, Eq. (19) becomes

$$dp_s(\mathbf{P}, t) = \frac{\cos(\gamma(\mathbf{R}_0, \mathbf{n}))}{2\pi c_0 R_0^2} p_0(D_i(\mathbf{R}_0)D_r(\mathbf{R}_0)) \times \mathcal{R}_{01}(\mathbf{R}_0, \mathbf{n}) e' \left( t - \frac{2R_0}{c_0} \right) d\mathbf{S}_{\mathbf{R}}. \quad (25)$$

The equation obtained above gives a rather simple and interesting expression for the time-pressure field response received at  $\mathbf{P}$  from an elementary surface for a given pulse shape in the range of the validity of the Kirchhoff approximation. On a larger scale, the backscattered pressure signal received at the source from the surface  $S$  on the bottom is given by the following integral which adds up the time contribution from each element of surface  $d\mathbf{S}_{\mathbf{R}}$ :

$$p_s(\mathbf{P}, t) = \int_{(S)} dp_s(\mathbf{P}, t). \quad (26)$$

Unfortunately, this integral is not analytically expressible without making crude assumptions and has to be numerically computed using a pre-defined surface height field.

### III. SEAFLOOR VOLUME SCATTERING

The analytical construction of the pressure field received at  $\mathbf{P}$  from an elementary volume is based on Ivakin's volume scattering approach.<sup>6</sup> The expression of the local volume backscattered pressure is based on Chernov's Small Perturbation theory.<sup>7,8</sup> This theory considers both sound speed and density variation in the volume. A detailed and updated formulation of this theory can be found in Ref. 9. The construction of this time-pressure field follows the same principles as for the seafloor surface modeling for the two crossings of the interface, but within the volume only the classical small perturbation theory is applied.

First, assuming that we have an incident pressure field  $p_t(\mathbf{R})$  transmitted through the surface and still allowing the tangent plane approximation,<sup>3</sup> we can write

$$p_t(\mathbf{R}) = (1 + \mathcal{R}_{01}(\mathbf{R}_0, \mathbf{n})) p_i(\mathbf{R}), \quad (27)$$

$$\frac{\partial}{\partial n} p_t(\mathbf{R}) = (1 + \mathcal{R}_{01}(\mathbf{R}_0, \mathbf{n})) \frac{\partial}{\partial n} p_i(\mathbf{R}). \quad (28)$$

Then, as in Eq. (6), the pressure field of the transmitted wave through the water-sediment interface takes this form at point  $\mathbf{R}$ :

$$p_t(\mathbf{R}) = p_0 G_i(\mathbf{R}_0) D_i(\mathbf{R}_0) \mathcal{T}_{01}(\mathbf{R}_0, \mathbf{n}), \quad (29)$$

with  $\mathcal{T}_{01}(\mathbf{R}_0, \mathbf{n}) = 1 + \mathcal{R}_{01}(\mathbf{R}_0, \mathbf{n})$  being the local water-sediment plane wave transmission coefficient.

Following Yamamoto<sup>9</sup> by assuming that sound speed and density variations are small compared to their actual respective mean values, in the backscattering case, the solution to the propagation equation in a medium containing variations of sound speed and density can be written as follows:

$$dp_v(\mathbf{R}) = \frac{k_1^2}{2\pi} \mu(\mathbf{R}') p_t(\mathbf{R}') G_v(\mathbf{R}_1) d\mathbf{V}_{\mathbf{R}}. \quad (30)$$

Here,  $k_1$  is the average wave number in the sediment and  $\mu(\mathbf{R}')$  being defined as the degree of inhomogeneities at the location  $\mathbf{R}' = \mathbf{R} + \mathbf{R}_1$  (Fig. 1):

$$\mu(\mathbf{R}') = \gamma_c(\mathbf{R}') + \gamma_\rho(\mathbf{R}'). \quad (31)$$

$\gamma_c$  and  $\gamma_\rho$  are, respectively, the relative fluctuation of the sound speed and of the density at the same location:

$$\gamma_c(\mathbf{R}') = \frac{c_1(\mathbf{R}') - \bar{c}_1}{\bar{c}_1}, \quad (32)$$

$$\gamma_\rho(\mathbf{R}') = \frac{\rho_1(\mathbf{R}') - \bar{\rho}_1}{\bar{\rho}_1}. \quad (33)$$

Here,  $\bar{c}_1$  and  $\bar{\rho}_1$  are, respectively, the average sound speed and density in the entire insonified volume. For core data analysis, it is possible to evaluate separately  $\gamma_c$  and  $\gamma_\rho$ . These core analyses (as noted by Hamilton<sup>10,11</sup> and Yamamoto,<sup>9</sup> for example) have shown that these two variables were not independent. As an approximation, it is possible to choose a relation of proportionality between  $\gamma_c$  and  $\gamma_\rho$ . This assumption is not always valid, but as density variation and sound speed variation approximations are difficult, it may be appropriate to use global estimation of the degree of inhomogeneities of volume  $\mu(\mathbf{R}')$  when separate measurements of  $\gamma_c$  and  $\gamma_\rho$  are not available. Considering now Eq. (30), the resulting pressure field created at the elementary volume  $d\mathbf{V}_{\mathbf{R}} = d\mathbf{S}_{\mathbf{R}} dR_1$  can be expressed as follows:

$$p_t(\mathbf{R}') = p_0 G_i(\mathbf{R}_0) D_i(\mathbf{R}_0) \mathcal{T}_{01}(\mathbf{R}_0, \mathbf{n}) e^{-jk_0 \bar{n}_1 R_1 - (BR_1/2)}. \quad (34)$$

Following Ivakin *et al.*,<sup>6</sup> spherical spreading in the sediment is neglected because propagation in the sediment is strongly attenuated due to the absorption coefficient  $\beta$ . However, this hypothesis may not be valid in the critical angle region. Rigorously,  $n_1(\mathbf{R}')$ , the refractive index at location  $\mathbf{R}'$  in the sediment, should also be used in the propagation loss term, but the averaged refractive index  $\bar{n}_1$  is used instead over the first meters of the sediment, disregarding the possible refraction effect within the sediment itself. According to Ivakin,<sup>6</sup> the Green's function  $G_v(\mathbf{R}_1)$  of Eq. (30) can be approximated by

$$G_v(\mathbf{R}_1) = e^{-jk_0 \bar{n}_1 R_1 - (BR_1/2)}. \quad (35)$$

Combining Eq. (35) and Eq. (30), the following expression of the backscattered pressure field just under the surface is obtained:

$$dp_v(\mathbf{R}) = \frac{k_0^2 n_1^2(\mathbf{R}')}{2\pi} \mu(\mathbf{R}') p_t(\mathbf{R}) e^{-2jk_0 \bar{n}_1 R_1 - BR_1} d\mathbf{V}_{\mathbf{R}}. \quad (36)$$

Then, back above the surface using again the tangent plane approximation with Eq. (29) and Eq. (36),  $dp_v(\mathbf{R})$  becomes

$$dp_v(\mathbf{R}) = \frac{k_0^2 n_1^2(\mathbf{R}')}{2\pi} \mu(\mathbf{R}') p_0 G_i(\mathbf{R}_0) D_i(\mathbf{R}_0) \times \mathcal{T}_{01}(\mathbf{R}_0, \mathbf{n}) \mathcal{T}_{10}(\mathbf{R}_1, -\mathbf{n}) \times e^{-2jk_0 \bar{n}_1 R_1 - \beta R_1} d\mathbf{V}_{\mathbf{R}}. \quad (37)$$

By analogy with what was done for the interface back-scattering at a given frequency  $f$ , the pressure field received at the source point  $P$  takes the form:

$$dp_v(\mathbf{P}, f) = \frac{k_0^2 n_1^2(\mathbf{R}')}{2\pi} \mu(\mathbf{R}') p_0 D_i(\mathbf{R}_0) D_r(\mathbf{R}_0) \times \mathcal{T}_{01}(\mathbf{R}_0, \mathbf{n}) \mathcal{T}_{10}(\mathbf{R}_1, -\mathbf{n}) G_i(\mathbf{R}_0) G_r(\mathbf{R}_0) \times e^{-2jk_0 \bar{n}_1 R_1 - \beta R_1} d\mathbf{V}_{\mathbf{R}}. \quad (38)$$

In the case of backscattering,  $G_i(\mathbf{R}_0) G_r(\mathbf{R}_0) = G(\mathbf{R}_0)^2$  is the squared Green's function ( $1/R_0^2$ )  $e^{-2jk_0 R_0}$ . For a non-cw signal of spectrum  $E(f)$ , we consider that the attenuation  $\beta$  in the seafloor is frequency dependent, as suggested by Clay and Medwin<sup>3</sup> (p. 260):

$$\beta = \alpha|f|, \quad (39)$$

where  $\alpha$  is the attenuation coefficient which is assumed to be fixed at a given porosity. Equation (38) becomes

$$dp_v(\mathbf{P}, t) = \frac{2\pi n_1^2(\mathbf{R}')}{c_0^2} \mu(\mathbf{R}') p_0 \times D_i(\mathbf{R}_0) D_r(\mathbf{R}_0) \mathcal{T}_{01}(\mathbf{R}_0, \mathbf{n}) \mathcal{T}_{10}(\mathbf{R}_1, -\mathbf{n}) \times \mathcal{F}^{-1}(e^{-\alpha|f|R_1} f^2 G(\mathbf{R}_0)^2 e^{-2jk_0 \bar{n}_1 R_1} E(f)) d\mathbf{V}_{\mathbf{R}}, \quad (40)$$

where the attenuation loss term  $e^{-\alpha|f|R_1}$  appears with the other frequency dependent terms. Let us define  $\Gamma_{(v)}(\mathbf{R}_0, t)$  as the following:

$$\Gamma_{(v)}(\mathbf{R}_0, f) = \mathcal{F}^{-1}(e^{-\alpha|f|R_1} (f^2 G(\mathbf{R}_0, f)^2 e^{-2jk_0 \bar{n}_1 R_1} \times E(f) e^{2\pi j f t})). \quad (41)$$

Let us also define  $\mathcal{K}_v(\mathbf{R}_0, t)$  as the following:

$$\mathcal{K}_v(\mathbf{R}_0, t) = \int_{-\infty}^{+\infty} f^2 G(\mathbf{R}_0, f)^2 e^{-2jk_0 \bar{n}_1 R_1} E(f) e^{2\pi j f t} df. \quad (42)$$

Expressing  $G(\mathbf{R}_0)^2$ , we find

$$\mathcal{K}_v(\mathbf{R}_0, t) = \frac{1}{R_0^2} \int_{-\infty}^{+\infty} f^2 E(f) e^{2\pi j f (t - 2((\bar{n}_1 R_1 + R_0)/c_0))} df. \quad (43)$$

By looking at Eq. (23), we can write

$$\mathcal{K}_v(\mathbf{R}_0, t) = -\frac{1}{(2\pi)^2 R_0^2} e'' \left( t - 2 \left( \frac{\bar{n}_1 R_1 + R_0}{c_0} \right) \right). \quad (44)$$

Fourier transform properties allow us to express  $\Gamma_v(\mathbf{R}, t)$  as follows:

$$\Gamma_v(\mathbf{R}_0, t) = \mathcal{F}^{-1}(e^{-\alpha|f|R_1} * \mathcal{K}_v(\mathbf{R}_0, t)), \quad (45)$$

with  $*$  being the convolution operator. Equation (44) gives an analytical expression of  $\mathcal{K}_v(\mathbf{R}, t)$  in the time domain. Then, Eq. (40) becomes

$$dp_v(\mathbf{P}, t) = \frac{-n_1^2(\mathbf{R}')}{2\pi R_0^2 c_0^2} \mu(\mathbf{R}') p_0 D_i(\mathbf{R}_0) D_r(\mathbf{R}_0) \mathcal{T}_{01}(\mathbf{R}_0, \mathbf{n}) \times \mathcal{T}_{10}(\mathbf{R}_1, -\mathbf{n}) \mathcal{F}^{-1}(e^{-\alpha|f|R_1}) * e'' \left( t - 2 \left( \frac{\bar{n}_1 R_1 + R_0}{c_0} \right) \right) d\mathbf{V}_{\mathbf{R}}. \quad (46)$$

This expression is less simple than Eq. (25) because it contains a convolution operation but accounts reasonably well for the frequency content of the time-backscattered signal from an elementary volume. From this last expression, the backscattered pressure signal received at the source from a volume ( $V$ ) located under the seafloor surface ( $S$ ) is given by the integral:

$$p_v(\mathbf{P}, t) = \int_{(V)} dp_v(\mathbf{P}, t). \quad (47)$$

As for the surface contribution, this last expression is not analytically computable. It should be numerically computed using a pre-defined volume inhomogeneities field.

#### IV. INTERFACE AND VOLUME SCATTERING INTEGRAL

The total pressure field received at  $P$  from the sea bottom is the sum of the pressure field from the surface Eq. (26) and the pressure field from the volume Eq. (47):

$$p(\mathbf{P}, t) = p_s(\mathbf{P}, t) + p_v(\mathbf{P}, t) = \int_{(S)} dp_{(s)}(\mathbf{P}, t) + \int_{(V)} dp_{(v)}(\mathbf{P}, t). \quad (48)$$

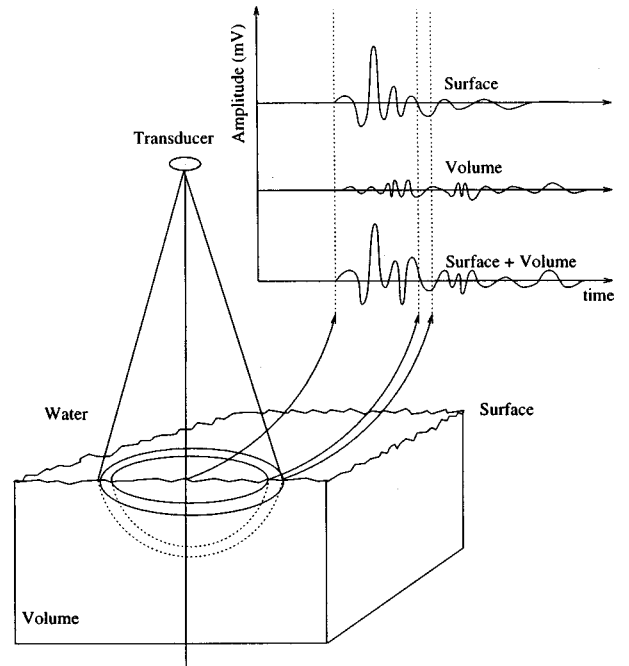


FIG. 2. Construction of the time series.

An example of the construction of the surface and volume time series is presented in Fig. 2 in the case of normal incidence. At a given time  $t$ , it shows which part of the seafloor surface and of the seafloor volume is contributing at source **P**.

## V. CONCLUSION

The prime objective of this model is to understand the physical processes occurring during echo sounding. To account as accurately as possible for the main factors influencing the time series, the domain of validity of this model is not completely defined. The pertinence of the Kirchhoff approximation for extreme cases such as low frequency or moderate-to-high incident angles is still subject to debate. Similarly, the Small Perturbation theory may not be applicable to highly inhomogeneous sediments. Another limitation of this model is that shear waves, interface waves, slow waves, and multiple scattering are not taken into account as these phenomena are not easily measurable at sea, and their significance is often negligible in high frequency applications. Gassy sediment creating nonlinear effects which may be of particular importance is not accounted for in this model. These limitations are largely compensated for by the possibility of generating the whole time-series backscattered from the seafloor for a well-described configuration, instead of focusing on the unpredictable and little understood average quantity of the seafloor backscattering strength. For many years, backscattering strength has been measured and modeled for various seafloor types but results are not consistent. For example, there is not a one-to-one relation between a seafloor and its corresponding evolution of the backscattering strength as a function of the incident angle at a given

frequency. The error in the ground truthing, in the backscattering strength measurement, and in the modeling is too important to allow sole use of backscattering strength for classification or characterization purposes. Both surface and volume contributions must be taken into account. The modeling of the time-pressure field is one way to achieve this.

## ACKNOWLEDGMENTS

This work is partially funded by the MAST-III-ISACS project. The authors wish to thank colleagues at SACLANT-CEN for helpful comments.

- <sup>1</sup>L. Brekhovskikh and Y. P. Lysanov, *Fundamentals of Ocean Acoustics* (Springer-Verlag, New York, 1982).
- <sup>2</sup>E. I. Thorsos, "The validity of the Kirchhoff approximation for rough surface scattering using a Gaussian roughness spectrum," *J. Acoust. Soc. Am.* **83**, 78–92 (1988).
- <sup>3</sup>C.S. Clay and H. Medwin, *Acoustical Oceanography: Principles and Applications* (Wiley, New York, 1970), Chap. 10, Append. 10.
- <sup>4</sup>N. P. Chotiros, "Reflection and reverberation in normal incidence echo-sounding," *J. Acoust. Soc. Am.* **96**, 2921–2929 (1994).
- <sup>5</sup>P. M. Morse, *Vibration and Sound*, 2nd ed. (McGraw-Hill, New York, 1948), p. 327.
- <sup>6</sup>A. N. Ivakin and Y. P. Lysanov, "Underwater sound scattering by volume inhomogeneities of a medium bounded by a rough surface," *Sov. Phys. Acoust.* **27**, 212–215 (1981).
- <sup>7</sup>L. A. Chernov, *Wave Propagation in a Random Medium*, Part 2 (McGraw-Hill, New York, 1960).
- <sup>8</sup>M. Bruneau, *Introduction aux theories de l'acoustique* (Universite du Maine, 1983).
- <sup>9</sup>T. Yamamoto, "Acoustic scattering in the ocean from velocity and density fluctuation in the sediments," *J. Acoust. Soc. Am.* **99**, 866–879 (1996).
- <sup>10</sup>E. L. Hamilton, "Compressional-wave attenuation in marine sediments," *J. Acoust. Soc. Am.* **37**, 620–646 (1972).
- <sup>11</sup>E. L. Hamilton and R. T. Bachman, "Sound velocity and related properties of marine sediments," *J. Acoust. Soc. Am.* **72**, 1891–1903 (1982).

# Time-evolution modeling of seafloor scatter. II. Numerical and experimental evaluation

O. Bergem, E. Pouliquen, G. Canepa, and N. G. Pace  
 SACLANTCEN, Viale S. Bartolomeo 400, I-19138 La Spezia, Italy

(Received 12 June 1997; accepted for publication 22 February 1999)

A time-evolution model of seafloor scatter is numerically implemented and experimentally evaluated. The model is based on analytically expressing the elementary time-backscattered response of every seafloor surface and every seafloor volume infinitesimal element. The implementation of the model is based on a statistical realization of the seabed interface and volume inhomogeneities, from which the time series are computed by coherent summation of the scatter from small elements over the insonified area and volume. The analytical expressions and the implementation are evaluated for the image solution case, for which an almost perfect agreement is found. Examples are shown of how the beam width and seabed roughness affect the time-series return from both the surface and from the volume. The results of the model are compared with data from two different bottom types recorded with a parametric sonar. Reasonable accordance is found between the model and the data. © 1999 Acoustical Society of America. [S0001-4966(99)02006-8]

PACS numbers: 43.30.Gv, 43.30.Hw, 43.30.Ft, 43.30.Pc [DLB]

## INTRODUCTION

In Part I,<sup>1</sup> analytical expressions of the time-evolution approach of the seafloor scatter were presented in which sound pressure level is represented as a function of time at the receiver. The equations account for scattering from the seafloor interface and from the inhomogeneities in the volume below the interface. The water column above the seafloor interface is assumed to be homogeneous with a constant sound velocity profile.

In this paper the implementation and evaluation of the model are discussed. Two main issues must be considered when trying to evaluate the model: the validity of the theoretical basis of the model and the validity of the implementation itself. The first issue is difficult, because only for very simple cases can an analytical expression for the time-domain signal be derived. One such case is the image solution, where the source and receiver are located at the same position above a flat, perfectly reflecting bottom. Verification of both the equations and the model implementations is carried out for this special case.

One form of evaluation is to observe the output of the model response for different input parameters. Since the model computes the return from the seafloor surface and volume separately, both returns can be studied and evaluated. For the surface part, the model is evaluated as a function of rms roughness, as a function of different transmit beam patterns and as a function of the surface spectra. For the volume part, the model is run varying the sediment attenuation, the transmit beam patterns, and the correlation length of the volume.

In order to further evaluate the model, the results are compared with data from two different bottom types: sand and clay.

## I. IMPLEMENTATION

The main principles of the implementation are described in this section in order to understand the way the model is

constructed. The computer model is called BORIS which is an abbreviation for "BOttom Response from Inhomogeneities and Surface." A detailed specification of the implementation of BORIS is given in Ref. 2.

The following integrals from Ref. 1 are the starting points for the implementation:

$$p(\mathbf{P}, t) = p_s(\mathbf{P}, t) + p_v(\mathbf{P}, t) = \int_S dp_s(\mathbf{P}, t) + \int_V dp_v(\mathbf{P}, t), \quad (1)$$

which expresses that the pressure field received at  $\mathbf{P}$  from the seafloor is the sum of the elementary pressure fields integrated over the seafloor surface ( $S$ ) and the seafloor volume ( $V$ ). For a monostatic source and receiver far from the seabed with directivity pattern  $D_i$  and  $D_r$ , the seafloor surface contribution is given by

$$dp_s(\mathbf{P}, t) = \frac{\cos(\gamma(\mathbf{R}_0, \mathbf{n}))}{2\pi c_0 R_0^2} p_0 [D_i(\mathbf{R}_0) D_r(\mathbf{R}_0)] \mathcal{R}_{01}(\mathbf{R}_0, \mathbf{n}) \times e' \left( t - \frac{2R_0}{c_0} \right) d\mathbf{S}_{\mathbf{R}}. \quad (2)$$

Here  $\mathcal{R}_{01}(\mathbf{R}_0, \mathbf{n})$  is the local water-sediment plane wave reflection coefficient at point  $\mathbf{R}$ ,  $c_0$  is the average sound speed in water,  $\gamma(\mathbf{R}_0, \mathbf{n})$  is the angle between the incident direction  $\mathbf{R}_0$  and the vector  $\mathbf{n}$  normal to the surface at  $(\mathbf{R})$ , and  $p_0$  is the source level.  $e'(t)$  is the time derivative of the transmitted pulse  $e(t)$ .

The volume contribution is given by

$$dp_v(\mathbf{P}, t) = \frac{-n_1^2(\mathbf{R}')}{2\pi R_0^2 c_0^2} \mu(\mathbf{R}') p_0 D_i(\mathbf{R}_0) D_r(\mathbf{R}_0) \mathcal{T}_{01}(\mathbf{R}_0, \mathbf{n}) \times \mathcal{T}_{10}(\mathbf{R}_1, -\mathbf{n}) \mathcal{F}^{-1}(e^{-\alpha|f|R_1}) * e'' \left( t - 2 \left( \frac{\bar{n}_1 R_1 + R_0}{c_0} \right) \right) d\mathbf{V}_{\mathbf{R}}. \quad (3)$$

In this expression,  $\bar{n}_1$  is the average refractive index in the first few meters of the bottom,  $n_1$  is the local refractive index

at location  $\mathbf{R}'$  into the sediment, and  $R_1 = |\mathbf{R}' - \mathbf{R}|$  is the distance of penetration in the sediment.  $\alpha$  is the attenuation coefficient, while  $T_{01}(\mathbf{R}_0, \mathbf{n})$  and  $T_{10}(\mathbf{R}_1, -\mathbf{n})$  are the downward and upward transmission coefficients. The double time derivative of the transmitted pulse  $e(t)$  is denoted by  $e''(t)$ .

These equations give the sound pressure level for a given time  $t$  at position  $\mathbf{P}$  by integration over the surface ( $S$ ) and the volume ( $V$ ). The local incident angle  $\gamma(\mathbf{R}_0, \mathbf{n})$ , the local reflection coefficient  $R_{01}(\mathbf{R}_0, \mathbf{n})$ , and the local degree of inhomogeneities in the volume  $\mu(\mathbf{R}')$  are assumed to be known. Instead of using average quantities for these parameters, a different approach has been taken. Before the calculations are carried out for the integral over  $S$ , one realization of the seafloor surface and volume is calculated (see Sec. II). In other words, one sample from the class of seafloors having the chosen statistical parameters is created and used as input to the model. In this way the local variations will be known, and the integration can be carried out.

The integration is carried out by approximating the integrals with Riemann sums. Discretization is done over square patches on the seafloor, all having the same size, independent of the incident angle. For each surface element, the local incidence angle is computed and the local reflection coefficient is calculated. The contribution from the volume is calculated for each surface element by taking the incident angle into account. This angle, which depends on the local surface slope and the reflection coefficient, gives the direction for the penetration into the volume. The total contribution from the surface and the volume element is then weighted by the actual beam pattern. This approach is followed for the whole surface ( $S$ ). All calculations are done in such a way that the actual phase of the contribution to the final result is preserved.

To be able to simulate data recorded at sea, the model includes a full 3D rotation and position matrices for the source and receiver. This allows for variations of the source position, heave, pitch, and roll.

It is important to notice that the model has a stochastic nature, and that the result from one run to another will be different, even with the same input parameters. This is because the seafloor realization will normally change between runs. The model allows seafloor shape to be constant to prevent this behavior. A simplified block diagram of the model is shown in Fig. 1.

One of the problems with the model is the storage requirement for the seafloor surface, and in particular for the seafloor volume. Depending on the resolution and frequency, the storage required for these arrays can easily be unmanageable. In the implementation this problem is solved by generating a smaller part of the surface and volume, and then building the total surface and volume by re-using the same part in different orientations and angles. The process is done such that no discontinuities are created in the interfaces between the parts. Care has also been taken to ensure that the symmetry of the bottom is minimized with respect to the beam pattern. However, this simulation artifact may introduce errors in the result because the surface and volume may be periodic. By keeping the block size reasonably large com-

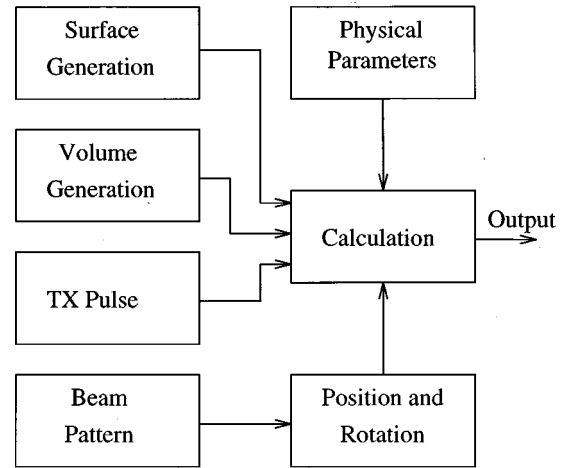


FIG. 1. Simplified box diagram of the model.

pared with the wavelength of the signal, the effect of the error can be reduced.

## II. SEAFLOOR GENERATION

### A. Seafloor surface generation

Limited information on seafloor surface characteristics is available. Only camera pictures, geologist experiences, and a few roughness spectrum measurements<sup>3</sup> contribute to build an idea about the morphology of the interface. This knowledge is often limited to a range of spatial wave numbers. However, within that range Briggs<sup>3</sup> found that seabed interface spectra were close to a power law. In this particular study, we propose a filtered power law spectrum  $W_s$  to produce an isotropic field  $h$ :

$$\begin{aligned}
 W_s(\mathbf{K}) &= \eta \mathbf{K}^{-\nu} & \text{if } K_{hp} < |\mathbf{K}| < K_{lp} \\
 &= 0 & \text{if } |\mathbf{K}| \leq K_{hp} \\
 &= 0 & \text{if } |\mathbf{K}| \geq K_{lp}.
 \end{aligned} \tag{4}$$

Here  $\mathbf{K}$  is a two-dimensional wave vector with a magnitude equal to the wave number  $K$ .  $K_{hp}$  is the high pass cutoff wave number,  $K_{lp}$  is the low pass cutoff wave number.  $\nu$  is related to the fractal dimension and accounts for the decay of the power spectrum.  $\eta$  is the normalization factor so that locally:

$$\int_{(s)} W_s(\mathbf{K}) d^2\mathbf{K} = \sigma_h^2. \tag{5}$$

Here,  $\sigma_h$  is the rms height of the isotropic generated height field  $h(\mathbf{r})$ , assuming that this process is locally stationary. The high and low pass filtering of the power spectrum has a physical meaning. By removing the low frequency content of the spectrum, the large scale oscillations are canceled, which is sensible if one considers that sedimentary regions are more often flat than featured with rolling hills. The low pass filtering prevents the generated interfaces from looking too rough at a very small scale.

The seabed interface generation is based on the Fourier synthesis,<sup>4</sup> which is often used for generating land terrain

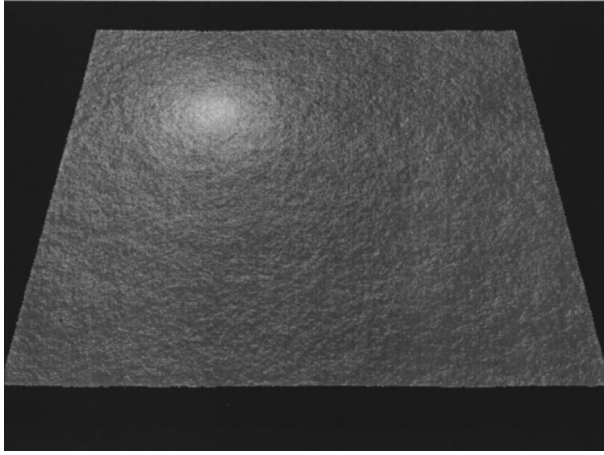


FIG. 2. 10 m×10 m generated surface of a muddy seafloor with high bioturbation:  $\sigma_h=2.5$  cm,  $\nu=3$ ,  $K_{hp}=1$  rad/m,  $K_{lp}=200$  rad/m.

models. It first consists of producing a random power spectrum  $W_s$ , which is converted into amplitude fields in the spatial domain using a Fourier transform. The same method is valid for any chosen spectrum, isotropic or not.

To illustrate the surface generation, two seafloor realizations are presented in Figs. 2 and 3 using image rendering techniques. The scenes are enlightened by a source light located 1 m above the average height of the seafloor. The first image shows an isotropic muddy sediment with high bioturbation, while the second image shows an isotropic sandy bottom.

### B. Seafloor volume generation

Marine sediments are inhomogeneous, which means that there are spatial fluctuations in the sound speed, density, attenuation, porosity, or grain size. These fluctuations have important consequences for the geo-acoustical properties of the seafloor, but unfortunately these variations are particularly difficult to measure to the required spatial resolution. The use of core sampling is not adequate because it only gives sparse and vertical information on the actual 3D random inhomogeneity field. The use of cross-well tomography

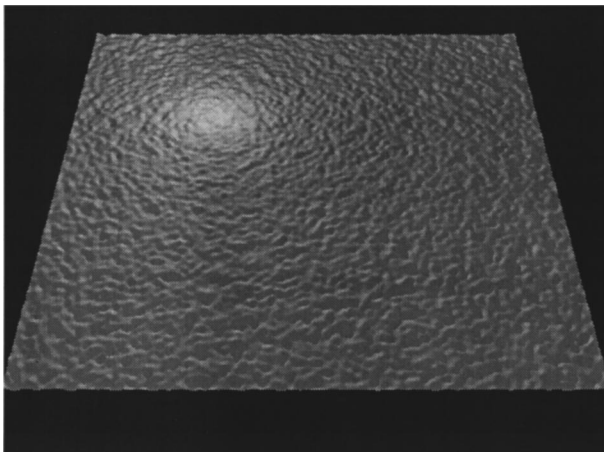


FIG. 3. 10 m×10 m generated surface of an isotropic sandy seafloor.  $\sigma_h=2.5$  cm,  $\nu=3$ ,  $K_{hp}=10$  rad/m,  $K_{lp}=40$  rad/m.

is an alternative way of measuring fluctuations in the sediment, but so far little information is available for small scale variations. Generally, measurement of volume inhomogeneities has been little addressed until recently because volume scattering in high frequency acoustics has been considered a secondary (and often negligible) physical phenomena. The physical entity for characterizing seafloor volume inhomogeneities is usually the 3D power spectrum of correlation  $W_v$ , which is the Fourier transform of the 3D correlation function  $B_v$  accounting for sound speed and density variations around mean values. Since knowledge of the correlation function is weak, there is little evidence for arguing for or against the selection of any particular function. In the present paper, an exponential correlation function  $B_v$  has been used<sup>5</sup> but any other spectra such as; for example, the ones proposed by Yamamoto<sup>6</sup> or Tang<sup>7</sup> could be used. This spectrum has this form:

$$B_v(\mathbf{R}''-\mathbf{R}') = \mu^2(\mathbf{R}') \exp\left(-\frac{|z_1|}{l_v} - \frac{r_1}{l_h}\right). \quad (6)$$

Here  $l_v$  and  $l_h$  are, respectively, named vertical correlation length and horizontal correlation length.  $|\mathbf{R}''-\mathbf{R}'|$  represents the local distance from point  $\mathbf{R}'$  to any point  $\mathbf{R}''$  into the sediment. The Fourier transform of the chosen correlation function can be expressed analytically:

$$W_v(\mathbf{K}) = \mu^2(\mathbf{R}') \left( \frac{l_v}{\pi(1+K_z^2 l_v^2)} \frac{l_h^2}{2\pi(1+K_r^2 l_h^2)^{3/2}} \right), \quad (7)$$

where  $\mathbf{R}''-\mathbf{R}' = \mathbf{z}_1 + \mathbf{r}_1$ ,  $r_1 = \sqrt{x_1^2 + y_1^2}$  and  $\mathbf{K} = \mathbf{K}_{z_1} + \mathbf{K}_{r_1}$ .  $B_v$  and  $W_v$  have a symmetrical behavior as regards the vertical axis. As far as the choice of  $l_v$  and  $l_h$ , it is reasonable to assume that sound speed and density are varying more in the vertical direction than in the horizontal direction. Gensane<sup>5</sup> finds that typically  $l_h$  is approximately five times  $l_v$ . This implies that, even in surficial sediments, the idea of layering persists; in other words, sediment fluctuations have longer spatial scales in the horizontal direction. This is valid for slightly perturbed sediments and does not apply for highly bioturbated sediments where isotropy can be assumed ( $l_v = l_h$ ).

### III. THE IMAGE SOLUTION AND REFLECTION LOSS

The simplest test of the model is a comparison with the image solution, which corresponds to the image of the source on the opposite side of a perfectly reflecting, flat surface. For this simple case, an analytical expression can be obtained. For simplicity, assume that  $p_0=1$  at 1 m from the source. With an omnidirectional receiver and transmitter and a flat, perfectly reflecting bottom, Eq. (2) becomes

$$dp_s(\mathbf{P}, t) = -\frac{\cos(\gamma(\mathbf{R}_0, \mathbf{n}))}{2\pi c_0 R_0^2} e' \left( t - \frac{2R_0}{c_0} \right) d\mathbf{S}_{\mathbf{R}}, \quad (8)$$

where  $\theta = \gamma(\mathbf{R}_0, \mathbf{n})$  is the incident angle. Looking at the full

integral with a source at height  $H$  above the bottom ( $\cos \theta = H/R_0$ ), one obtains

$$p_s(\mathbf{P}, t) = - \int_{(S)} \frac{\cos(\theta)}{2\pi c_0 R_0^2} e' \left( t - \frac{2R_0}{c_0} \right) d\mathbf{S}_{\mathbf{R}} \quad (9)$$

$$= \int_{(S)} \frac{H}{2\pi c_0 R_0^3} e' \left( t - \frac{2R_0}{c_0} \right) d\mathbf{S}_{\mathbf{R}}. \quad (10)$$

Using polar coordinates, the above equation can be written as

$$p_s(\mathbf{P}, t) = - \int_0^\infty \int_0^{2\pi} \frac{H}{2\pi c_0 R_0^3} e' \left( t - \frac{2R_0}{c_0} \right) r d\theta dr \quad (11)$$

$$= - \frac{H}{c_0} \int_0^\infty \frac{e'(t - 2R_0/c_0)}{R_0^3} r dr. \quad (12)$$

Using the fact that

$$R_0^2 = r^2 + H^2, \quad (13)$$

one can write

$$\frac{de}{dt} = - \frac{c_0 \sqrt{r^2 + H^2}}{2r} \frac{de}{dr}. \quad (14)$$

Using Eq. (13) and Eq. (14), Eq. (12) can be written as

$$p_s(\mathbf{P}, t) = \frac{H}{2} \int_0^\infty \frac{de/dr}{r^2 + H^2} dr. \quad (15)$$

Integrating by parts, one obtains

$$p_s(\mathbf{P}, t) = \frac{He(t - 2\sqrt{r^2 + H^2}/c_0)}{2(r^2 + H^2)} \Big|_{r=0}^\infty + \int_0^\infty \frac{He(t - \sqrt{r^2 + H^2}/c_0)r}{(r^2 + H^2)^2} dr. \quad (16)$$

The first part equals

$$- \frac{e(t - 2H/c_0)}{2H}, \quad (17)$$

which is the image solution expected from a flat, perfectly reflecting surface at a distance  $H$ . The contribution of the last integral will depend on the transmitted pulse as well as height  $H$  and sound speed  $c_0$ . If the pulse has a dominant frequency  $f_0$  the last term of Eq. (16) will be small compared to the image solution if  $\lambda_0 \ll H$ , which reflects the far field assumption made for the Kirchhoff approximation.

In Fig. 4 the result of a run from the model is plotted with the theoretical solution for a 5-kHz Ricker pulse. The source and receiver are situated 10 m above a flat, infinitely hard bottom. The sampling rate of the signal was 1 MHz and the surface resolution 0.005 m. These values were selected in order to minimize numerical inaccuracies. There is no discernible difference between the two pulses for this special case. As the bottom is perfectly flat and homogeneous, the seafloor realization will be unique, and no variations will be seen in the model between different runs.

Another simple test of the model can be carried out by comparing the model return with the Rayleigh reflection coefficient given by

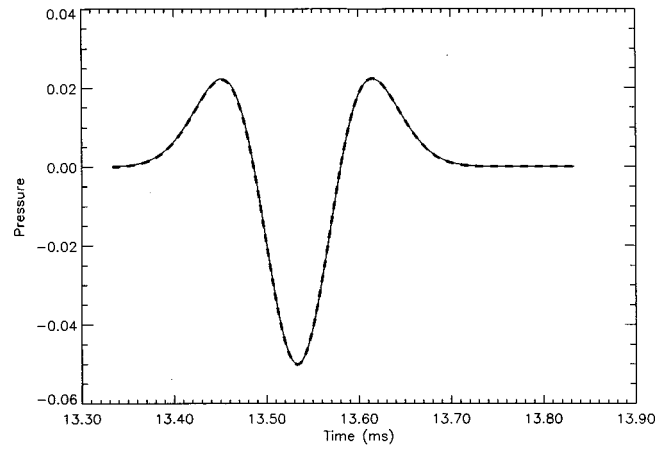


FIG. 4. Comparison between the theoretical solution (solid) and the model (dashed) for a flat, infinitely hard bottom. The pulse is a Ricker pulse and the source–receiver location is 10 m above the bottom.

$$\mathfrak{R} = \frac{(Z_1 - Z_0)}{Z_1 + Z_0}, \quad (18)$$

where  $Z_1$  and  $Z_0$  are the acoustic impedances of the seafloor and the water column, respectively. This expression is valid for plane waves at normal incidence. If the bottom is flat and homogeneous, the source is located far from the seafloor (in term of wavelengths) and the beam width is wide, the reflection coefficient for a spherical wave can be approximated by

$$\mathfrak{R} \approx \frac{\int_{2H/c_0}^{2H/c_0+l} p_r(t)^2 dt}{2H \int_0^l p_t(t)^2 dt}. \quad (19)$$

Here  $p_r(t)$  and  $p_t(t)$  are the reflected and transmitted pressures, respectively, and  $l$  the length of the transmitted pulse.

The model was run using the previous Ricker pulse for three different bottom types: sand, silt, and clay. The reflection coefficients were calculated using Eq. (19), and the results are shown in Table I. The change in impedance is easily seen in the results computed by the model; for the softer bottoms the energy penetrates into the sediment instead of being reflected.

The values in Table I agree with the reflection coefficients given by Eq. (18) to the third decimal.

## IV. MODEL SIMULATIONS

### A. Seafloor surface contribution

The first simulation examines the effect of the roughness of a sandy seafloor surface on the received pulse. The volume contribution is set to zero to emphasize the effect on the surface roughness. For all runs, the shape of the sea-

TABLE I. Calculated reflection coefficient using Eq. (19) for a flat surface.

Bottom type	Sound speed (m/s)	Density (g/cm <sup>3</sup> )	$\mathfrak{R}$
Sand	1720	1.9	0.371
Silt	1610	1.7	0.292
Clay	1510	1.5	0.203



TABLE II. Model parameter settings.

Parameters	Surface runs	Volume runs	Sand	Clay
Surface increment $dp_s$ (m)	0.03	0.03	0.03	0.03
Surface block length (m)	15	15	15	15
Volume increment $dp_v$ (m)	...	0.025	0.025	0.025
Volume block length (m)	3	3	3	3
Depth $H$ (m)	25.0	25.0	14.5	11.9
Sound speed, $c_1$ (m/s)	1720	1510	1270	1475
Density $\rho_1$ (g/cm <sup>3</sup> )	1.90	1.50	1.90	1.55
Surf. rms rough. $\sigma_h$ (m)	0.03(Varying)	0.025	0.01	0.01
Surf. pow. exp $\nu$	3.0	3.0	3.0	3.0
LP cutoff $K_{lp}$ (rad/m)	40 (Varying)	200	40	200
HP cutoff $K_{hp}$ (rad/m)	10	1	10	1
Volume inhomog. $\mu$	...	0.04	0.02	0.02
Volume Hor. Cor. $l_h$ (m)	...	0.05 (Varying)	0.1	0.05
Volume Ver. Cor. $l_v$ (m)	...	0.01 (Varying)	0.02	0.01
Attenuation $\beta$ (dB/m)	...	0.5 (Varying)	0.9	0.5
Beam pattern	omnidir. (Varying)	10° (Varying)	Topas	Topas

floor interface is the same, i.e., the differences between the surfaces correspond to a scaling in the  $z$ -direction. The beam is omnidirectional, and the transducer is situated 25 m above the bottom. The transmitted pulse is a Ricker pulse centered at 5 kHz. The value of all parameters used is shown in column 2 (surface runs) of Table II. Figure 5 shows the result of the simulation. For all the simulations shown in this section, the y-axes show the pressure relative to a transmitted pulse with maximum absolute value of 1.0. An offset is added to the signals in order to separate them from each other.

For a flat bottom ( $\sigma_h=0.0$  cm in Fig. 5) the returned signal shape closely resembles the transmitted signal with the amplitude reduced by the transmission loss  $1/2H=0.02$  and the reflection loss  $\mathfrak{R}=0.371$ . With an rms roughness of only 0.5 cm, the effect of returns from off-normal incidence is clearly visible. As the roughness increases, the energy is spread out and the shape of the coherent reflection changes. At the extreme case of 5–6 cm rms roughness the coherent reflection has almost disappeared, and the incoherent scattering from off normal incidence dominates the signal.

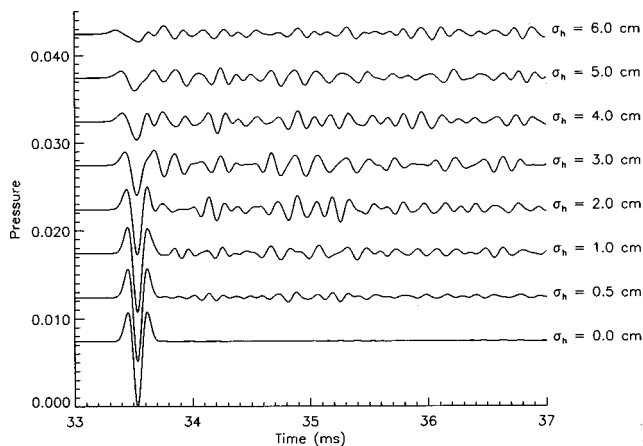


FIG. 5. Returned signal from the seafloor surface as function of rms roughness for an omnidirectional beam.

The second simulation examines the effect of the transmit beam pattern on a sandy seafloor with rms roughness 3.0 cm (Fig. 6). For the purpose of this illustration, the beam patterns of amplitudes versus angles have a Gaussian shape, and the angles in the figure are the half-beam width measured at  $e^{-1}$ . The other parameters for the run are the same as in column 2 (surface runs) in Table II. With a beam width of 30°, the result looks similar to the omnidirectional beam, except that the tail of the signal has less energy due to the reduced beam width. As the beam narrows, the scattering from the side of the beam is reduced, and consequently less energy is seen in the tail. The effect of roughness is, however, not removed. Looking at the amplitude of the 3° beam, one can see that compared to the flat bottom case in Fig. 5 the amplitude of the signal is reduced by a factor of 2 due to the roughness.

Figure 7 shows an example of varying the low pass cutoff frequency  $K_{lp}$  of the surface spectrum. The beam is omnidirectional and the surface rms roughness is 3.0 cm. The other parameters are unchanged with respect to the previous runs. For the low cutoff frequencies, the seafloor is

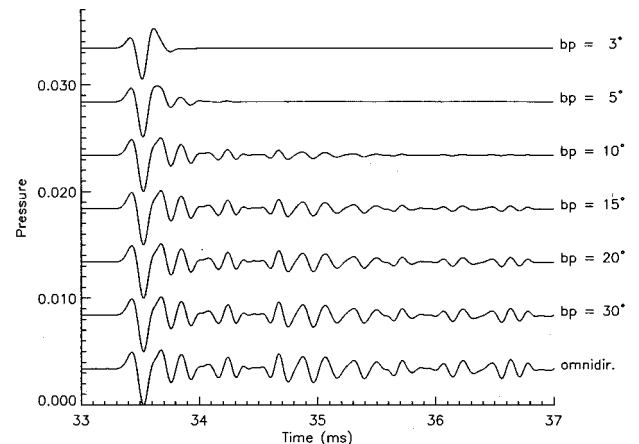


FIG. 6. Returned signal from the seafloor surface as function of the transmit beam pattern. Surface rms roughness 3.0 cm.

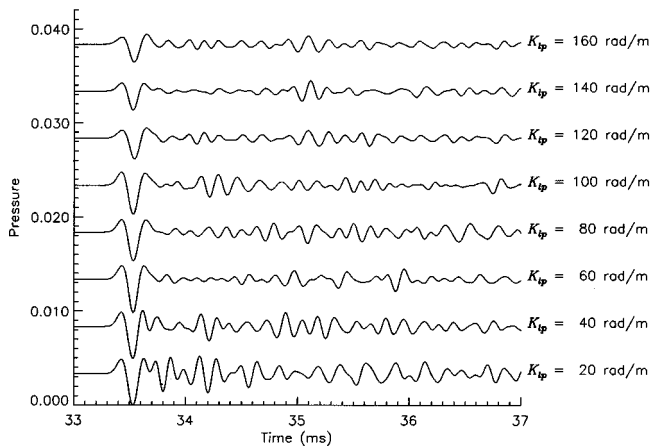


FIG. 7. Returned signal from the seafloor surface as function of the LP cutoff frequency  $K_{lp}$ . Surface rms roughness 3.0 cm, omnidirectional beam.

dominated by low frequency variations which give significant contribution to the returned signal. When the high frequency components are added to the surface, there are more but smaller random variations causing the total return to decrease. This is the case both for the coherent part of the signal and for the tail. It should be noted that a direct comparison between the signals cannot be done since a change in the cutoff frequency changes the shape of the surfaces.

### B. Volume contributions

Volume scattering contributes to the total scatter at a level which is directly proportional to the degree of inhomogeneities  $\mu^2(\mathbf{R}')$  [Eq. (3)]. The result is a weighted and time shifted combination of the second derivative of the transmitted pulse, convolved with the frequency dependent attenuation.

As with the surface contribution, a few test cases have been selected to illustrate the results from the model. The model parameters for the runs are shown in column 3 (volume runs) in Table II. The bottom is composed of silt, the depth is 25 m, and the transmit beam width is  $10^\circ$ . The same realization of the seafloor volume and surface are used in

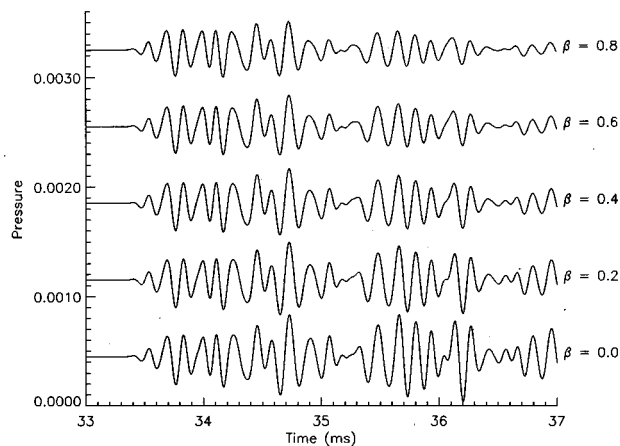


FIG. 8. Returned signal from the seafloor volume as function of the sediment attenuation  $\beta$ .

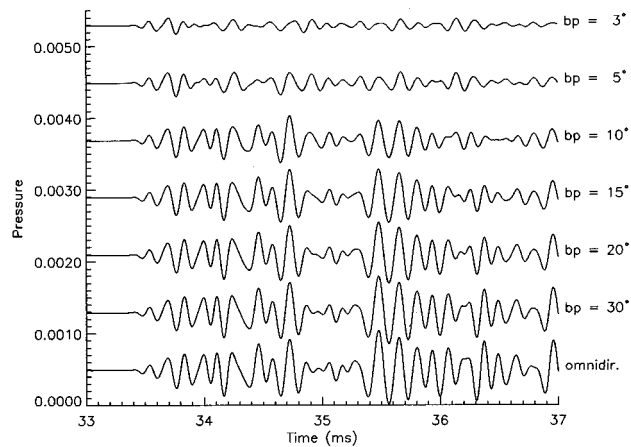


FIG. 9. Returned signal from the seafloor volume as function of the transmit beam width.

each run in order to allow comparison of results. The first case is presented in Fig. 8, where the return from the volume is shown as a function of the attenuation in the sediment. A higher attenuation reduces the energy from the deeper parts of the volume as expected. Looking carefully at the last parts of the signals, one can see that the higher frequencies are more attenuated than the lower (e.g., the knee at 36.05 ms). This is a consequence of using Eq. (3) in the model which includes the convolution process and effectively works as a low pass filter on the returned scatter.

The beam pattern is also important when looking at the volume return. Figure 9 shows examples of the volume return for different transmit beam patterns. The results are interesting, because they show that the scatter from the outer part of the beam contributes extensively to the final result. Hence, a narrow beam does not only reduce the surface scatter, but also the volume contribution.

Figure 10 shows the effect of varying the correlation length in the volume. The correlation is varied such that the ratio between the horizontal and vertical correlation length is equal to 5. The other parameters for the run are given by Table II. It should be noted that changing the correlation lengths implies a change in the realization of the

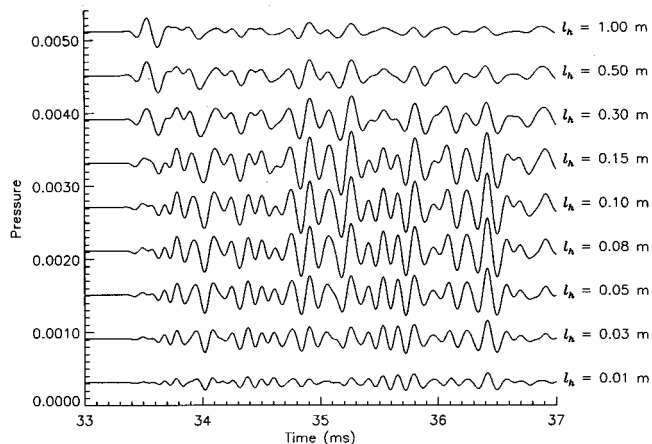


FIG. 10. Returned signal from the seafloor volume as function of the horizontal volume correlation  $l_h$ . The vertical volume correlation is scaled accordingly such that  $l_v = 0.2l_h$ .

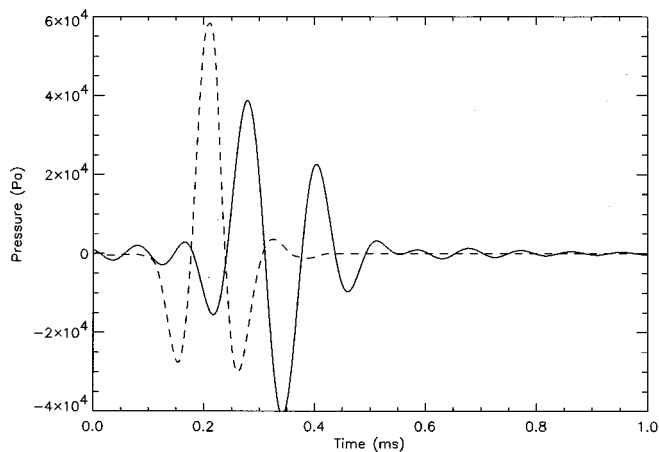


FIG. 11. The transmitted pulse (dashed) and the resulting pulse (solid) after application of the filter which accounts for the effects of the recording system.

shape of the volume, which means that the result cannot be directly compared. It is interesting to see that the maximum volume return is obtained when the vertical correlation length is in the same order as the wave length of the signal. For the very short correlation lengths, the total energy in the volume contributions decreases. This result is expected, because the shorter correlation means increased randomness in the volume. With no horizontal or vertical correlation at all, the total contribution from the volume should be zero. For the long correlation lengths the variations are low with respect to the wavelength, and the scatter from the volume decreases. An infinite correlation length corresponds to an homogeneous volume which has zero volume contribution.

## V. COMPARISON WITH DATA

The Simrad TOPAS 040 parametric system has been used to acquire data from different bottom types in the Gulf of La Spezia, Italy. The parametric sonar has a primary frequency centered at 40 kHz; the secondary frequencies are in the 1–12 kHz range. The sonar was calibrated and mounted in a free floating buoy during the experiments.<sup>8</sup> A Ricker pulse, similar to the one used in the previous simulations, was input to the model. The center frequency of the pulse is around 8 kHz and the source level around 206 dB Ref. 1  $\mu$  Pa. The pulse and the beam pattern used as input to the model are based on data recorded during the calibration. A filter was used during the acquisition of the data; in order to visually compare the results, the same filter has been applied to the time series from the model. The filter includes the effects of the reflections in the buoy and the band pass filter of the recorder. The transmitted pulse together with the effect of the filter is shown in Fig. 11.

The parameters used as input for the simulations are based on core analysis, seismic profiling, and underwater photographs. The calculated values for the sound speed and density should be close to correct, while the roughness parameters from the surface are based on visual estimates from the under water photographs. The volume parameters are based on sound velocity measurements in the cores. How-

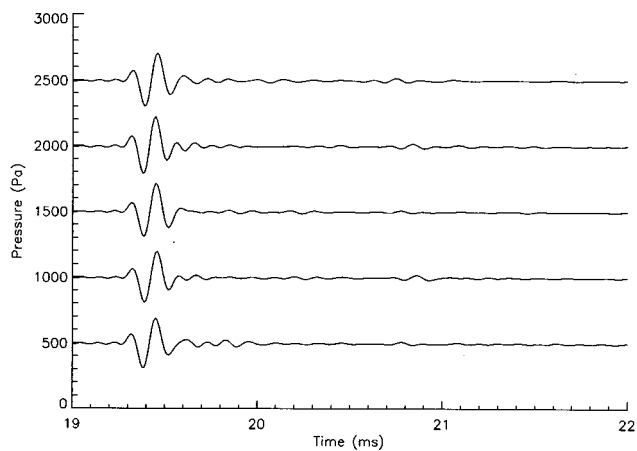


FIG. 12. Recorded data. Bottom: compacted sand.

ever, the measured volume values are uncertain, both due to insufficient resolution in the probes and to disturbance caused by the coring process.

### A. Compacted sand

The Tellaro site is characterized by flat, compacted sand close to the shore with a transition to silt as one moves away from the coast. The data presented here were acquired close to the shore. A traditional seismic profile shows that there are no major sublayers up to 9 ms after the interface, but a relatively high scattering density is seen from the upper part of the sediment. Photographs show very low roughness and some biological activity. Several cores have been taken from the site, but as the corer only penetrates a few decimeters into the sediment, the volume inhomogeneity profile can only be roughly estimated in the upper part. The parameters used as input to the model are shown in column 4 (sand) in Table II.

In Fig. 12, five pings of data recorded at the site are shown. They are all taken with the ship anchored at the same position, but with small differences in heading and position. The strong seafloor reflection dominates the signal, and very little energy is present in the tail.

Figure 13 shows the results of five runs from the BORIS model. The model was run with the same parameters each time, but the realizations of the surface and the volume were

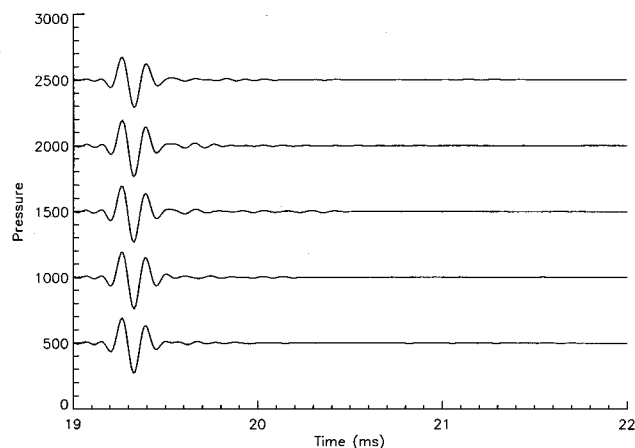


FIG. 13. BORIS simulated data. Bottom: compacted sand.

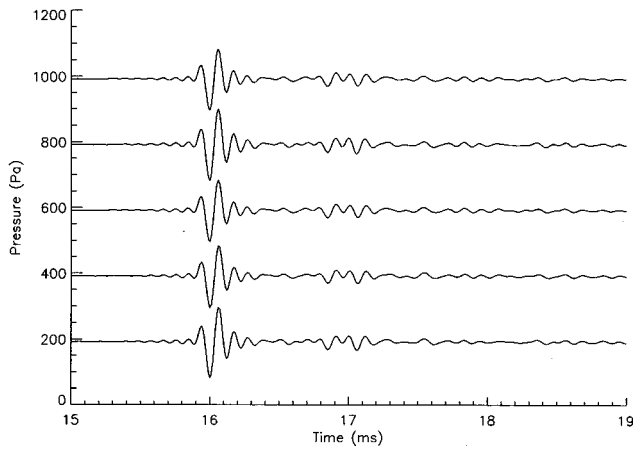


FIG. 14. Recorded data. Bottom: clay. A layer is clearly seen at 17 ms.

different. The correspondence between the signals from the data and the model is good, and the predicted amplitude is very close to the measured values. The runtime for five simulations was of the order of 1 min on a workstation.

### B. Clay

The bottom of the bay of Portovenere consists of a relatively homogeneous layer of clay, but to the south side of the bay, more scattering from the volume is observed. The data presented here are recorded in the latter part. The underwater camera shows some biological activity, but generally the surface is relatively flat. The parameters which have been used are shown in column 5 (clay) of Table II. The data (Fig. 14) show a clear indication of sublayers. This can be observed at all pings around 17 ms. Compared to the sandy site, the amplitudes of the signals are lower and the volume returns are stronger.

The results from five runs of the model are shown in Fig. 15. The variation between the pings is higher than in the data, and the volume return appears at different places because no distinct sublayers have been input to the model. The amplitude of the signals is close to the recorded amplitude.

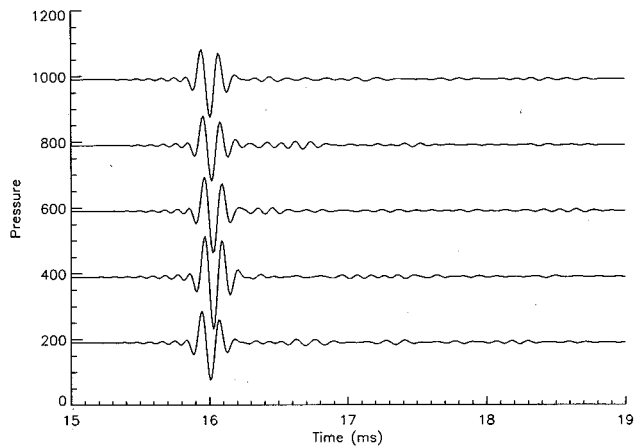


FIG. 15. BORIS simulated data. Bottom: clay. No specific layer is seen due to the different realizations of the volume.

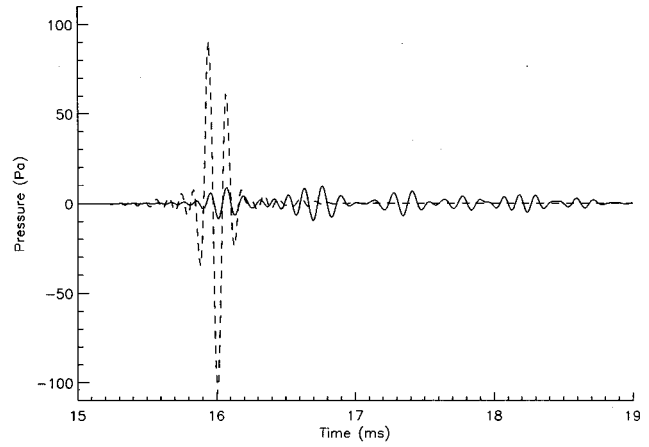


FIG. 16. Simulated surface and volume contributions. Bottom: clay. The solid line shows the volume contribution, the dashed line shows the surface contribution.

To illustrate the contribution from the volume and surface more clearly, the simulated responses are plotted separately in Fig. 16. The solid line shows the volume contribution and the dashed line shows the surface contribution. The coherent sum of these two contributions is the same as the lowest ping in Fig. 15. It is interesting to note that in this case the maximum amplitude of the surface response is about ten times stronger than for the volume.

### C. Comments

The model seems to include the main physical mechanisms of an interaction of a short pulse with the seafloor. The shape and the amplitude of the times series predicted by the model are close to the data recorded at sea. The volume contribution from the model is slightly too low and the surface contribution slightly to high for the clay site. It is difficult to judge whether this is due to errors in the input parameters to the model or to the model itself.

A complication in the comparison between model and data is the effect of the bottom generation. The data are recorded with slight changes in position and heading over the same bottom, while the model is run with a different realization of the bottom for each run. The result is that the change from one run to the other is higher for the model than the change in the recorded data between two pings. An alternative approach would be to use one bottom realization for the model, but to move the source as in the experiment. This may or may not give a better result, depending on the similarity between the real bottom and the generated bottom. Essentially, the problem arises from the fact that a statistical description can only give at best a statistically satisfactory result. A one-to-one correspondence between snapshot data and model cannot be expected.

### VI. SUMMARY

In this paper we have implemented the time-evolution model for seafloor scatter presented in Part 1.<sup>1</sup> The model calculates the time-series return from the seafloor based on scattering from the interface and from inhomogeneities in the

volume. The interface and volume are described with statistical parameters, and for each run the model creates one realization of each of these and calculates the return. Fourier synthesis is used to produce both the surface height field and the volume variations in sound speed and density. For the surface part a filtered power law spectrum has been used, while the volume part uses an exponential correlation function. The final return is a summation of contributions from elementary surface and volume elements based on analytical expressions.

The model was run and results given for different input parameters. The effect of interface roughness was illustrated, and it was shown how different transmit beam patterns affect the time-domain signal. Examples of the volume contributions were also shown. Comparison with data recorded by a calibrated parametric array was carried out for two different bottom types consisting of compacted sand and clay. Good agreement was found between the data and the model both with respect to the shape and the amplitude of the signal. A problem with the comparison was that some of the input parameters to the model were difficult to estimate from the available knowledge of the seafloor. This was especially apparent for the volume inhomogeneity distribution.

The model appears to predict well the scattering from the seafloor interface and volume, and it would appear to be a valuable tool for studying the geometric and environmental effects on the time-domain signals. Further work is needed in order to evaluate the model against different sonar systems, frequencies, and oblique angles of incidence.

## ACKNOWLEDGMENTS

This work is partially supported by EU, under the MAST-III initiative, project "ISACS," Contract No. MAS3-045.

- <sup>1</sup>E. Pouliquen, O. Bergem, and N. G. Pace, "Time-evolution modeling of seafloor scatter. I. Concept," *J. Acoust. Soc. Am.* **105**, 3136–3141 (1999).
- <sup>2</sup>G. Canepa, O. Bergem, and E. Pouliquen, "The implementation of BORIS-3D: BOTTOM Response from Inhomogeneities and Surface, Version 1.0," SAACLANTCEN Special Report, M-125 (1997).
- <sup>3</sup>K. B. Briggs, "Microtopographical roughness of shallow-water continental shelves," *IEEE J. Ocean Eng.* **14**, 360–367 (1989).
- <sup>4</sup>M. F. Barnsley, R. L. Devaney, B. B. Mandelbrot, H. O. Peitgen, D. Saupe, and R. F. Voss, *The Science of Fractal Images* (Springer-Verlag, New York, 1988).
- <sup>5</sup>M. Gensane, "Sea-bottom reverberation: the role of volume inhomogeneities of the sediment," in *Ocean Reverberation*, edited by D. D. Ellis, J. R. Preston, and H. G. Urban (Kluwer Academic, New York, 1993), pp. 59–64.
- <sup>6</sup>T. Yamamoto, "Acoustic scattering in the ocean from velocity and density fluctuation in the sediments," *J. Acoust. Soc. Am.* **99**, 866–879 (1996).
- <sup>7</sup>D. J. Tang, "Small scale volumetric inhomogeneities of shallow water sediments: Measurements and discussion," in "High Frequency Acoustics in Shallow Water," edited by N. G. Pace, E. Pouliquen, O. Bergem, and A. P. Lyons, SAACLANTCEN Conference, Lerici, July 1997, pp. 539–546.
- <sup>8</sup>O. Bergem and N. G. Pace, "Installation and calibration of a parametric array for shallow water backscatter measurements," *Proc. Oceans 96*, Fort Lauderdale, pp. 773–778 (1996).

# Acoustic hot-spots and breaking wave noise in the surf zone

Grant B. Deane

*Marine Physical Laboratory, Scripps Institution of Oceanography, UCSD, La Jolla, California 92093-0238*

(Received 16 March 1998; accepted for publication 1 March 1999)

Measurements of the horizontal directivity of the underwater sound made behind individual breaking waves in the surf zone are presented. To an observer behind a breaking wave, the breaking noise appears to radiate from acoustic hot-spots, which are compact sources at the ends of the breaking wavecrest. The hot-spots move through the surf zone as the length of breaking wavecrest increases, leading to systematic temporal variations in the wave noise directionality. Measurements of the acoustical properties of the mixture of water and bubbles left behind the breaking wave are presented. Sound speeds as low as 500 m/s and acoustic absorption up to 50 dB per meter occur in the bubbly region. Based on these measurements and a propagation model for the surf noise, it is proposed that the observed horizontal directivity is caused by the interaction between the noise radiated by the breaking wavecrest and the residue of bubbly water left behind the wave. © 1999 Acoustical Society of America. [S0001-4966(99)02206-7]

PACS numbers: 43.30.Nb, 43.30.Pc, 43.30.Es [SAC-B]

## INTRODUCTION

The central topic of this paper is the radiation pattern of underwater sound generated by breaking surf. A number of surf noise studies have been published,<sup>1-3</sup> but the experiments reported here represent the first measurements of the horizontal directivity of individual breaking waves in the surf zone. The noise field directionality turns out to be a very interesting property of surf noise. As will be shown, it is intimately related to the large populations of bubbles left in the water column by the passage of the wave generating the sound.

The air entrained by a breaking wavecrest forms bubbles which range in radius from tens of micrometer to centimeters.<sup>3</sup> The bubbles are organized into plumes, which, in the surf zone, can extend a meter or more into the water column. During their creation the plumes radiate sound, with oscillating bubbles on the plumes' outer edges being the primary source of wave noise above about 500 Hz.<sup>3</sup> Newly formed bubbles emit a short pulse of sound before joining the population of quiescent bubbles in the plume and behind the breaking wavecrest. Although quiet, the quiescent bubbles are strong sound absorbers<sup>4</sup> and the close proximity of the newly forming and radiating plumes to the absorbing, bubbly residue behind the wave causes the wave noise to exhibit a distinctive radiation pattern. The breaking noise observed from behind the wave appears to radiate from compact acoustic sources, or "acoustic hot-spots," which follow the ends of the breaking wavecrest.

The major part of this paper explores the interaction between the noise radiated by a breaking wavecrest and the region of bubbly water left behind the wave. This interaction depends on a number of phenomena, including the spatial extent and acoustical properties of the residue of bubbles and the propagation of sound through the range-dependent surf zone. Thus a study of surf noise directionality naturally leads to a discussion of the fluid dynamics of wave breaking, the acoustical properties of bubbly mixtures, and range-dependent propagation phenomena. These diverse topics all

play a role in the physics of surf noise generation and are discussed below in this context.

Section I describes an experiment to measure the sound radiated by individual breaking waves in the surf zone, and presents measurements of the breaking wave noise coherence and horizontal directivity. The properties of the horizontal directivity and the formation of acoustic hot-spots are discussed in Sec. II. The remainder of the paper studies the generation of the hot-spots, which is attributed to the interaction between the noise radiated by the breaking wavecrest and the residue of bubbles left in the water column behind the breaking wave. The results of an experiment to measure the acoustical properties of the bubbly residue are presented in Sec. III. A model for the spatial extent of the bubbly residue and the temporal evolution of the length of a breaking wavecrest are presented in Sec. IV. The results of these two sections are combined in Sec. V, which describes a range-dependent propagation model to compute the interaction between acoustic sources in the wavecrest and the absorbing properties of the residue. The implications of the modeling for the horizontal directivity of the breaking wave noise are discussed in Sec. VI. Concluding remarks can be found in Sec. VII.

## I. MEASUREMENTS OF SURF NOISE DIRECTIONALITY

Figure 1 illustrates the various terms used in connection with breaking waves throughout the paper. The figure shows a wave breaking just north of Scripps Pier. A length scale for the picture can be obtained from the spacing between the pier pilings, which is 9.1 m. The breaking wavecrest is the region of foamy water at the leading edge of the wave. The wave break points are at the edges of the breaking crest, and delineate the region of breaking crest from unbroken crest. The bubbly residue is the region of bubble-filled water trailing the wave. Bubbles rising to the surface from this region create foam which remains visible for a few seconds after the

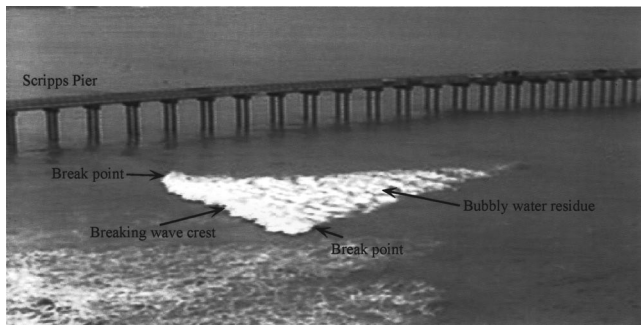


FIG. 1. Annotated picture of a breaking wave crest, illustrating the breaking wavecrest, the wave break points, and the bubbly water residue.

passage of the wave. The edges of the bubbly residue mark the trajectory of the break points through the surf zone.

### A. The passive acoustic experiment

During November 1996 an experiment was conducted in the surf zone off Red Beach at Camp Pendleton, Southern California to measure surf noise and the acoustical properties of bubble plumes entrained by breaking surf as part of the Adaptive Beach Monitoring Experiment. Two International Transducer Corporation 6050 C broadband hydrophones and an ITC 1007 acoustic source were deployed on a frame jettied into the sea floor, as shown in Fig. 2. The hydrophones were fixed 90 cm above the sea floor, and they were placed on an axis aligned in the seaward–shoreward direction. A pressure sensor was deployed to measure wave height, but failed early in the experiment and did not provide usable data. Sound pressure time-series data were brought to shore through a steel armored cable and stored on a digital audio tape recorder. The wave conditions in the region of the frame were recorded by a video camera mounted on a bluff overlooking the experiment site. The video and audio tapes were synchronized at the start of every recording to allow a comparison of the two data sets.

The frame and sensors were deployed in roughly 30 cm of water at low tide. As the tide rose, the frame was gradually submerged—the positions of the low and high tide levels relative to the sensors are shown in Fig. 2. The effect of the tidal variations was to sample different wave breaking conditions over a tidal cycle. At high tide, the water depth was

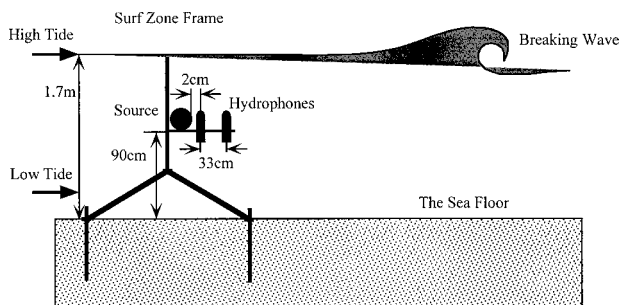


FIG. 2. The surf zone frame and sensor geometry for the surf noise and bubble plume measurements. At high tide, waves tended to break between the frame and the shore line. Wave noise recordings were made during these times. At mid-tide, waves tended to break over the frame allowing measurements of the acoustical properties of the bubble plumes.

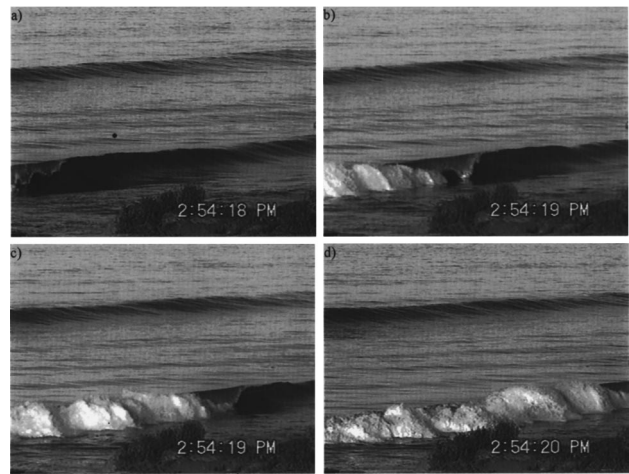


FIG. 3. Pictures of a breaking wave event recorded by a video camera overlooking the experiment site. The image sequence runs from (a) through (d) with a 0.5-s delay between frames. The wave break point in the field of view is moving north. The small black dot near the center of the upper, left hand image marks the location of the sensor frame.

approximately 2 m and the surf broke between the frame and shore line. These conditions permitted measurement of the noise radiated by waves breaking about 25–50 m shoreward of the frame. A little less than 4 h of breaking wave noise was recorded over three consecutive days during high tide. At mid-tide, the waves tended to break over the frame allowing active acoustic measurements of the acoustical properties of bubble plumes. The subject of this section is the wave noise measurements; the bubble plume measurements are discussed in Sec. III.

### B. The spectrum, coherence, and horizontal directivity of breaking wave noise

Figure 3 shows a sequence of four images, spaced at half-second intervals, of a wave breaking event that occurred during the wave noise measurements at Camp Pendleton on 11/9/1996 at 6:54:18 A.M. Pacific Standard Time. The time stamp on the images is different because the clock internal to the video camera was set to Universal Time, which is 8 h ahead of Pacific Standard Time. The still water depth at the time of the event was 1.7 m, which placed the recording hydrophones about 80 cm below the mean water surface. The sequence of images shows the wave break point in the field of view moving north along the shore line (see the Introduction to this section for a definition of the term “break point”). The location of the surf zone frame with the recording array is indicated by the black dot near the center of Fig. 3(a).

A spectrogram of the ambient noise recorded during the breaking event is shown in Fig. 4(a). The spectrogram shows a frequency versus time plot of the ambient noise at the shoreward hydrophone before, during, and after the breaking event. The spectral levels are in dB relative to  $1 \mu\text{Pa}^2/\text{Hz}$ . The noise time series was sampled at 50 kHz and formed into spectral estimates by averaging 3, 50% overlapped, 1024-point Fourier transforms. Individual time-series segments were multiplied by a Hanning window before trans-

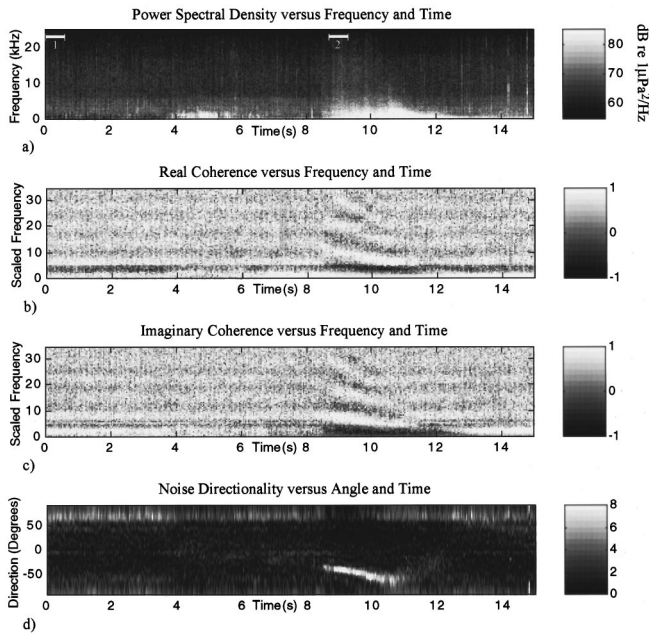


FIG. 4. Analysis of the ambient noise power spectrum and directionality in the surf zone. Noise from a wave breaking event between the array and shore line can be seen between 8.5 and 12 s. (a) Spectrogram of the ambient noise. A broadband rise in the noise level occurs during the breaking event. (b), (c) The real and imaginary parts of the ambient noise coherence plotted as a function of scaled frequency and time. Note the change in sign of the initial excursion of the imaginary part of the coherence at the beginning of the breaking event. (c) The horizontal directionality of the noise field calculated from the broadband coherence, showing that energy from the shoreward side of the array predominates during wave breaking.

forming. Each spectral estimate represents an average of 30.72 ms of data and has a frequency bin width of 48.83 Hz.

The acoustic signature of the breaking event begins approximately 8.5 s into the record, and lasts for roughly 3 s. The noise of the breaking event is clearly distinguishable above the background ensemble of more distant breaking surf and the noise propagating to the array from deeper water. The low frequency sound which occurs between 4 s and 6 s into the record is the noise generated by the passage of the shoaling wave over the frame. Since the pressure sensor failed early in the experiment, the interference caused by both flow noise and disturbance of the hydrophone cables provides a convenient reference for the time of the wave over the array.

Half-second averages of the power spectral density of the noise field before and during the breaking event are shown in Fig. 5, labeled “PSD before breaking” and “PSD during breaking.” The time intervals for these averages are indicated in Fig. 4(a) by the short, white lines at the top of the plot labeled “1” and “2.” The most significant differences in level between the before-breaking and during-breaking spectra occur in the frequency band between 400 Hz and 10 kHz, over which the breaking wave noise can be 15 dB higher than the background levels. It can be concluded that the ambient noise field over this frequency band is dominated by wave noise during the breaking event. The rapid roll-off in the two curves above 22.5 kHz is due to the recording system anti-aliasing filters.

Figure 4(b) and (c) shows plots of the real and imagi-

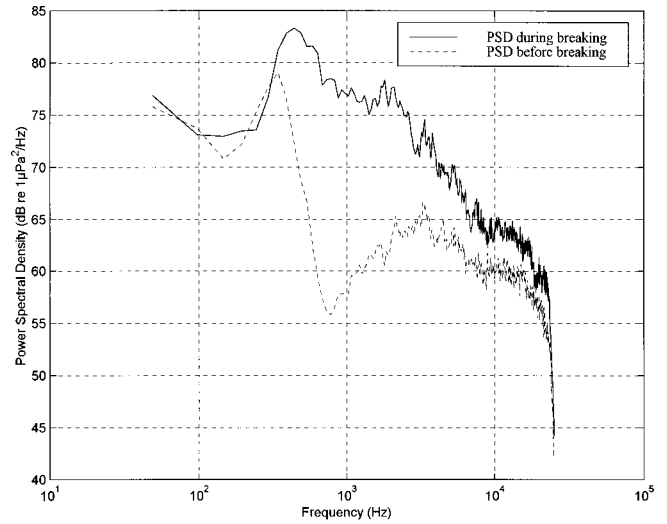


FIG. 5. The ambient noise power spectral density averaged over two 0.5-s intervals before a breaking event and during a breaking event. The time intervals for the before-breaking and during-breaking averages are marked by the two short, white lines near the top of Fig. 4(a), respectively, labeled “1” and “2” on the plot. Note that the ambient noise levels are higher during the breaking event from 400 Hz up to 22.5 kHz, which was the upper limit of the recording system. The rapid roll-off above 22.5 kHz is due to the recording system anti-aliasing filters.

nary parts of the noise field coherence estimated from 30.72 ms segments of data, and plotted as a function of scaled frequency and time. The dimensionless, scaled frequency arises naturally from an analysis of the noise field coherence, and is given by  $\Omega = \omega L/c$ , where  $\omega$  is angular frequency,  $L$  is the separation between the two points at which the coherence is measured, and  $c$  is the sound speed in the water column. The formula used to estimate the coherence is Eq. (A1) of the Appendix, with individual spectral estimates being calculated using the same processing scheme as for the power spectral density estimates (described above). Each of the coherence functions plotted vertically in Fig. 4(b) and (c) is an average of three individual coherence estimates. The figures show that the noise field coherence is stable in frequency and time before and after the breaking event, and exhibits the oscillations in frequency which are characteristic of oceanic ambient noise fields. The coherence of the noise field during the wave breaking event is also oscillatory, but is clearly different from the noise coherence before and after the event, with the difference extending throughout the entire 22.5-kHz frequency band. In addition, the coherence evolves in time throughout the event, implying a time variance in the horizontal directionality of the ambient noise field (see the Appendix for a discussion of the relationship between broadband coherence and noise field directionality).

Averages of the real and imaginary parts of the coherence function, taken over the same intervals as the before-breaking and during-breaking power spectrum averages, are shown in Fig. 6. Note that the initial excursion of the imaginary component of the coherence is positive before the breaking event but becomes negative during the breaking event. This can be clearly seen in the plots of coherence versus time where the horizontal band at the bottom of Fig.



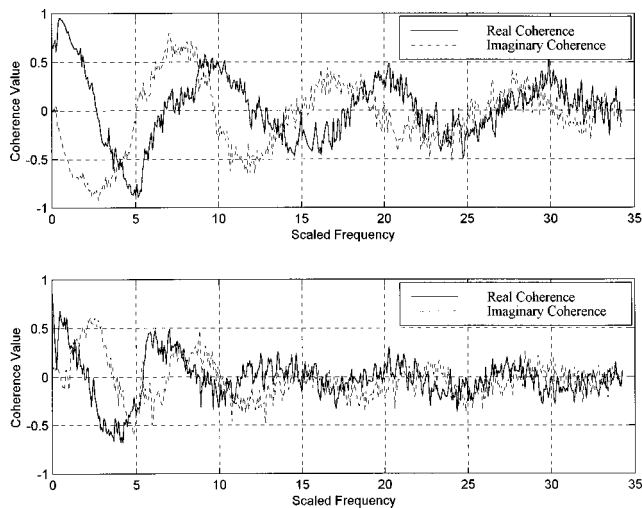


FIG. 6. Plots of the real and imaginary parts of the ambient noise coherence, averaged over the same 0.5-s intervals as the power spectral density averages shown in Fig. 5. Curves such as these form the basis of the noise field directionality calculations. (a) Average coherence during wave breaking. (b) Average coherence before wave breaking. Note that the initial excursion of the imaginary part of the coherence is negative during the breaking event and positive before the breaking event. The change in sign indicates that the main component of ambient noise energy is incident on the array from the seaward direction before breaking and from the shoreward direction during breaking.

4(c) changes from light to dark at the beginning of the breaking event. As discussed in the Appendix, the significance of this change in sign is that the main component of energy incident on the array has changed from seaward to shoreward in origin. The fine horizontal line appearing at roughly  $\Omega = 5$  in Fig. 4(c) is almost certainly caused by the close proximity of the seaward hydrophone to the acoustic source used for the active transmission measurements (see Fig. 2 for the geometry). This interference is an unavoidable consequence of the geometry required for the active acoustic bubble plume measurements reported in Sec. III.

The horizontal noise field directionality has been calculated from each of the coherence estimates with Eq. (A10), and is plotted as a function of angle and time in Fig. 4(d). Negative and positive angles, respectively, correspond to energy from the shoreward and seaward sides of the array. The coherence functions were passed through a low-pass filter and then interpolated to provide a weighted-mean estimate of the coefficients in Eqs. (A8) and (A9). Note that the directional density function as computed satisfies a normalization constraint, and is independent of the total energy incident on the array. The angular resolution of the broadband, two element hydrophone array is determined by the order of the last term in the Fourier series expansion in Eq. (A10). The hydrophone separation of 0.327 m and maximum frequency of 22.5 kHz yield an angular resolution of about 9 degrees for the horizontal directionality estimate.

During the breaking event, the noise field directionality is strongly peaked, with noise from the shoreward direction dominating. The peak in the directionality corresponds to the white line in Fig. 4(d) that extends from 8.5 s to 11 s in the plot. The angle of the peak varies continuously with time, starting at  $-40$  degrees and ending at roughly  $-90$  degrees.

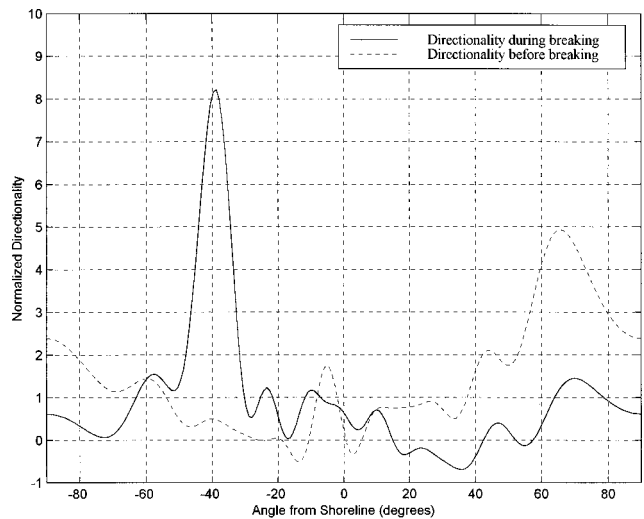


FIG. 7. Plots of the noise horizontal directionality averaged before and during a breaking event. The averages were taken over the same 0.5-s intervals as the power spectral density averages shown in Fig. 5. Before the wave breaks, the sound field is dominated by noise from a kelp bed to the northwest of the array. After the wave breaks, the sound field is dominated by an apparent source at the wave break point which, at the instant shown, is shoreward of the array at an angle of 40 degrees to the shore.

This shows that the wave noise, observed from the seaward side of the breaking wave, is strongly directional and time dependent. I have called this phenomenon an ‘‘acoustic hot-spot’’ because of the apparent existence of a noisy, localized source associated with the breaking wave. In fact, it will be argued later that the hot-spot is created by the screening effect of the bubbly residue left behind the wave rather than a variation in source strength inherent in the breaking wavecrest.

Line plots of the noise field directionality averaged over the same half-second intervals as the coherence and power spectral density estimates before the breaking event and during the breaking event are shown in Fig. 7. The noise field directionality before the breaking event shows a broad peak at 65 degrees, indicating that radiation from sources on the seaward side of the array dominate the noise field before the breaking event. This peak can be seen as a gray, horizontal line near the top of the plot in Fig. 4(d) and was probably due to biological activity within a kelp bed on the northwest side of the array. The directional density function during the breaking event, corresponding to the solid line in the plot, is quite different from the directionality before the breaking event and shows a well defined peak just before  $-40$  degrees. The change in angle of the peak between the two plots shows that before the wave breaks, the noise field is dominated by sources on the seaward side of the array and during the breaking event, the field is dominated by noise from the breaking wave. A small peak can still be seen at 65 degrees in the directional density function during the breaking event, showing that the off-shore sources are still present, but largely masked by the breaking wave noise. The width of the directionality peak during the breaking event corresponds to the minimum beam width of the broadband coherence array, and the actual beam width of the horizontal directionality may have been narrower than the array resolution.

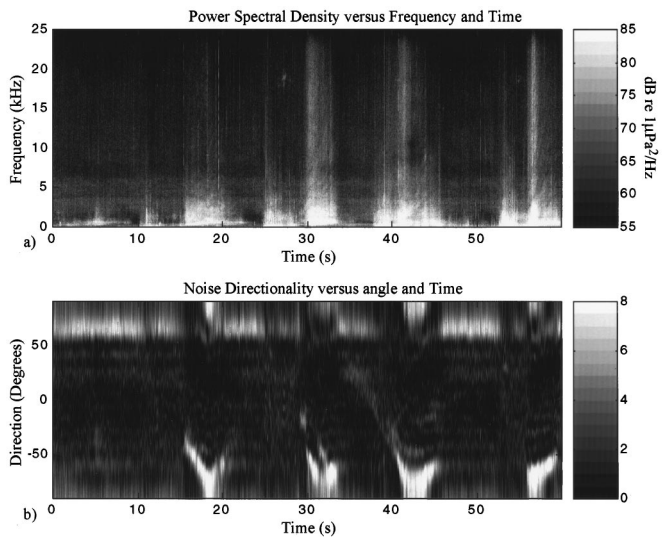


FIG. 8. One minute of surf noise summarized into (a) a spectrogram and (b) horizontal directionality. The noise from four breaking events can be seen in the spectrograms. Each of the breaking events generated a “hot-spot” in the noise field directionality.

The small-scale oscillations in the directional density estimates are a function of the highest order term of the Fourier series expansion at which the expansion was truncated, and should not be interpreted as a feature of the actual noise field. The true directional density function is a measure of power and therefore always positive. Our plots of this function are only estimates, and in fact do go negative in Fig. 7. This is an artifact of the limited bandwidth of the inversion, which forces a truncation of the Fourier series expansion. The negative values do not have any physical significance, and our directionality estimates have been plotted on a positive scale in the directionality versus time plot in Fig. 4(b).

Although an analysis of coherence and horizontal directionality for only a single event has been shown so far, it is representative of the kinds of waves which generated hot-spots, which are plunging breakers on the seaward side of the array. For example, Fig. 8 shows four consecutive breaking events over a 1-min interval. All four events show a strongly peaked directional density function, although the details of how the directionality evolves in time varies from one event to another. The hot-spot phenomenon was consistently observed throughout the data set for plunging breakers on the shoreward side of the surf frame. Occasionally, two hot-spots were observed in the noise directionality from a single breaking wave. An example of such an event is shown in Fig. 9.

A few waves had sufficient height to first break on the seaward side of the array, and these events showed a different pattern in their horizontal directivity than plunging breakers on the shoreward side. Figure 10 shows the power spectral density and horizontal directivity for a wave that first broke to the southwest of the array, on the seaward side. The video recordings show that the entire length of the breaking crest passed on the southern side of the array on its shoreward journey. The horizontal directionality of the wave noise, as estimated from the noise field coherence, changes character as the breaking crest passes by the array (the passing of the array is indicated in the plot by a change from

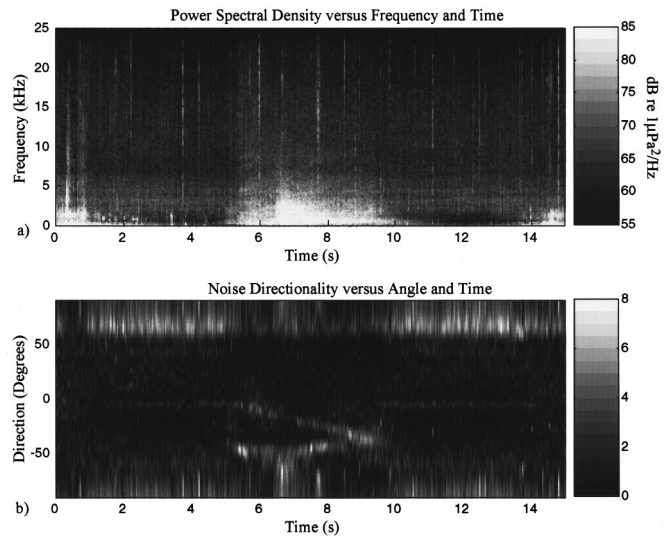


FIG. 9. An example of a breaking event that formed two converging hot-spots. (a) Ambient noise spectrogram. (b) Ambient noise directionality. Only a few events showing directionality patterns like this were observed in the data set. The origin of the faint peak around 0 degrees apparent before and after the breaking event is unknown.

positive to negative angles). As the breaking crest moves to the shoreward side of the array, the directionality shows a more well-defined peak.

## II. THE INTERPRETATION OF ACOUSTIC HOT-SPOTS

The directional density estimates of breaking wave noise show that sound radiated by a relatively small segment of the breaking wavecrest dominates the noise field at the coherence array. An estimate of the length scale of the dominant region can be made as follows. Linear, shallow water wave theory predicts the wave speed to be  $v_g = \sqrt{gh}$ , where  $g$  is the acceleration due to gravity and  $h$  is the water depth. This is only a crude expression for the wave motion, but is sufficient for our present purpose. Setting  $h = 2.2$  m yields a wave

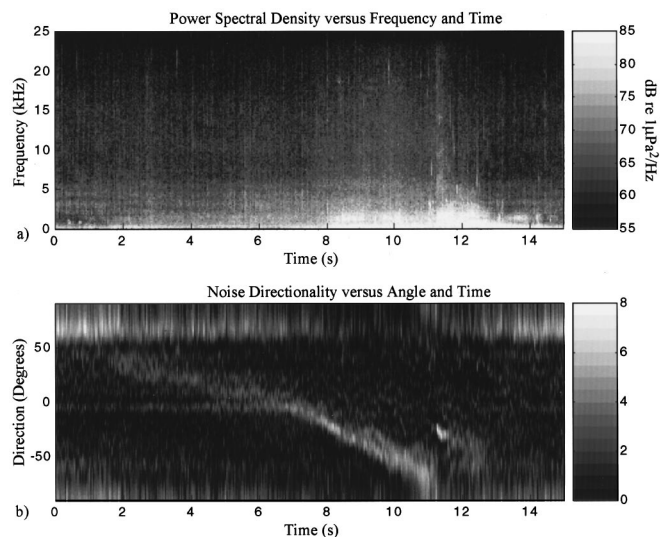


FIG. 10. An example of a breaking event that first broke on the seaward side of the array and then passed on the southern side of the array on its shoreward journey. (a) Ambient noise spectrogram. (b) Ambient noise directionality.

speed of 4.6 m/s. The time interval between the passage of the wave over the array to the middle of the breaking event can be estimated from the ambient noise spectrogram, and is approximately 5 s. Thus the wave was roughly 23 m shoreward of the array by the middle of the acoustically active part of the breaking event. The directionality at this time shows a narrow, 9 degree band of energy. This beam width corresponds to the maximum resolution of the array, so the actual beam may have been narrower than 9 degrees. At a range of 23 m, the 9 degree beam has a 3.6-m footprint, which provides an upper limit for the length scale of the hot-spot. This length is considerably less than the total length of spilling/overturning region, which was 20 m or more at this time.

The conclusion is that only a fraction of the total length of breaking wave crest contributes to the noise field observed behind the wave. On the face of it, there are at least two equally reasonable explanations for this, which are that the phenomenon is an effect inherent to the source of noise (the breaking wavecrest), or the hot-spots are caused by a propagation effect. Consider the first hypothesis, that the hot-spots are a property of the source. The video recordings show that the hot-spots tend to track the trajectory of the wave break points, implying that the wave break points are more acoustically energetic than the rest of the breaking crest. Under this hypothesis, hot-spots should always be observed in pairs, one for each break point, and the hot-spot trajectories should diverge as the break points separate. The data set, however, shows very few hot-spot pairs, and those pairs that are observed tend to converge (see Fig. 9, for example) indicating that the acoustically active regions are merging. As will be shown, this behavior can be explained in terms of a screening effect but is not consistent with radiation from pairs of break points dominating the noise field because they are louder than the rest of the breaking crest.

The second hypothesis is that the phenomenon is caused by a propagation effect. The bubbly residue left behind the breaking wave is strongly absorbing and may be screening the seaward-propagating sound radiated by all of the breaking crest except for the wave break points, which radiate into relatively bubble-free water. Under this assumption, hot-spots are formed in pairs but they are not expected to be *observed* in pairs, at least not at a single array. Typically, the sound from one of the two break points will have a clear transmission path to the array while sound from the other will be screened by the bubbly residue. An exception is if the array is directly behind a symmetrically breaking wave, in which case both break points may have a clear transmission path to the array. However, our method of estimating horizontal directivity has a north-south ambiguity, mapping the hot-spots from such symmetrical events into a single peak.

The fact that hot-spots are occasionally observed in converging pairs is explained by noting that a single shoaling wave can initially overturn at multiple points along the wavefront. Each of these points develops into lengths of breaking crest which merge as the wave shoals. A wave with two regions of breaking crest will have two converging break points and two diverging break points. A suitably positioned array will receive energy from both the converging break

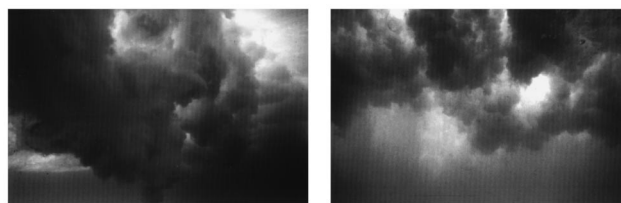


FIG. 11. Underwater photographs taken beneath breaking waves off La Jolla Shores Beach. The images illustrate the plumes of bubbles injected into the water column by breaking surf. The breaking crest is moving from right to left in the left hand figure, and directly into the page in the right hand figure.

points while sound from the outer, diverging break points is absorbed by the bubble screen. Without the screening effect, an array positioned behind a wave with two lengths of breaking region should observe four hot-spots, two of which are converging and two of which are diverging. The directionality patterns seen in Fig. 9 do not show four hot-spots, and thus favor the screening hypothesis.

Another implication of the screening hypothesis is that waves breaking on the seaward side of the array should not show hot-spots, since there is no bubble screen between the breaking wavecrest and the array. This prediction is largely consistent with the few events in the data set which were observed to break on the seaward side of the array (Fig. 10 shows an example of such an event). This aspect of the data set, however, should not be overemphasized as the interpretation of these events is not straightforward. The theory leading to the measurement of horizontal directivity in the Appendix shows that the angular resolution of the estimate depends on the highest frequency component of a given source heard at the array. The ambient noise spectrograms show an increase in the higher frequency components of wave noise as the breaking crest moves from the seaward to the shoreward side of the array, improving the angular resolution of the directivity estimates.

In the remainder of the paper, the screening hypothesis will be examined in greater detail. Section III details an experiment to measure the sound speed and absorption in the bubbly residue, and this is followed by a simple theory for the trajectories of the wave break points in Sec. IV. The results from these two sections are combined into a propagation model in Sec. V, which demonstrates the bubble screening effect.

### III. AN EXPERIMENT TO MEASURE SOUND SPEED AND ABSORPTION IN THE BUBBLY RESIDUE

The acoustical properties of the bubbly residue behind a breaking wave need to be quantified before the effect of the residue on the noise radiated by the breaking crest can be calculated. In this section, an experiment is described in which active acoustic measurements of the sound speed and absorption in the bubbly residue were made. The active acoustic experiment was performed at Camp Pendleton during times when there were waves actively breaking over the surf frame. To give an impression of the bubbly residue, Fig. 11 shows two photographs of the underside of breaking waves taken in a more convenient location, which was the

surf zone off La Jolla Shores Beach. The photographs illustrate the plumes of bubbles which comprise the major portion of the bubbly residue.

### A. The active acoustic experiment

The geometry for the active acoustic experiment is shown in Fig. 2. The measurements were made by exciting an acoustic source with a broadband signal and recording the response on two horizontally separated hydrophones. The seaward hydrophone was only centimeters from the source and acted as a reference. Travel time measurements were made by subtracting the time delays estimated from cross-correlations of the signal received at the seaward and shoreward hydrophones with a template signal. Similar travel time sonars have been used in earlier experiments by Lamarre and Melville,<sup>5-7</sup> and Vagle and Farmer<sup>8</sup> to study wave-induced bubble plumes.

In the interest of brevity, only a summary of the signal processing scheme for the travel time sonar is presented here. The signal used to excite the acoustic source was recorded synchronously with the response from the hydrophones providing a time reference for subsequent pulse compression. The excitation signal was a 64 bit, pseudo-random noise sequence composed of 180° phase-modulated sine cycles (1 cycle per bit) centered on 10 kHz. The result was a broadband ping with significant frequency components between 4 kHz and 20 kHz. The spectral components of the signal were adjusted prior to transmission to compensate for the frequency response of the source. Each ping lasted 6.4 ms, and 100 pings were transmitted every second. The resulting data stream was sampled at 50 kHz, providing approximately 5 points across a correlation peak after pulse compression. The correlation peaks were interpolated by comparison with a theoretical correlation function to improve the temporal resolution.

### B. Measurements of sound speed reductions and increased absorption

Figure 12 shows the auto-correlation of the source excitation signal and the cross-correlation of the excitation signal with the signals recorded at the reference and measurement hydrophones for a ping transmitted through bubble-free water. The two cross-correlation functions show a clear peak corresponding to the arrival time of the ping at the reference and measurement hydrophone locations. The direct-path arrival can be seen just after 0.3 ms on the measurement hydrophone cross-correlation. The smaller, negative peak at 1.4 ms corresponds to the surface reflected arrival. The time interval between the direct-path peaks on the reference and measurement cross-correlations is a measure of the propagation time of the ping between the two hydrophones. The sound speed of the medium between the hydrophones (water or bubbly mixture) is estimated by dividing the separation between the hydrophones by the propagation time. The sound speed estimates presented below were averages over 50 pings to reduce noise and filter out variations occurring on time scales of less than 0.5 s.

Figure 13(a) shows a 20-min record of the sound speed at a depth of 70 m below the mean surface during a period

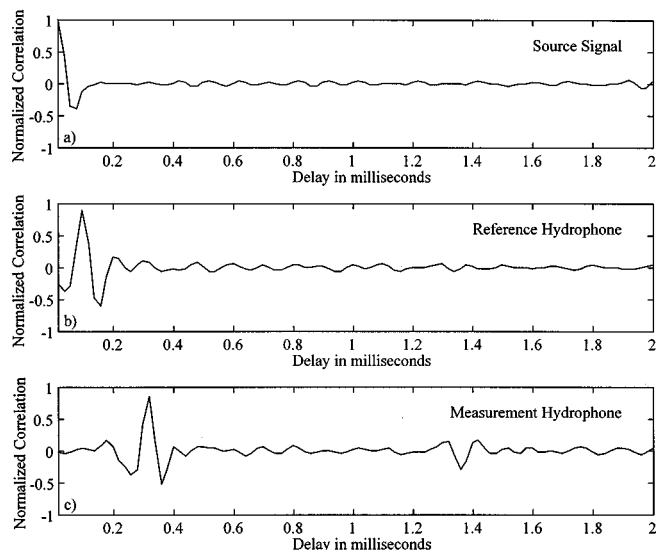


FIG. 12. Cross-correlation functions showing the acoustic arrivals at the reference and measurement hydrophones. (a) Auto-correlation function of the source excitation. (b) Cross-correlation of the source excitation with the reference hydrophone. (c) Cross-correlation of the source excitation with the measurement hydrophone.

when waves were breaking intermittently over the deployment frame. The mean water depth was approximately 1.7 m. Large, episodic reductions in sound speed can be seen in the record. The video recordings show that each of the reductions occurred a few seconds after the passage of a breaking wave over the frame. The reduction in sound speed caused by the wave-induced bubble plumes is significant, approaching 50% for the two events shown. As the bubbles entrained in the water column diffused, degassed, and possibly dissolved, the sound speed gradually recovered to its

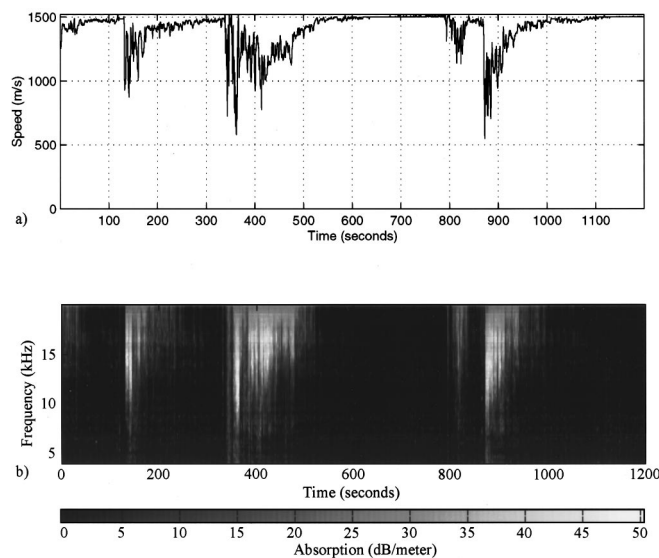


FIG. 13. The acoustic properties of the water column 70 cm below the surface during a series of breaking events over the travel time sonar. (a) Sound speed versus time. The reductions in sound speed that occur just after 100 s, 300 s, and 800 s in the record are associated with the passage of breaking waves over the sound speed sonar. (b) Acoustic absorption versus frequency and time. Large increases in acoustic absorption occur simultaneously with the reductions in sound speed.

bubble-free value of just under 1500 m/s. Note that the time scale for this recovery is about 200 s, which is roughly two orders of magnitude longer than the duration of the noise accompanying wave breaking.

Figure 13(b) shows the acoustic absorption over the same path and during the same time as the sound speed measurements. The absorption was estimated by dividing the spectrum of the sequence of pings arriving at the measurement hydrophone by a reference spectrum recorded in bubble-free water. The absorption measurements are limited to frequencies at which the acoustic source radiates efficiently, which was 4 kHz to 20 kHz. The white regions in the plot correspond to periods of strong absorption, approaching 50 dB per meter, and are strongly correlated with the periods of sound speed reduction.

The data show that after the passage of a breaking wave, the water column is filled with acoustically absorbing bubbles which penetrate to at least half the mean water depth. The increased absorption and reduced sound speed caused by the bubbles persist for roughly two orders of magnitude longer than the generation of the wave noise. The effects are most pronounced immediately after the injection of the bubbles, with sound speed reductions approaching 500 m/s and absorptions of up to 50 dB per meter. Large regions of bubbly water with these properties have a pronounced effect on sound propagation and need to be included in models of surf noise.

#### IV. A SIMPLE THEORY FOR THE TRAJECTORIES OF WAVE BREAK POINTS

The source of bubbly water behind a breaking wave is obvious. Spilling, plunging, collapsing, and surging waves all create air bubbles in the water column as they overturn and break, although the small-scale physics of the air entrainment processes remains largely a subject of speculation.<sup>9</sup> Our concern here is with the large-scale spatial distribution of the bubbly water behind a breaking wave, and also the trajectories of the wave break points which define the edges of the bubbly region. The goal is to develop a simple model for the velocity of the break points which delineate the region of breaking crest from unbroken crest.

##### A. The formation of wave break points

The fluid dynamics of shoaling and breaking waves is a complicated process that has been the subject of study for over 100 years.<sup>10,11</sup> The nonlinear fluid dynamics leading to wave breaking are now largely understood, but the breaking process itself remains an area of active investigation. Gravity waves incident on very steep beaches may be reflected without breaking, but for the beach slopes of interest here, essentially all of the incident wave energy is dissipated by breaking. As a wave shoals on a beach, it evolves through a series of phases. Synolakis and Skjelbreia<sup>12</sup> describe spatial zones of gradual shoaling leading to rapid shoaling until the wave overturns. After overturning, there is a zone of rapid and then gradual decay. These last two zones were identified in an earlier work by Svendsen *et al.*<sup>13</sup> as inner surf and outer surf regions. There is no single description of wave evolution

through these phases, and the literature tends to treat the dynamical behavior of pre-breaking and post-breaking waves as separate problems. A full mathematical description of wave breaking is not available at this time, but a recent paper by Grilli *et al.*<sup>14</sup> shows that the behavior of shoaling solitary waves in the laboratory can be accurately reproduced by a nonlinear numerical model, up to the point of wave breaking.

Most of the extensive literature on breaking wave dynamics is concerned with waves that have a constant amplitude along their crest; that is, waves that do not show any height variation in a direction normal to that of their motion. Casual observation shows, however, that waves shoaling on beaches frequently exhibit variations in wave height along the wavecrest. This is not surprising as the instantaneous surface elevation is, in general, the result of a superposition of different wave number and frequency components,<sup>15,16</sup> which cause spatial interference patterns. Wave height variations have a significant effect on breaking and the subsequent patterns of air entrainment, and thus need to be included in a model connecting the breaking process with wave noise directionality. What is commonly discussed are the effects of variable bathymetry. The differences between the effects of variable bathymetry and variable wave height are discussed in Sec. IV F.

The breaking process begins when the velocity of the water at the top of the crest exceeds the mean speed of the wavefront, and this occurs when the height of the wavecrest is some fraction of the mean water depth. If the wave height varies along the wavecrest, say a single maximum falling off to either side, then the wave initially breaks at a point along the crest corresponding to the highest point of the wavecrest reaching its breaking depth. As the wave continues to shoal, the single break point grows into a line of breaking wavecrest delimited by a break point at each end. Figure 1 shows an example of this process.

##### B. Coordinate frames and variables

The analysis which follows is based on two coordinate frames, one that is stationary with respect to the sea floor, which will be called the earth frame, and another that moves with shoaling wave front, which will be called the wave frame. For clarity, wave frame variables will always be denoted by an overbar. In addition, the value of variables evaluated at the wave break point will be subscripted by *b*. For example, the variable  $\bar{s}$  denotes distance along the wavecrest in the wave frame, and  $\bar{s}_b(t)$  is the distance along the wavecrest to the wave break point(s) at time *t*. The origin for *s* is the point on the wavecrest that initially broke. The other wave frame variables are the wave height  $\bar{\eta}(\bar{s}, t)$  and the still water depth  $\bar{h}(\bar{s}, t)$ . The earth frame variables are *x* and *y*, which, respectively, correspond to across-shore and along-shore distances. The origin for *x* and *y* is the point at which the wavecrest initially broke. See Fig. 14 for an illustration of the earth and wave frame variables. For a plane wavefront incident to the shore line at an angle  $\alpha$ , the coordinate transformations between the position variables in the two frames are given by

$$x = ct \cos \alpha + s \sin \alpha, \quad y = s \cos \alpha - ct \sin \alpha, \quad (1)$$

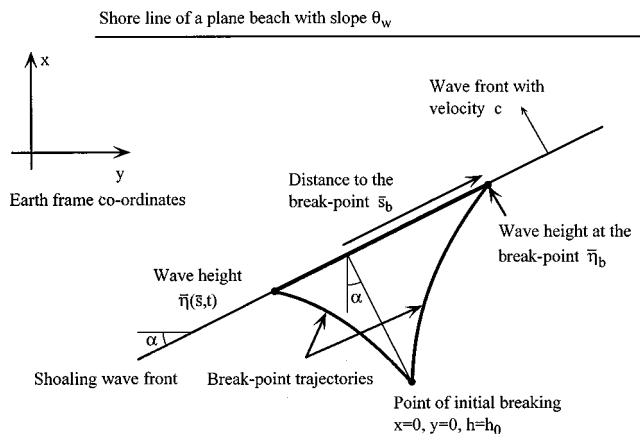


FIG. 14. The earth and wave frame variables. The earth frame is fixed relative to the sea floor and the wave frame moves with the (planar) shoaling wavefront.

where  $t$  is time after the wave initially broke and  $c$  as the wave speed normal to the wave front. The water depth at which the wave first breaks is denoted by  $h_0$ .

### C. The McCowan breaking criterion

The central assumption underlying our model for the break point velocity is the McCowan breaking criterion,<sup>17</sup> which states that the height of a wavecrest at incipient breaking is some constant fraction of the water depth at which breaking occurs. If  $\bar{\eta}_b$  is the height of the wavecrest just before overturning and  $\bar{h}_b$  is the still water depth at which breaking occurs, then the McCowan breaking criterion is

$$\bar{\eta}_b = K_m \bar{h}_b, \quad (2)$$

where  $K_m$  is a constant of order unity. Much of what follows is predicated on the validity of Eq. (2). McCowan originally determined  $K_m$  to be 0.78, but subsequent studies have shown that the value lies in a range of 0.7–1.5, depending on beach slope and wave period.<sup>18</sup> The analysis that follows also assumes that both wave refraction and variations in wave speed along a wavefront can be neglected for the acoustically active duration of a wave breaking event, which is a few seconds. A simple, linear analysis shows that wave refraction can be neglected (e.g., a planar wavefront remains planar) on length scales along the wavecrest of about 50 m provided that the incident angle of the wave to the beach is small and the beach slope is less than about 1:25. Although significant water depth variations can occur along a wavecrest incident at an angle to the beach, the numerical study presented by Grilli *et al.*<sup>14</sup> goes some way toward justifying the assumption of a constant wave speed.

### D. The trajectory of wave break points

Consider a wavecrest shoaling at an angle  $\alpha$  to the shore line of a planar beach with slope  $\theta_w$ . The height of the wavecrest  $\bar{\eta}(\bar{s})$  is assumed to vary with distance  $\bar{s}$  along the wavecrest, which is assumed to be the cause of the wave break points and their movement through the surf zone. A discussion of the role that variable bathymetry plays in the formation of break points can be found in Sec. IV F—it is

sufficient for our present purpose to assume that the beach is plane. In the wave frame, an expression for the still water depth at the wave break point is

$$\bar{h}_b(\bar{s}_b, t) = h_0 - ct \cos \alpha \tan \theta_w - \bar{s}_b(t) \sin \alpha \tan \theta_w, \quad (3)$$

where the wave frame variables are defined in Sec. IV B. The penultimate term in Eq. (3) accounts for shoreward movement of the wave past its point of initial overturning, and the last term accounts for variations in still water depth with position along the wavecrest. Substituting the McCowan breaking criterion into Eq. (3) yields an expression for the wave height at the break point:

$$\bar{\eta}_b(\bar{s}_b, t) = K_m (h_0 - ct \cos \alpha \tan \theta_w - \bar{s}_b(t) \sin \alpha \tan \theta_w). \quad (4)$$

Equation (4) defines a relationship between the wave height at the break point, time after initial overturning, and distance along the wavecrest to the break point, which is itself a function of time. Our goal is to determine the relationship between  $t$  and  $\bar{s}_b$ , and an additional constraint to Eq. (4) is required in order to achieve it. The additional constraint comes from recognizing that the wave height profile is actually a free variable, determined by the frequency and wave number spectrum of the incident wave field and the dynamics of wave shoaling leading up to breaking. Although it would be possible, in principle, to compute the wave height profile for a given bathymetry and incident wave field, the wave height variation will be simply parameterized directly by the two-parameter set of exponential profiles:

$$\bar{\eta}_b = K_m h_0 \exp(-|\bar{s}_b/s_0|^\gamma), \quad (5)$$

where  $s_0$  is a length scale characterizing the change of height along the wavecrest and  $\gamma$  is a dimensionless parameter of order unity that will be called the profile index. Equation (5) provides a useful representation for the wave height variation but does not have any theoretical significance. In the absence of a comprehensive data set, a family of exponential profiles is a reasonable choice, and results in break point trajectories which are qualitatively similar to those observed in the surf zone. In any case, our general conclusions concerning the interaction of the bubbly residue with the noise radiated by the breaking wavecrest do not depend on a particular functional dependence chosen for the wave height profile. Equations (4) and (5) together define a transcendental equation relating  $\bar{s}_b$  and  $t$ :

$$h_0 (1 - \exp(-|\bar{s}_b/s_0|^\gamma)) - ct \cos \alpha \tan \theta_w - \bar{s}_b \sin \alpha \tan \theta_w = 0. \quad (6)$$

In general, Eq. (6) needs to be solved numerically to determine  $\bar{s}_b$  as a function of time for a given beach slope, angle of incidence, profile index, and initial breaking depth. Closed form solutions can, however, be obtained in two limiting cases. The first of these is for waves of constant height, corresponding to the limit of  $s_0 \rightarrow \infty$  in Eq. (5). In this case, Eq. (6) simplifies to

$$\bar{s}_b = -ct \cot \alpha, \quad (7)$$

which shows that in the case of a constant wave height there is a single break point which moves along the wavecrest in

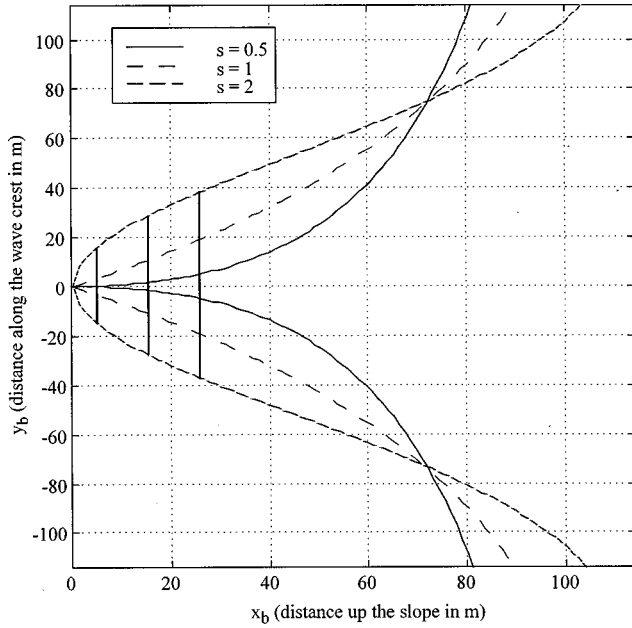


FIG. 15. Spatial extent of the bubbly region generated by breaking surf normally incident in a plane beach calculated assuming an exponential variation in wave height profile. The lines track the break point loci for waves with different profile indices.

the direction of negative  $\bar{s}_b$ . Applying the coordinate transformations defined in Eq. (1) to Eq. (7) results in equations for the trajectory of the break point in the earth frame:

$$x_b = 0, \quad (8a)$$

$$y_b = \frac{-ct}{\sin \alpha} (1 - 2 \sin^2 \alpha). \quad (8b)$$

Equation (8a) is intuitively correct since, by assumption, a wave of constant height always breaks at the same depth, requiring  $x_b$  to be constant. For waves normally incident on the beach,  $\alpha \rightarrow 0$ , and from Eq. (8b),  $y_b \rightarrow \infty$ . Physically, this limit corresponds to the wave breaking along its entire length simultaneously.

The second limiting case is for waves normally incident on the beach, so that  $\sin \alpha = 0$ ,  $\cos \alpha = 1$ , and Eq. (6) can be rearranged to yield

$$\bar{s}_b = \pm \left\{ \ln \left( \frac{h_0}{h_0 - ct \tan \theta_w} \right) \right\}^{1/\gamma}. \quad (9)$$

The  $\pm$  sign arises in Eq. (9) because there are breakpoints located on either side of  $\bar{s}_b = 0$ . For  $\bar{s}_b$  to be real, Eq. (9) requires that  $0 < ct \tan \theta_w / h_0 < 1$ , which amounts to requiring that time be positive and less than the value that would take the wave onto the shore.

Equation (6) is the main result of this section. Together with the coordinate system transformations, it defines the trajectory of the break points through the surf zone. As the break point trajectories define the edges of the bubbly residue, it also allows the spatial extent and shape of the residue behind the wave to be calculated.

Figure 15 shows the spatial extent of the bubbly region generated by waves with profile indices of  $\gamma = 0.5, 1$ , and  $2$ . The vertical lines in the figure are discussed in Sec. VB. A

scaling length of  $s_0 = 75$  m and an initial breaking depth of  $h_0 = 2$  m were used for all the profiles, and the wave speed was rounded up to 5 m/s. The profile index  $\gamma = 2$  corresponds to a nearly uniform wave height across the greater part of the wavecrest, which results in an initially rapid expansion of the breaking crest length. The wave height corresponding to a profile index of 0.5 varies rapidly across the wavecrest, and gives rise to a slowly spreading breaking region.

## E. The velocity of wave break points

Although an explicit expression for the velocity of the break points is not required for the acoustical analysis in Sec. V, some physical insight into break point behavior can be gained from the result, which does not require the assumption of a wave height profile. Rearrangement of Eq. (4) allows for an explicit solution of the time at which a region of wavecrest breaks as a function of wave height and distance along the wavecrest:

$$t = \left( \frac{h_0 - K_m \bar{\eta}_b(\bar{s}_b, t)}{\tan \theta_w} - \bar{s}_b(t) \sin \alpha \right) \frac{1}{c \cos \alpha}. \quad (10)$$

The velocity of the break points is found by differentiating Eq. (10) with respect to  $\bar{s}_b$ :

$$\frac{d\bar{s}_b}{dt} = \left( \frac{dt}{d\bar{s}_b} \right)^{-1} = -c \cos \alpha \left/ \left( \frac{K_m}{\tan \theta_w} \frac{\partial \bar{\eta}_b}{\partial \bar{s}_b} + \sin \alpha \right) \right. \quad (11)$$

For the special case of waves incident normally to the shore line, the break point velocity is proportional to the wave speed and inversely proportional to the rate of change of wave height along the wavecrest. A normally incident wave with uniform height breaks simultaneously along the entire crest, and the break point velocity becomes infinite. Conversely, waves with rapid variation in height along the wavecrest have break points which move slowly.

## F. The effect of variable bathymetry

The main results of this section have been derived under the assumption that a plane wave is shoaling on a plane beach, and that wave break points are formed because of variations in wave height along a wavecrest. This is a convenient assumption as it allowed the formulation of calculable expressions for the trajectory of wave break points. Clearly, few real beaches are plane and so the question of how the analysis can be modified to accommodate this fact is important. For any real beach, the bathymetry  $h(x, y)$  can always be expressed as

$$h(x, y) = h_0 - x \tan \theta_w + \epsilon(x, y), \quad (12)$$

where the notation of earth frame variables has been retained and  $\epsilon(x, y)$  is a function that represents the deviation of the real beach bathymetry from a plane beach. In wave coordinates, Eq. (12) becomes

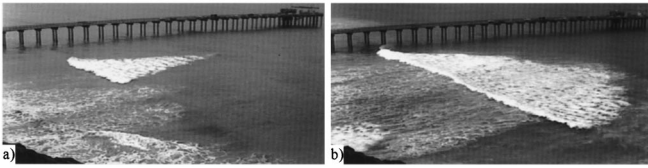


FIG. 16. The foam patterns generated by breaking surf off La Jolla Shores Beach. The foam indicates the extent of the air–water mixture behind the breaking wave crest. (a) A narrow, wedgelike pattern indicating a profile index less than unity. (b) A pattern showing an initially rapid and then more gradual growth, consistent with a profile index greater than unity.

$$\begin{aligned} \bar{h}_b(\bar{s}_b, t) = & h_0 - ct \cos \alpha \tan \theta_w - \bar{s}_b \sin \alpha \tan \theta_w \\ & + \epsilon(ct \cos \alpha + \bar{s}_b \sin \alpha, \bar{s}_b \cos \alpha - ct \sin \alpha), \end{aligned} \quad (13)$$

Eq. (13) then replaces Eq. (3) in the break point trajectory analysis of Sec. IV D. The net result is a more complicated transcendental equation which replaces Eq. (6) and which must also be solved numerically.

In addition to the question of the analysis, there is a more interesting physical question concerning the relative importance of variations in bathymetry versus spatial interference structure in determining break point behavior. It might seem from Eq. (13) that they can be considered as essentially equivalent, but there are, in fact, important physical differences. If the detailed behavior of break points is determined primarily by bathymetry alone, then one would expect to see very similar patterns of wave breaking over the same region of sea floor, at least over time scales on which the bathymetry is stable. If, however, the primary causative agent is spatial interference structure, itself caused by a superposition of incident wave number components and the refraction of those components as they shoal, then the wave breaking patterns can be expected to be quite variable both over short time scales and over the same region of sea floor.

It seems likely that both bathymetry and variations in wave height play a role in determining the size and shape of the bubbly regions formed behind shoaling waves. At any rate, there is clear evidence that it is not variations in bathymetry alone that are responsible. Figure 16 shows two breaking waves with very different patterns of breaking observed at the same location and within 10 min of each other. It is difficult to imagine how the differences could be attributed to changes in bathymetry alone over such a short time scale, leading to the conclusion that the kind of variations in wave height considered here are important.

### G. A discussion of the break point model

The physics of break point generation is more complicated than our simple model suggests. A realistic surf model would include a wave train with multiple wave frequencies and a wave number spectrum. Smith and Vicent<sup>19</sup> have shown that two wave trains with different frequencies interact, with the higher frequency peak decaying more rapidly than it would in isolation. Wind also affects breaking surf, as discussed by Douglas.<sup>20</sup> He studied the influence of wind on near shore breaking waves in a laboratory flume. Douglas reported that breaker location, geometry, and type depended

strongly on wind direction. Onshore winds caused waves to break in deeper water and to spill while offshore winds caused waves to break in shallower water and to plunge. A final point to note is that the McCowan breaking criterion is being applied to individual waves with a variable wave height. Even assuming that this criterion can be applied to single waves, it is not altogether clear that the empirical law holds when breaking is well under way and the break point delineates unbroken crest from breaking crest. Also note that the break point analysis depends on the assumption of negligible wave refraction and constant wave speed during breaking.

Notwithstanding the issues raised above, a beginning needs to be made somewhere, and the results of the break point analysis presented here will be used to study the screening effect of the bubbly region on the noise radiated by the breaking wavecrest. In addition, there is some evidence for the large-scale patterns of air entrainment suggested by the model and family of exponential profiles. The foam shapes formed by the breaking waves in Fig. 16 are qualitatively similar to those plotted from the theory in Fig. 15. The narrow, wedgelike pattern in Fig. 16(a) developed from a slowly spreading breaking region, and the boundaries of the region have a concave shape similar to those obtained from the theory with a profile index of less than unity. The convex boundary in Fig. 16(b) was generated by a breaking crest which initially grew rapidly in length, which is consistent with a profile index greater than unity.

## V. A MODEL FOR THE INTERACTION BETWEEN WAVE NOISE AND THE BUBBLY RESIDUE

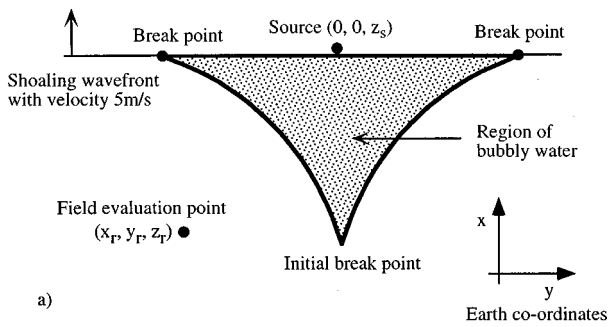
In this section, a propagation model for the noise radiated by the breaking wavecrest is presented that accounts for the screening effect of the region of bubbly water left behind the wave.

### A. A normal mode, range-dependent propagation model

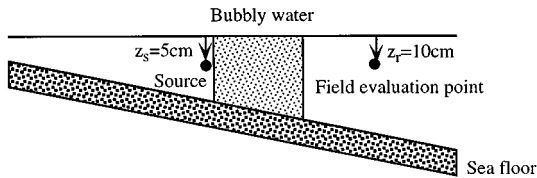
The geometry for the propagation calculations is shown in Fig. 17. Figure 17(a) shows a plan view of a shoaling wave normally incident on a plane beach. The wave height along the wavecrest is assumed to vary in the manner described in Sec. IV, giving rise to the region of bubbly water behind the wave. A receiver is located on the seaward side of the wave at the position  $(x_r, y_r, z_r)$  where  $x$  and  $y$  define position in the horizontal plane and  $z$  is depth measured down from the still water surface. Figure 17(b) shows a side view of the water column that transects the bubbly region.

The model depicted in Fig. 17 includes several approximations. The first of these is that the sound speed and absorption of the bubbly water left behind the wave is assumed to be independent of position, and remain constant while the wave is breaking and radiating noise. The active acoustic experiment described in Sec. III showed that the bubble plumes left behind the breaking wave absorb sound and reduce the sound speed for about 200 s after the passage of the wave. The acoustically active phase of the breaking event lasts for a few seconds, or approximately two orders of mag-





a)



b)

Material	Sound Speed (m/s)	Density (Kg/m <sup>3</sup> )	Absorption (dB/wavelength)
Sea floor	1700	1500	0.5
Seawater	1500	1000	0
Bubbly water	1000	1000	2

FIG. 17. Geometry for the Green's function calculation. (a) Top view. (b) Side view.

nitude less than the time scale for the decay of the bubbly residue. Given the disparity between the time scales, the assumption that the plumes are in frozen equilibrium is reasonable.

The second approximation is that the plumes fill the entire vertical extent of the water column, as shown in Fig. 17(b). The bubble plume properties reported in Sec. III were measured in the middle of the water column, where absorptions as high as 50 dB per meter were recorded. The acoustic calculations which follow are based on very mild values of absorption by comparison, a value of 2 dB per meter being sufficient to demonstrate the screening effect. Given the much higher values of absorption observed at mid-water depth, it is not unreasonable to assume that sufficient numbers of bubbles will be found throughout the entire water column to provide the assumed absorption.

The acoustic field due to an harmonic, point source of angular frequency  $\omega$  in the range-dependent wave guide can be calculated with an adiabatic mode expansion of the Green's function solution of the Helmholtz equation. The adiabatic mode expansion can be written as

$$p_s = i\sqrt{2\pi} \sum_{n=1}^{\infty} \frac{E_n N_n(x_r)}{\sqrt{k_n r_r}} \sin(\nu_n z_r) \exp\left(i \int_0^{r_r} k_n dr\right), \quad (14)$$

where the path for the integral runs in a straight line connecting the source and receiver,  $i = \sqrt{-1}$ ,  $r_r = \sqrt{x_r^2 + y_r^2}$ , and the other symbols are defined below. Equation (14) is a two-dimensional solution to what is strictly a three-dimensional propagation problem. It does, however, yield an accurate solution for the acoustic field for the shallow beach slopes and

short ranges of interest here. The terms in the sum on the right hand side of Eq. (14) correspond to normal modes which resonate between the ocean's surface and the sea floor, and propagate along radials away from the source. A given mode  $n$  has amplitude  $N_n$ , horizontal wave number  $k_n$  and vertical wave number  $\nu_n$  which depend on the local water depth and acoustical properties of the water column and sea floor. The excitation amplitude of mode  $n$  at the source is

$$E_n = N_n(0) \sin(\nu_n(0)z_s). \quad (15)$$

Although an infinite number of modes have been included in the sum, in practice only a finite number of modes,  $N_{\max}$ , propagate through the water column for any given frequency. Mode of higher order are radially evanescent in the water column and do not contribute to the acoustic field more than a few wavelengths from the source.

Zhang and Tindle's complex effective depth method<sup>21</sup> has been used to compute the mode eigenvalues and amplitudes. The mode eigenvalues are the solutions of the dispersion equation

$$2\nu_n(k)h(x) - i \ln(V(k_n)) - \pi = 2(n-1)\pi, \quad (16)$$

where

$$\nu_n = (\omega^2/c_w(x,y)^2 - k_n^2)^{1/2} \quad (17)$$

is the vertical wave number of mode  $n$ ,  $c_w(x,y)$  is the sound speed in the water column,  $h(x)$  is the local water depth, and  $V$  is the plane wave reflection coefficient for the water column-sea floor interface, given by

$$V(k) = \frac{\rho_s \nu - i \rho_w \xi}{\rho_s \nu + i \rho_w \xi}, \quad (18)$$

where  $\rho_w$  and  $\rho_s$ , respectively, are the density of the water column and the sea floor,  $\xi$  is the vertical wave number in the sea floor given by

$$\xi = (k^2 - \omega^2/c_s^2)^{1/2}, \quad (19)$$

and  $c_s$  is the sound speed in the sea floor. The mode amplitudes are given by

$$N_n^{-2} = \frac{1}{2} \left[ H + \frac{\pi - i \ln(V)}{2\nu_n} \right]. \quad (20)$$

The effect of the air-laden water behind the wave is to decrease the sound speed and increase sound absorption. These effects are incorporated into the acoustic model by making  $c_w$  in Eq. (17) complex. Strictly speaking, the discontinuity in sound speed in range introduced by this step invalidates the assumption of adiabatic propagation, which is that the environmental parameters, such as water depth and sound speed, vary slowly in range. In practice, this means that the effects of mode coupling and scattering at the boundaries of the bubbly region are not accounted for. These effects are difficult to model without a detailed knowledge of the fine-scale spatial distribution of the bubbles behind the wave, which are not available. In any case, the effects of mode coupling and scattering would not alter the main properties of the screening phenomenon of interest here.

## B. The formation of acoustic shadows behind the bubbly residue

The geometry used to calculate the effect of the bubbly region on breaking waves noise is shown in Fig. 17. A point source representing a small length of breaking wavecrest is placed just in front of a region of absorbing bubbles and a receiver is assumed to lie in a region of bubble-free water. The acoustic field was computed in a plane 10 cm below the ocean's surface using Eq. (14). The choice of receiver depth is arbitrary and has no real effect on the results. The field was computed for three times after the initial wave break using waves with two profiles. Two source locations were used for each wave profile and time to simulate radiation from break points, near the edge of the breaking crest, and a region of breaking crest mid-way between the two break points. The modal expansion was calculated along radials from the source. The values of sound speed, density, and acoustic absorption for the sea floor, sea water, and bubbly residue used in the calculations are given in Fig. 17. The density of the bubbly region was taken to be that of seawater on the grounds that the void fraction of air is low, typically around 5% or less, away from the breaking crest and does not significantly alter the mixture density. The values used for the sea floor are typical for a medium grained sand. The value of absorption used for the bubbly residue, 2 dB per meter, is considerably lower than the typical mid-water depth values observed during the active acoustic experiment. This was a deliberate choice to demonstrate that the screening effect is quite robust, and does not rely on particularly high values of absorption. The beach slope was set to 1 degree, and the source depth was 5 cm. The source frequency was set to 2 kHz, which was chosen to be well above the modal cutoff for the waveguide at any of the source locations.

The spatial extent of the bubbly residue was computed using Eq. (6). Two sets of results are presented for the two choices of profile index  $\gamma=0.5$  and  $\gamma=2$ . A scaling length  $s_0=75$  m and initial breaking depth  $h_0=2$  were used for both profiles and the wave speed rounded up to 5 m/s. These are the same parameters as those chosen for Fig. 15, which shows the break point trajectories and spatial extent of the bubbly region in the horizontal plane predicted by the break point model. The three vertical lines in the figure correspond to the along-shore location of the breaking crest at the times chosen for the field evaluations. The extent of the wavecrest at each time and profile index is the length of the vertical line which lies between the lower and upper break point trajectories.

The results of the calculations are shown in Figs. 18 and 19. Each figure contains six Green's function calculations corresponding to the three times and two source locations chosen. The times after wave breaking are 1 s, 3 s, and 5 s, which, respectively, correspond to the left, middle, and right plots. The Green's function is shown as a gray scale contour map in dB referenced to a free-space source strength of unity at 1 m. The dark regions in the figures correspond to areas of low sound intensity. The shore line runs vertically, with the deepest water on the left hand side of the plots. The source is positioned in the middle of each plot.

The most obvious feature of the plots is the region of

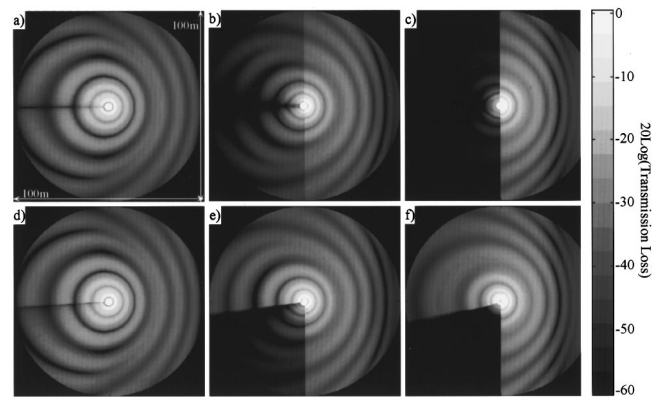


FIG. 18. Green's function calculations for the wedge geometry including the effect of the evolving air–water mixture behind a breaking wave, travelling to the right, with a profile index of 0.5. The other geometry parameters are defined in the text. (a) through (c) correspond to 1, 3, and 5 s after the initial wave break. The source is located mid-way between the ends of the breaking wavecrest. (d) through (f) correspond to 1, 3, and 5 s after the initial wave break and a source located at the uppermost edge of the breaking wavecrest.

low intensity that forms behind the breaking crest as the bubbly residue develops. Consider the sequence of plots 18(a) through (c). The source is located mid-way between the two ends of the breaking wavecrest and the profile index is 0.5. From Fig. 15, the width of the bubbly region after 1 s is less than 1 m and very little acoustic absorption occurs. After 3 s, the bubbly region is 1.4 m wide, which is sufficient to cause a 30 dB or so drop in the acoustic intensity directly behind the breaking region. After 5 s, the bubbly region width has increased to 4 m and very little sound penetrates through to the seaward water column. Plots 18(d) through (f) show the acoustic field for a source located on one end of the breaking crest. The screening effect of the bubbly region on this source is quite different—the upper, left quadrant of the water column behind the wave is insonified. The overall effect of the bubbly residue is to mask seaward propagating

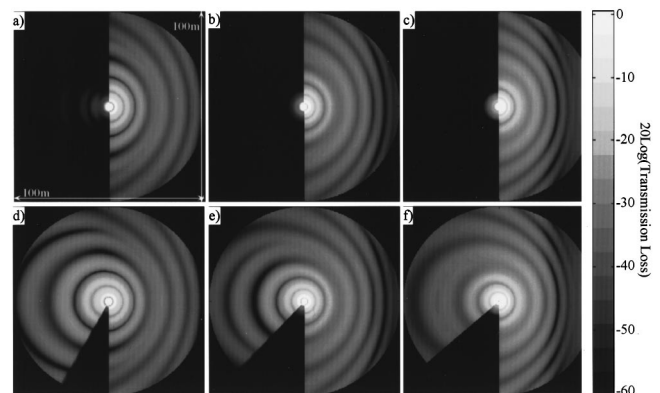


FIG. 19. Green's function calculations for the wedge geometry including the effect of the evolving air–water mixture behind the breaking wave, traveling to the right, with a profile index of 2. The other geometry parameters are the same as for Fig. 18, and are defined in the text. (a) through (c) correspond to 1, 3, and 5 s after the initial wave break. The source is located midway between the ends of the breaking wavecrest. (d) through (f) correspond to 1, 3, and 5 s after the initial wave break and a source located at the uppermost edge of the breaking wavecrest.

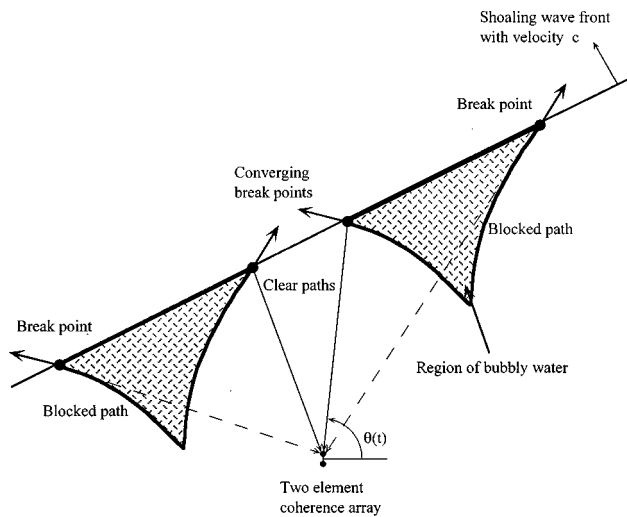


FIG. 20. An interpretation of the acoustic hot-spots in terms of the interaction of the noise radiated by the breaking wavecrest and the acoustic shadow cast by the absorbing, bubbly residue behind the wave. A shoaling wave with two regions of breaking wavecrest are shown, giving rise to two converging and two diverging break points. The noise from the diverging break points is absorbed by the bubbly residue and only the converging break points appear at the coherence array as acoustic hot-spots.

sound from all sources along the wavecrest except for those at the wave break points.

Figure 19 shows the same screening effect for a profile index of 2. The spatial extent of the bubbly region grows more rapidly during the initial stages of wave breaking than the  $\gamma=0.5$  profile, but the overall masking effect is similar.

It is clear from these results that under most circumstances only a small fraction of the breaking wavecrest at the edge of the breaking region—the wave break point—dominates the wave noise radiating to the seaward side of the wave. The exceptions are an observer placed directly behind the point at which a symmetrical wave broke, who would perceive both break points, and a wave with a flat profile, which initially breaks along an extended region of wavecrest. In the latter case, the entire region of breaking crest will initially be audible, fading to just the break points as the bubbly residue develops.

Some of the more subtle features of the acoustic field plots are worth noting in passing. The light and dark bands evident in the acoustic field are due to the constructive and destructive interference of the propagating acoustic modes. As the modes propagate into shallow water, they eventually reach their cutoff depth, after which they are strongly attenuated. This effect can be seen in Fig. 18(c), for example. Only the fundamental mode is left propagating on the far right hand side of the plot, as indicated by the absence of interference bands.

## VI. THE CONNECTION BETWEEN BREAKING NOISE, THE BUBBLY RESIDUE, AND ACOUSTIC HOT-SPOTS

Our interpretation of the acoustic hot-spot phenomenon is shown in Fig. 20. A wave normally incident to the shore line of a plane beach is shown moving from right to left across the figure. The wave is assumed to vary in height

along its crest and to have two peaks in its wave height profile. As the crest breaks, it injects plumes of bubbles, much like those illustrated in Fig. 8, into two triangular regions the water column. When the plumes are first created, they are acoustically active and radiate sound into the water column along the entire length of the breaking crest. Not all of the sound, however, is audible on the seaward side of the wave. The bubbly water left behind the wave is a strong sound absorber that casts an acoustic shadow, screening the seaward regions from all of the breaking crest noise except for the break points at the crest ends, which radiate into relatively bubble-free water. Thus the wave noise appears to radiate from a compact source at either end of the breaking wavecrest. As the wave break points move through the surf zone, those audible at a monitoring array generate a time-dependent peak in the wave noise horizontal directivity [indicated by  $\theta(t)$  in Fig. 20]. The figure also shows how the screening effect leads to converging hot-spots when these are observed in pairs.

The motion of the break points through the surf zone, and therefore the noise directionality, is a function of the wave height profile of the incident wave field, and also the beach topography. The model developed in Sec. IV describes the trajectory of break points for waves with an exponential wave height profile that are incident on a plane beach. The case of multiple peaks in a wave height profile, such as the event illustrated in Fig. 20 and as casually observed from the beach, could be handled by segmenting the wavecrest into lengths that contain a single peak and then analyzing the break point motion for each of these lengths separately.

The hypothesis that the observed phenomenon is a screening effect and not an effect intrinsic to the breaking crest is largely consistent with the observed properties of hot-spots. The hot-spots were not usually observed in pairs, counter to what would be expected if the wave break points were simply more acoustically energetic than the rest of the breaking wavecrest. The few events that did exhibit pairs of hot-spots showed *converging* radiating regions; another result that is inconsistent with a hypothesis based on energetic break points alone. In addition, the few events observed on the seaward side of the array did not show hot-spots, consistent with the idea that there is no air residue in front of the wave to absorb the sound.

## VII. CONCLUDING REMARKS

The principle result of this study is the discovery of acoustic hot-spots behind breaking surf. To an observer behind a shoaling wave, the breaking wave noise apparently radiates from compact acoustic sources which move through the surf zone. The acoustic measurements of the bubbly residue behind the wave and the propagation modeling reported here both provide strong evidence that the hot-spots are caused by an acoustic shadow created by the bubbly residue.

A simple model for the movement of the wave break points through the surf zone has been proposed based on the McCowan breaking criterion and a family of exponential profiles for the variation of the wave height at the break points along the wavecrest. A detailed reconciliation between this model and acoustical measurements of the wave

noise directionality must await future studies with more comprehensive measurements of beach bathymetry and the wave height profiles of the incident wave field.

The results of the present work have a number of broad implications for the study of acoustics in the surf zone. There is clearly a relationship between the dynamics of wave breaking and the directionality of the resulting noise which could be exploited to learn more about the wave breaking process itself. For example, simultaneous measurement of the wave height profile and noise field directionality would allow study of the connection between the wave height, bottom bathymetry, and the formation of break points. The contribution of surf noise 100 m or more off shore is dependent on the temporal and spatial distribution of the population of quiescent bubbles left behind the wave, and needs to be included in predictive models for the surf noise.

The role of acoustically active bubbles in generating ambient oceanic noise has been the subject of several studies,<sup>3,22,23</sup> and the background layer of micro-bubbles in the ocean is thought to affect wave noise propagation.<sup>24-26</sup> The effects of absorption are even more pronounced in the surf zone, and the population of quiescent bubbles entrained by breaking surf is as important to the properties of the breaking wave noise as the acoustically active bubbles that are generating the sound. The effects are pronounced in the surf zone because the large length to height ratio inherent in the geometry accentuates the screening effect of the bubble plumes. An effect analogous to the acoustic hot-spots observed in the surf zone should also be seen in measurements of the horizontal directivity of breaking wave noise made within a meter or so of the surface in the open ocean.

## ACKNOWLEDGMENTS

I am grateful to Dr. Robert T. Guza for a number of helpful discussions. I am pleased to thank Rod Milan and James Uyloan who assisted with the experiment design and deployment, and Dr. M. Dale Stokes who took the photographs in Fig. 11. This work was conducted as part of the Adaptive Beach Monitoring Experiment, with the assistance of the military and civilian personnel at the Camp Pendleton Marine Base. The support of the Office of Naval Research, Contract Nos. N00014-96-0120 and N00014-96-1-0852, is gratefully acknowledged.

## APPENDIX: THE RELATIONSHIP BETWEEN NOISE COHERENCE AND DIRECTIONALITY

A broadband, two element hydrophone array can be used to obtain an estimate of the ambient noise field directionality in the following way. Consider the array geometry shown in Fig. A1. Two hydrophones, immersed in a homogeneous noise field, are separated by a distance  $L$  along an axis  $A$ . The noise field coherence  $\Gamma(\Omega)$  at the hydrophone array, measured as a function of frequency and fixed separation, is given by<sup>27</sup>

$$\Gamma(\Omega) = \frac{\langle S_1(\omega)S_2^*(\omega) \rangle}{\sqrt{\langle S_1(\omega)S_1^*(\omega) \rangle \langle S_2(\omega)S_2^*(\omega) \rangle}}, \quad (\text{A1})$$

where  $S_1(\omega)$  and  $S_2(\omega)$ , respectively, denote the Fourier transforms of the pressure time series induced by the noise field respectively at hydrophones 1 and 2, and the dimensionless, scaled frequency  $\Omega = \omega L/c$ , where  $\omega$  is angular frequency and  $c$  is the sound speed in the water column. The symbols\* and  $\langle \rangle$  in Eq. (A1), respectively, denote complex conjugation and ensemble average.

The noise field directionality integrated over a cone centered on the array axis (see Fig. A1) and coherence are related via the spatial Fourier transform

$$\Gamma(\Omega) = \frac{1}{2\pi} \int_{-\pi}^{\pi} F(y) \exp(i\Omega y/\pi) dy, \quad (\text{A2})$$

where  $y = \pi \sin(\theta)$ ,  $\theta$  is angle measured from a normal to the array axis and  $F(y)$  is a measure of the incident acoustic intensity. Note that our angle convention is different from that normally adopted in the literature, but is convenient for our array geometry and orientation with respect to the shore line. The more standard convention is to measure angle from the array axis. In general,  $F$  will be a function of both spherical polar angular coordinates, but any variability in power around the array axis is averaged out by the omni-directional response of the hydrophones. It has been tacitly assumed that  $F$  is independent of frequency. Equation (A2) holds true regardless, but the assumption is required in order to express  $F$  as a function of  $\Gamma$ . The practical implications of this assumption are discussed at the end of this Appendix.

It follows directly from Eqs. (A1) and (A2) that the directional density function obeys a normalization constraint. As the hydrophone separation  $L$  tends to zero,  $S_1 \rightarrow S_2$ , and from Eq. (A1) it follows that  $\Gamma(0) \rightarrow 1$ . It follows from Eq. (A2) that  $F(y)$  must satisfy the constraint:

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} F(y) dy = 1. \quad (\text{A3})$$

To obtain an expression for the directional density function in terms of the broadband coherence,  $F(y)$  is expanded as a Fourier series:

$$F(y) = \frac{a_0}{2} + \sum_{j=1}^{\infty} (a_j \cos(jy) + b_j \sin(jy)), \quad (\text{A4})$$

where

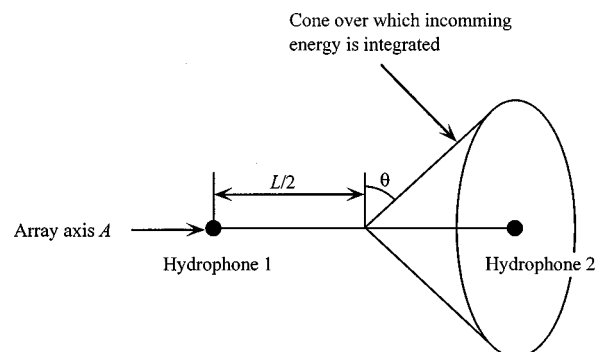


FIG. A1. Two-element hydrophone array geometry.

$$a_j = \frac{1}{\pi} \int_{-\pi}^{\pi} F(y) \cos(jy) dy, \quad j=0,1,2,\dots, \quad (\text{A5})$$

and

$$b_j = \frac{1}{\pi} \int_{-\pi}^{\pi} F(y) \sin(jy) dy, \quad j=1,2,3,\dots \quad (\text{A6})$$

The Fourier coefficients  $a_j$  and  $b_j$  can be expressed in terms of the coherence by noting that

$$a_j + ib_j = \frac{1}{\pi} \int_{-\pi}^{\pi} F(y) \exp(ijy) dy = 2\Gamma(\pi j), \quad j=1,2,3,\dots \quad (\text{A7})$$

It follows immediately that:

$$a_j = 2 \operatorname{Re}(\Gamma(\pi j)), \quad b_j = 2 \operatorname{Im}(\Gamma(\pi j)). \quad (\text{A8})$$

Substituting for the coefficients in Eq. (A4), and noting that the normalization constraint requires that  $a_0=2$ , an expression is obtained for the directionality in terms of the noise coherence:

$$F(y) = 1 + 2 \sum_{j=1}^{\infty} (\operatorname{Re}(\Gamma(\pi j)) \cos(jy) + \operatorname{Im}(\Gamma(\pi j)) \sin(jy)). \quad (\text{A9})$$

Equation (A9) is an expression for the ambient noise directionality expressed in terms of the noise coherence sampled at discrete values of normalized frequency. Making the substitution  $y = \pi \sin(\theta)$  in Eq. (A9), the directional density function can be expressed in terms of angle:

$$F(\theta) = 1 + 2 \sum_{j=1}^{\infty} (\operatorname{Re}(\Gamma(\pi j)) \cos(\pi j \sin(\theta)) + \operatorname{Im}(\Gamma(\pi j)) \sin(\pi j \sin(\theta))). \quad (\text{A10})$$

In any practical measurement, the sum in Eq. (A10) will be truncated at some maximum value of frequency corresponding to  $j=j_{\max}$ , which sets the maximum angular resolution of the array. The array used for the noise measurements had a separation of  $L=0.327$  m and a bandwidth of 22.5 kHz which, assuming a sound speed of 1500 m/s, sets  $j_{\max}=10$  and the maximum angular resolution of the array is about 9 degrees.

Equation (A10) can be used to compute the noise field directionality in a plane from broadband measurements of the noise field coherence. The expression is exact provided that the noise field directionality is independent of frequency. In physical terms, this means that the noise field must be dominated by a single source, or multiple sources with the same spectral shape at the array location. This constraint is, in fact, reasonably well satisfied for our data set but Eq. (A10) cannot be used, in general, without some information about the nature and distribution of the sources generating the ambient noise.

Since  $\cos$  and  $\sin$ , respectively, are even and odd functions of their arguments, it follows from Eq. (A9) that the real part of the coherence determines the symmetrical component of the directional density function and the imaginary

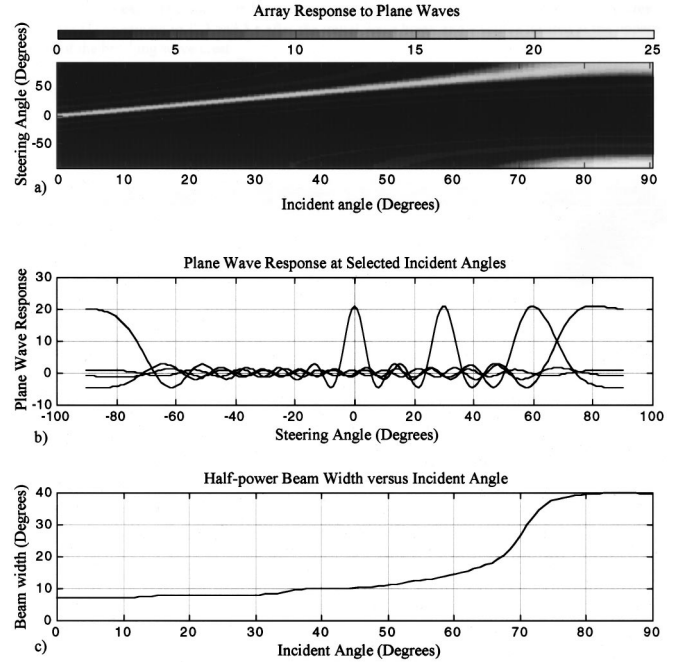


FIG. A2. The theoretical response of the two-element, broadband hydrophone array. (a) The array response to broadband plane waves plotted as a function of steering angle and angle of incident energy. (b) Line plots of the array response at selected angles of incidence. (c) The half-power beam width of the array plotted as a function of incident angle.

part determines the asymmetrical component of the directionality. For example, consider a plane wave incident on the array at the transformed angle  $y_0$  with the corresponding directional density function

$$F(y) = \alpha \delta(y_0 - y), \quad (\text{A11})$$

where, from the normalization constraint,  $\alpha=2\pi$  for  $-\pi < y < \pi$  and  $\alpha=\pi$  for  $y_0 = \pm\pi$ . The coherence function for the plane wave calculated from Eq. (A2) is

$$\Gamma(\Omega) = \frac{\alpha}{2\pi} \exp\left(i \frac{\Omega y_0}{\pi}\right). \quad (\text{A12})$$

For  $y_0 = -\pi$ , corresponding to energy incident from the shoreward side of the array in the experiment, the initial excursion of the imaginary component of the coherence is negative. For energy incident from the seaward side of the array the converse is true, and the initial excursion is positive. If the wavefront of the incident wave is aligned with the array axis, then  $y_0=0$  and the imaginary part of the coherence is identically zero.

The beam forming response of the array over a limited bandwidth is found by substituting Eq. (A12) into Eq. (A9) and evaluating the resulting sum up to the  $j_{\max}$  term. The result of setting  $j_{\max}=10$ , which is the maximum Fourier component available from the array deployed in the experiment, is shown in Fig. A2(a). The color contour plot shows the estimate of directionality plotted versus steering angle,  $\theta$ , and angle of the incident plane wave  $\theta_0$ . The formation of a beam at  $\theta = \theta_0$  is evident from the red strip running diagonally across the plot. The beam width varies depending on the rate of change of the trigonometric function arguments,  $\pi j \sin(\theta)$ , with respect to  $\theta$ . The effect of beam broadening is

most noticeable as  $\theta_0 \rightarrow \pi/2$  where the resolution of the array is significantly degraded. Figure A2(b) shows line plots of the beam forming response as a function of  $\theta$  at selected angles of incidence.

The final point concerns the maximum value that the directional density function can take for a band limited array. Consider a plane wave incident on the broad side of an array with a maximum frequency response corresponding to  $j_{\max}$ . The coherence of the plane wave is unity, and the estimate for the directional density function is

$$F(y) = 1 + 2 \sum_{j=1}^{j_{\max}} \cos(jy). \quad (\text{A13})$$

The maximum response occurs for  $y=0$ , and has a value of  $2j_{\max}+1$ . Thus the maximum response for the array used in the experiment is 21.

- <sup>1</sup>O. B. Wilson, M. S. Stewart, J. H. Wilson, and R. H. Bourke, "Noise source level density due to surf—Part I: Monterey Bay, CA," *IEEE J. Ocean Eng.* **22**, 425–433 (1997).
- <sup>2</sup>O. B. Wilson, S. N. Wolf, and F. Ingenito, "Measurements of acoustic ambient noise in shallow water due to breaking surf," *J. Acoust. Soc. Am.* **78**, 190–195 (1985).
- <sup>3</sup>G. B. Deane, "Sound generation and air entrainment by breaking waves in the surf zone," *J. Acoust. Soc. Am.* **102**, 2671–2689 (1997).
- <sup>4</sup>C. S. Clay and H. Medwin, *Acoustical Oceanography: Principles and Applications* (Wiley, New York, 1977).
- <sup>5</sup>E. Lamarre and W. K. Melville, "Instrumentation for the measurement of void-fraction in breaking waves: laboratory and field results," *IEEE J. Ocean Eng.* **17**, 204–215 (1992).
- <sup>6</sup>E. Lamarre and W. K. Melville, "Void fraction measurements and sound-speed fields in bubble plumes generated by breaking waves," *J. Acoust. Soc. Am.* **95**, 1317–1328 (1994).
- <sup>7</sup>E. Lamarre and W. K. Melville, "Sound-speed measurements near the ocean surface," *J. Acoust. Soc. Am.* **96**, 3605–3616 (1994).
- <sup>8</sup>S. Vagle and D. M. Farmer, "A comparison of four methods for bubble size and void fraction measurements," *IEEE J. Ocean Eng.* **23**, 211–222 (1998).
- <sup>9</sup>G. B. Deane and M. D. Stokes, "Air entrainment processes and bubble size distributions in the surf zone," *J. Phys. Oceanogr.* (in press, 1999).
- <sup>10</sup>D. H. Peregrine, "Breaking waves on beaches," *Annu. Rev. Fluid Mech.* **15**, 149–178 (1983).
- <sup>11</sup>J. A. Battjes, "Surf zone dynamics," *Annu. Rev. Fluid Mech.* 257–293 (1988).
- <sup>12</sup>C. E. Synolakis and J. E. Skejlbreia, "Evolution of maximum amplitude of solitary waves on plane beaches," *Journal of Waterway, Port, Coastal and Ocean Engineering* **119**, 323–342 (1993).
- <sup>13</sup>I. A. Svendsen, P. A. Madsen, and J. B. Hansen, "Wave characteristics in the surf zone," in *Proceedings of the 16th Coastal Engineering Conference* (American Society of Civil Engineers, New York, 1978), pp. 520–539.
- <sup>14</sup>S. T. Grilli, R. Subramanya, I. A. Svendsen, and J. Veeramony, "Shoaling of solitary waves on plane beaches," *Journal of Waterway, Port, Coastal and Ocean Engineering* **120**, 609–628 (1994).
- <sup>15</sup>M. S. Longuet-Higgins, "The refraction of sea waves on shallow water," *J. Fluid Mech.* **1**, 163–177 (1956).
- <sup>16</sup>A. J. Bowen and R. T. Guza, "Edge waves and surf beat," *J. Geophys. Res.* **83**, 1913–1920 (1978).
- <sup>17</sup>J. McCowan, "On the highest wave of permanent type," *Philos. Mag. J. Sci.* **38** (1894).
- <sup>18</sup>J. R. Weggel, "Maximum breaker height," *Journal of Waterways, Harbors Coastal Engineering Division., ASCE* **98** (1972).
- <sup>19</sup>J. M. Smith and C. L. Vincent, "Shoaling and decay of two wave trains on beach," *Journal of Waterway, Port, Coastal and Ocean Engineering* **118**, 517–533 (1992).
- <sup>20</sup>S. L. Douglas, "Influence of wind on breaking waves," *Journal of Waterway, Port, Coastal and Ocean Engineering* **116**, 651–663 (1990).
- <sup>21</sup>Z. Y. Zhang and C. T. Tindle, "Complex effective depth of the ocean bottom," *J. Acoust. Soc. Am.* **93**, 205–213 (1993).
- <sup>22</sup>M. L. Banner and D. H. Cato, "Physical Mechanisms of Noise Generation by Breaking Waves—A Laboratory Study," in *Sea Surface Sound*, Series C: Mathematical and Physical Sciences (Kluwer Academic, Dordrecht, 1988), pp. 429–436.
- <sup>23</sup>M. R. Loewen and W. K. Melville, "A model of the sound generated by breaking waves," *J. Acoust. Soc. Am.* **90**, 2075–2080 (1991).
- <sup>24</sup>M. J. Buckingham, "On acoustic transmission in ocean-surface waveguides," *Philos. Trans. R. Soc. London, Ser. A* **335**, 513–555 (1991).
- <sup>25</sup>D. M. Farmer and D. D. Lemon, "The influence of bubbles on ambient noise in the ocean at high wind speeds," *J. Phys. Oceanogr.* **14**, 1762–1778 (1984).
- <sup>26</sup>D. M. Farmer and S. Vagle, "Waveguide propagation of ambient sound in the ocean-surface bubble layer," *J. Acoust. Soc. Am.* **86**, 1897–1908 (1989).
- <sup>27</sup>H. Cox, "Spatial correlation in arbitrary noise fields with application to ambient sea noise," *J. Acoust. Soc. Am.* **54**, 1289–1301 (1973).

# Target localization for a three-dimensional multibeam sonar imaging system

Jules S. Jaffe

*Marine Physical Lab, Scripps Institution of Oceanography, La Jolla, California 92093-0238*

(Received 31 July 1997; accepted for publication 22 February 1999)

One advantage of three-dimensional multibeam imaging systems is that they have the ability to image and track many targets concurrently in their field of view. In this regard, the accurate estimation of both position and target strength are important data which can be derived from the echo responses of an active multibeam imaging system. This article considers an algorithm which can be used to yield more accurate positional estimates for targets than available from the preformed beams of a given multibeam system: FishTV [Jaffe *et al.*, *Deep-Sea Res.* **42**, 1495 (1995)]. The method refines a least squares residual of the predicted data versus observed data by adjusting the parameters of a model. The parameters are the position and target strength of the target. It is demonstrated that this procedure is a maximum likelihood estimator for position and target strength in high signal-to-noise. In addition, the shape of the least squares surfaces indicates that the algorithm is somewhat robust with respect to noise. © 1999 Acoustical Society of America. [S0001-4966(99)03406-2]

PACS numbers: 43.30.Sf, 43.30.Ft, 43.30.Vh [DLB]

## INTRODUCTION

Oceanographic three-dimensional sonar imaging systems have great potential for measuring animal abundance, and also behavior. In the case of abundance, the number density of animals is of primary interest. In the case of behavior, animal trajectories provide a measure of the relevant information. Three-dimensional sonars with fast frame rates can provide unambiguous measurements of animal positions as a function of time, which can be used for both purposes.

In general, the resolution requirements for the observation of animals can be quite high with respect to the diffraction limited beamwidths of the array components. Considering the abundance measurements, it is desirable to have the targets resolvable to a density which is high enough to be able to count individuals unambiguously with a wide field of view. Most conventional echo counting systems use either split beam or dual beam techniques which have limited ability to resolve multiple animals at the same range (Barange *et al.*, 1996). Certainly, as a lower limit, it is desirable for the system to be able to resolve animals at densities that are higher than the mean value. On the other hand, one cannot discount the possibility that animals prefer to associate together. In this case, the resolving capability would place a severe constraint on the system construction and performance. In practice, some compromise between these two situations is usually dictated by the increased complexity of the hardware needed for finer and finer discrimination with a wide field of view.

Considering the trajectory measurements, the ability to discriminate small positional changes of an animal might be necessary to observe any behavior whatsoever. This is likely to be the case for animals which cannot swim very fast, such as zooplankton. In addition, since the observation of target velocity requires the computation of the finite differences between successive positions, any error in positional infor-

mation will be amplified in computing velocity. Here, enhanced discrimination against target clutter is not the problem—better target position estimate is desired.

In this article, an algorithm is presented whose goal is to obtain better positional estimates of targets in the three-dimensional field of view of a sonar system. The algorithm is somewhat similar to one that has been proposed for the localization of targets in the ocean (Baggeroer *et al.*, 1988). However, in this case, the method is applied to data that have been collected by a specific imaging system (FishTV, Jaffe *et al.*, 1995). Specifically, the geometry of the sonar system is incorporated into a scheme which predicts data for a given target position. Refinement of the predicted output of the system against the observed data, as a function of target position and target strength, produces an accurate estimate of both parameters.

The FishTV sonar system was specifically designed to measure the three-dimensional locations of macrozooplankton in the ocean. As such, it operates at a frequency of 445 kHz. At this frequency, energy reflected from animals with target strengths as small as  $-90$  dB at ranges as great as 10 m can be measured (Jaffe *et al.*, 1995). A 2-cm euphausiid has an approximate target strength of  $-70$  dB at this frequency (Jaffe *et al.*, 1998). To date, the absence of adequate technology has prevented the measurement of *in situ* behavior of these animals. Interesting information can be obtained by tracking them, such as their foraging behavior and rate of energy expenditure over their daily activities. Euphausiids are an important component of the oceanic food chain and information about their behavior can be used to infer the degree to which these animals respond to their environment or interact with each other.

Previous results obtained with the FishTV system have allowed the inference of animal trajectories (McGehee and Jaffe, 1996). These animals were likely micronekton and

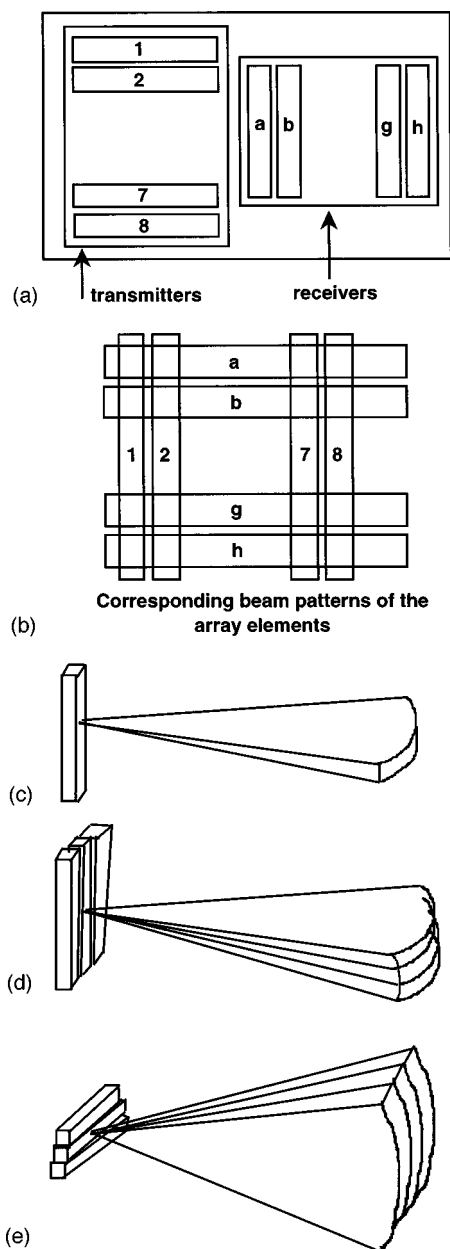


FIG. 1. The geometry of the arrays and their corresponding beam patterns. (a) A view of the placement of the transducers and receivers from the front. (b) A front view of the beam patterns of the elements shown in (a). (c) A single transmit element and its beam pattern. (d) A set of three transmitters showing the orientations of three of the eight transducers and their composite beam patterns. (e) Three of the receive elements and their corresponding beam patterns. [(c), (d), and (e) are after McGehee and Jaffe, 1996].

large euphausiids with fast swimming speeds. Additional information about the swimming behaviors of smaller, more slowly swimming animals with the FishTV system requires the development of more advanced processing algorithms which will allow the inference of target positions to a resolution that is smaller than the physical beamwidth of the transducers. This article presents an algorithm which can be used for this purpose.

Previous results of other researchers specifically concerned with improving the performance of the FishTV system have been aimed at using adaptive beamforming techniques to increase system resolution (Papazoglou and Krolik,

1996). In addition, methods have been proposed to track particles in three dimensions (Frenkel and Feder, 1995). Certainly a vast literature exists on various techniques for increasing the resolution of sonar arrays over the diffraction limit. Standard texts outline many of the techniques (Steinberg, 1976; Johnson and Dugeon, 1993) that have been successfully used for several applications. One possible difference in applying these methods to our system is that, since the transmitters are sequentially pulsed, the target may move a reasonable fraction of a wavelength during the acquisition of the complete image. Without modeling this movement, the appropriate algorithm in this case is an *incoherent* technique since the system is not temporally stationary over the time it takes to collect a single frame.

## I. DETAILS OF THE IMAGING SYSTEM

The FishTV imaging system has been described in previous documents; however, for completeness we reformulate here the expressions for the imaging properties of the system. The system uses unshaded rectangular apertures (16) of dimensions  $27\lambda \times 3\lambda$  for both projecting (8) and receiving (8) sound, each of which have a beam pattern of 2 degrees  $\times$  20 degrees.

The elements are placed in a two-dimensional array with the receivers and transmitters located side by side [Fig. 1(a)]. When the transmit elements are horizontally oriented they produce beam patterns which are narrow in the horizontal and wide in the vertical [Fig. 1(b)]. Since the receivers are rotated by 90 degrees with respect to the transmitters, their beam patterns are also rotated 90 degrees.

The elements are arranged in a spiral fashion so that the beam patterns of both the transmitters and receivers fill the field of view. Figure 1(c) (after McGehee and Jaffe, 1996) illustrates the transmit pattern of a single element. Figure 1(d) shows three of the receivers, which are rotated by 2 degrees (the beamwidth of the transducers). Figure 1(e) illustrates the geometry of the transmitters and their beam patterns. The composite pattern for the entire system is the product of the set of transmitters and the receivers. Since the transmitters and receivers each insonify a 2 degree  $\times$  20 degree field of view, their composite pattern fills a  $16(8 \times 2)$  degree  $\times$   $16(8 \times 2)$  degree area.

All of the receivers are used simultaneously. Numbering the transmitters from one to eight, the system transmits first on transmitter one and then receives on all eight. Next, the system transmits on transmitter two, and receives on all eight. This process is continued until all of the transmitters are pulsed and the information is assembled into a single three-dimensional frame. FishTV assembles an "image" as a three-dimensional matrix  $(i,j,k)$  of dimensions  $8 \times 8 \times 512$ , where the  $(i,j,k)$  element is the  $i$ th transmitter coupled with the  $j$ th receiver. Here,  $k$  is the range coordinate which results from the temporal sampling of the system. The system forms images by transmitting and receiving on a grid with sequential transmissions taking place on the columns and the receipt of sound taking place simultaneously on all of the rows. The geometry has been considered by other researchers and is known as a Mills crossed array (Urlick, 1983).



The physical geometry of the system is quite straightforward. The far field pattern of each of the rectangular apertures of dimensions  $L \times K$  can be written as

$$BP_{\text{trans}}(\theta_x, \theta_y) = \frac{\sin(\pi S_x L / \lambda) \sin(\pi S_y K / \lambda)}{(\pi S_x L / \lambda)(\pi S_y K / \lambda)}, \quad (1)$$

where  $S_x = \sin \theta_x$  ( $\theta_x$  being the angle from broadside in  $x$ ) and  $S_y = \sin \theta_y$  ( $\theta_y$  being the angle from broadside in  $y$ ).

Note that the beam pattern has been expressed in terms of the amplitude of the spatial pattern, not the intensity, as is more common. This has been done to be consistent with the maximum likelihood estimator (MLE) approach which uses this quantity, as discussed below.

When the transducer is rotated by angle  $\theta_i$  in the  $\theta_x$  direction, the transmit beam pattern can be described as

$$BP_{\text{trans}}(\theta_x - \theta_i, \theta_y). \quad (2)$$

Since the receiving transducers of dimensions  $L \times K$  are rotated by 90 degrees relative to the transmitting ones [Fig. 1(a)], their beam patterns can be expressed as

$$BP_{\text{recv}}(\theta_x, \theta_y) = \frac{\sin(\pi S_x K / \lambda) \sin(\pi S_y L / \lambda)}{(\pi S_x K / \lambda)(\pi S_y L / \lambda)}. \quad (3)$$

Rotating the receivers by  $\theta_j$  yields

$$BP_{\text{recv}}(\theta_x, \theta_y - \theta_j). \quad (4)$$

Now the composite beam patterns for the transmitters and receivers can be represented as

$$BP_{\text{comp}}(\theta_x, \theta_y; \theta_i, \theta_j) = BP_{\text{tran}}(\theta_x - \theta_i, \theta_y) BP_{\text{recv}}(\theta_x, \theta_y - \theta_j), \quad (5)$$

where  $i = \{1, 8\}$  and  $j = \{1, 8\}$ .

All together, the eight pointing angles of the transmitters and the eight pointing angles of the receivers form a two-dimensional matrix of dimensions  $8 \times 8$  at each range bin.

The system transmits the same selected transmit wave (generated with an arbitrary waveform generator) on each of the transmitted channels. These transmit waves are either chirp waveforms or short sine wave pulses. Assuming that the envelopes of such waves are of Gaussian form (or the matched filtered output from them are) with a temporal extent of standard deviation  $\sigma_t$ , the temporal dependence of such a wave can be represented as:

$$a(t) = \exp\left(-\frac{t^2}{2\sigma_t^2}\right). \quad (6)$$

The complete expression for the received wave as a function of both time and space can be written as the product of these waveforms:

$$S(\theta_x, \theta_y, t; \theta_i, \theta_j) = a(t) BP_{\text{comp}}(\theta_x, \theta_y; \theta_i, \theta_j). \quad (7)$$

Here, we ignore the attenuation and spatial spreading of the sound waves as these effects are systematic and do not effect the relative amplitude of the reflected energy at a particular range. Before digitization, the reflected signal is

quadrature demodulated, resulting in both sine and cosine channels. Given  $\psi_0$  as the carrier frequency, and  $\omega_d$  as the demodulating tone, the complex quadrature components can be written with noise added (after low pass filtering) as

$$a(t) \exp(2\pi i(\omega_0 - \omega_d)t) + n_c(t). \quad (8)$$

Digitization of the both the real and the imaginary components of this waveform produces a set of data with two data streams for each of the receive transducers. The waveforms can be processed to extract the relative phase of the output signal, or to compute either the magnitude (modulus) or the intensity of the output for each of the channels. If the statistics of the noise component are such that the real and imaginary components are Gaussian, the noise variable  $n_c(t)$  is a circular complex Gaussian random variable (Goodman, 1985). The total system response to an object at range  $z_0$  and position  $\theta_{x0}, \theta_{y0}$  can be computed by taking expression (8) and combining it with  $BP_{\text{comp}}(\theta_x, \theta_y; \theta_i, \theta_j)$  to yield:

$$A_{\text{OBS}}(\theta_{x0}, \theta_{y0}, \theta_i, \theta_j, t) = BP_{\text{comp}}(\theta_{x0}, \theta_{y0}; \theta_i, \theta_j) a(t - t_0) \times \exp(2\pi i(\omega_0 - \omega_d)t) + n_c(t) \quad (9)$$

for the set of transmit and receive beams  $\{i, j\}$ , where  $t_0 = 2z_0/c$ , and  $c$  is the speed of sound in the medium. In this expression the target reflectivity has been normalized to one.

## II. PROPOSED ALGORITHM

The basic idea of the algorithm is to regard the estimation of the position of an object as a parameter estimation task. Note that this is somewhat in distinction to beamforming algorithms which seek to create a better image (Papazoglou and Krolik, 1996). In order to simplify the expressions in the following discussions, the range is fixed and it is assumed that there is a single target in the three-dimensional matrix at that known range. In a practical sense, the range of a single target in the three-dimensional field of view of the system can be estimated by simply looking for the maximum value in the array. On the other hand, although the assumption of a single target simplifies the discussion, the problem of extracting multiple target positions from the three-dimensional matrix is very interesting and more complicated. This subject will be addressed in Sec. IV. In this paper the algorithm is limited to the single target case.

In order to compute the maximum likelihood estimator (MLE) for the position of a target given the observed data, the probability density function for the observed data as a function of target position needs to be determined. Here, we make use of "standard" results. The probability density function for the observation of a vector of length "a" given true length "s" with circularly complex Gaussian noise with standard deviation  $\sigma$  added is a Rician probability distribution function (Goodman, 1985):

$$p(a|s) = \frac{a}{\sigma^2} \exp\left(-\frac{a^2 + s^2}{2\sigma^2}\right) I_0\left(\frac{as}{\sigma^2}\right), \quad (10)$$

where  $I_0$  is a modified Bessel function of the first kind, zero order. In a typical application (Kay, 1993), the log likelihood function is differentiated and set to zero in order to compute

the minimum of this function as a closed form solution. Alternatively, the probability density function can be incorporated into a norm whose minimum is the MLE for the problem. Although quite achievable, in this case, a much simpler solution exists when the signal-to-noise is high: that is, the fact that the distribution approaches a Gaussian (Goodman, 1985) (in even moderate signal-to-noise) can be used to justify the use of a maximum likelihood estimator which uses a least square criteria:

$$\text{minimize } \sum_{i=1}^8 \sum_{j=1}^8 (s_{ij} - a_{ij})^2 \quad \text{for } i = \{1,8\}, j = \{1,8\}. \quad (11)$$

Therefore, in order to obtain a maximum likelihood estimate for the position of a single object, look for a minimization of the least square residual between the observed and predicted data. Under the assumption that the “true” position of the target is  $\theta_l$  and  $\theta_m$ , a global minimum for such a function will be the maximum likelihood estimator and will result in estimated values for the unknown parameters  $\theta_l$  and  $\theta_m$  as  $\hat{\theta}_l$  and  $\hat{\theta}_m$ .

The least squares estimator is

$$\sum_{i=1}^8 \sum_{j=1}^8 [A_{\text{OBS}}(\theta_l, \theta_m, \theta_i, \theta_j) - S_{\text{comp}}(\hat{\theta}_l, \hat{\theta}_m, \theta_i, \theta_j)]^2, \quad (12)$$

where

$$S_{\text{comp}}(\hat{\theta}_l, \hat{\theta}_m, \theta_i, \theta_j) = BP_{\text{comp}}(\hat{\theta}_l, \hat{\theta}_m, \theta_i, \theta_j) \quad (13)$$

is the noiseless data, obtained from an accurate model.

### III. RESULTS

In order to explore the application of the minimization of expression (12), a forward model was used to compute the  $8 \times 8$  matrix for a given test target position, according to Eq. (5). The actual model used for the computation was a bit more complicated as the sonar arrays for transmit and receive are actually displaced from each other, as can be seen in Fig. 1. Next, the field of view of the sonar was discretized into a  $64 \times 64$  matrix of potential positions. Sonar data, predicted for the position of a single test target at each of these  $(64)^2$  locations, were then computed. Following this, a set of  $(64)^2$  least squares residuals was computed between the set of predicted data for the single target and the set of  $(64)^2$  potential positions for each of the  $8 \times 8$  matrices of target positions. Mathematica (Wolfram, Inc.) was used to generate both the algebraic expressions for the reflected energy as a function of beam pattern and geometry and to also compute the numerical values of the data. In all cases, the synthetic data were run without adding noise to the observed data.

To examine the performance of the estimator the least squares surface was computed for a number of different test target locations. The small size of the matrices (both data and also test target grid) permits the computation of the entire least squares surface as a function of the  $64 \times 64$  set of potential target positions for a given test target location. In optimization theory, this would be referred to as using an

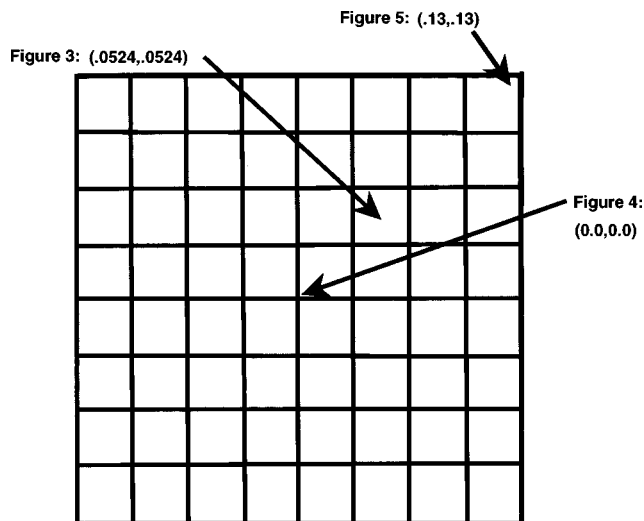


FIG. 2. The two-dimensional field of view of the system showing the  $8 \times 8$  set of beams that the sonar uses and the positions of the test targets (in radians) used for the simulations.

exhaustive search procedure to search for a global minimum of expression (12) as a function of target position. Several cases were computed (Fig. 2) which consisted of placing the test target either directly in the middle of the intersection of the sonar beams (Fig. 3), between the set of beams (Fig. 4), or in the periphery of the field of view of the system (Fig. 5).

Figures 3(a), 4(a), and 5(a) show the predicted  $8 \times 8$  images at the range of the target. Figures 3(b), 4(b), and 5(b) show contour plots of the reciprocal of the least squares residuals for the set of computed  $64 \times 64$  locations for each of these test target cases. The plots are contoured in levels of 0.99, 0.5, 0.1, 0.05, and 0.01 of the maximum value for each data set. One-dimensional slices through this least squares surface are perhaps more quantitatively interpretable, as shown in Figs. 3(c), 4(c), and 5(c). Since the functions are symmetric in the vertical and horizontal, only one slice is needed. As seen in the figures, there is only one global minimum for each of the positions which is at the true position of the test target. In addition, for the  $64 \times 64$  case, the shape of the minimum seems to be quite steep, indicating that the algorithm can be expected to compute reasonable values for the positions of the objects for moderate levels of noise.

An interesting feature of this ability to localize targets is the additional concurrent capability to compute target strength more accurately. The previous method that was used by our group to compute target strength was to simply take the largest response of the data matrix at the range of the target. If a target was exactly in the center of both the transmit and receive beams, an accurate estimate of target strength is obtained. However, for targets located between beams, the target strengths are biased low with the worst discrepancy occurring when the target is exactly in the center of the four neighboring beams (Fig. 4). This results in a  $-10$  dB underestimate of target strength in this worst case situation.

Equation (9) can be rewritten to include the variation in target reflectivity by multiplying it by a scalar factor:

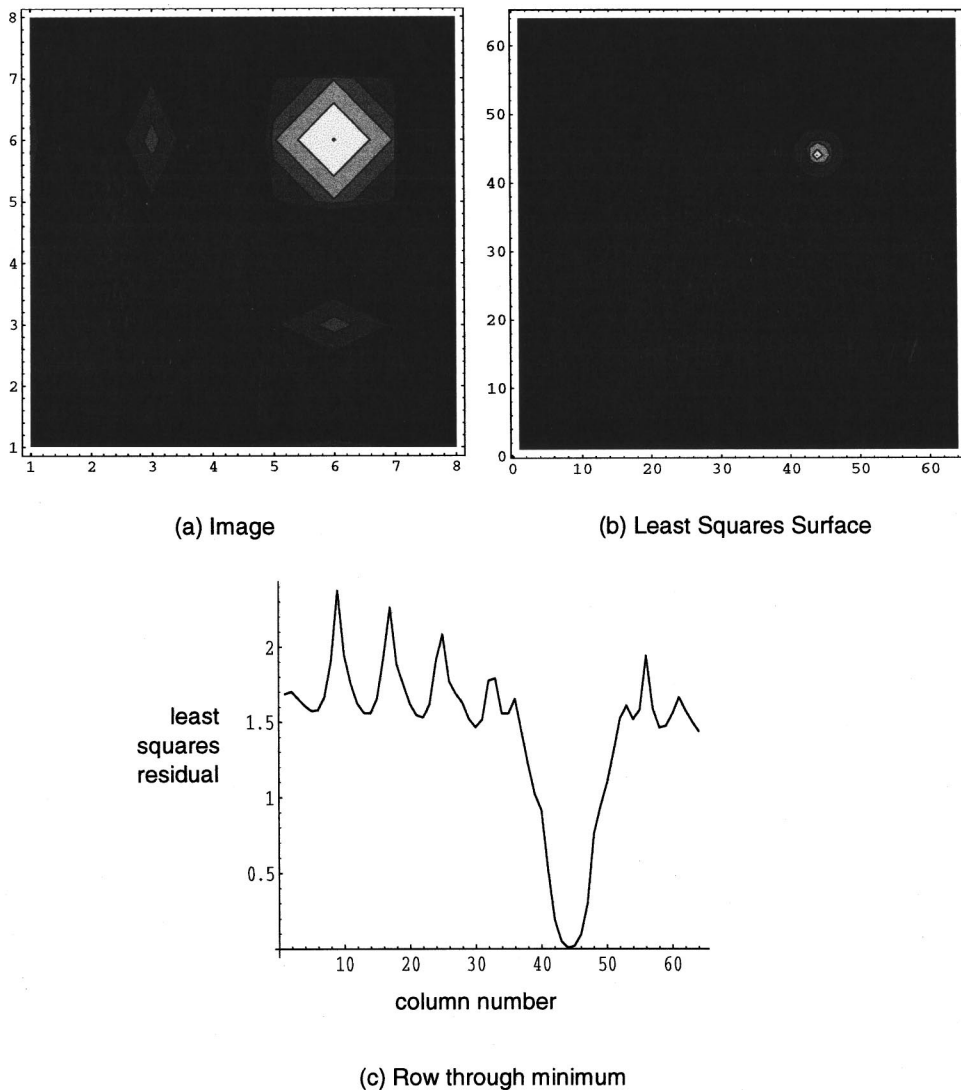


FIG. 3. A graph of the least squares error surface for radial target position (0.0524, 0.0524). (a) The  $8 \times 8$  image as collected by the sonar. (b) A contour diagram of the reciprocal least squares error surface (contours are 0.99, 0.5, 0.1, 0.05, 0.01 of the maximum value). (c) A one-dimensional slice through the error surface at the minimum error position.

$$\rho(\theta_{x0}, \theta_{y0})BP_{\text{comp}}(\theta_{x0}, \theta_{y0}, \theta_i, \theta_j)a(t-t_0) \times \exp(2\pi i(\omega_0 - \omega_d)t) + n_c(t). \quad (14)$$

Now, an estimate of target strength (proportional to reflectivity) can be included in the least squares procedure by varying this parameter in addition to the location. In order to estimate target strength, an approximate scale factor was computed by taking the ratio of the maximum value of each of the  $64 \times 64$  test matrices to the observed data before computing the least squares summation. The least squares surface was then computed and the minimum was used in order to estimate the position of the object. Given this object position, a set of object reflectivities was then used to compute the minimum least squares value. This is equivalent to a one-dimensional parameter search for the target strength variable. Figure 6 shows the dependence of the least squares residual on this value. The graph shows that the original value of 0.82 was not an optimal fit for computing the least squares residual as the target was actually between two of the  $64 \times 64$  beams. Following this, this best value for the reflectivity was then used to compute a new  $64 \times 64$  set of values for the position. As seen here, in one iteration the algorithm con-

verges to the correct value of 1 (a unit reflector was used for this simulation).

#### IV. DISCUSSION

In this article the problem of computing the position and target strength of a single target within the field of view of the FishTV system has been considered. The motivation is provided by the ecological significance of measuring the abundance, spatial distribution, and *in situ* behavior of animals. The procedure presented is an optimization method which minimizes a least squares norm (shown to be approximately a maximum likelihood estimator) of the target position.

Certainly one of the odd features of the FishTV system is that it uses relatively large multiple wavelength apertures for both transmitting and receiving sound. In this context, it is interesting to observe that the system can achieve substantially better resolution than the beamwidth of the transducers (in high SNR). Since the technique proposed here is incoherent in the sense that information between receiving elements is not being combined, it cannot be claimed that an effective aperture is synthesized which is larger than the physical ones

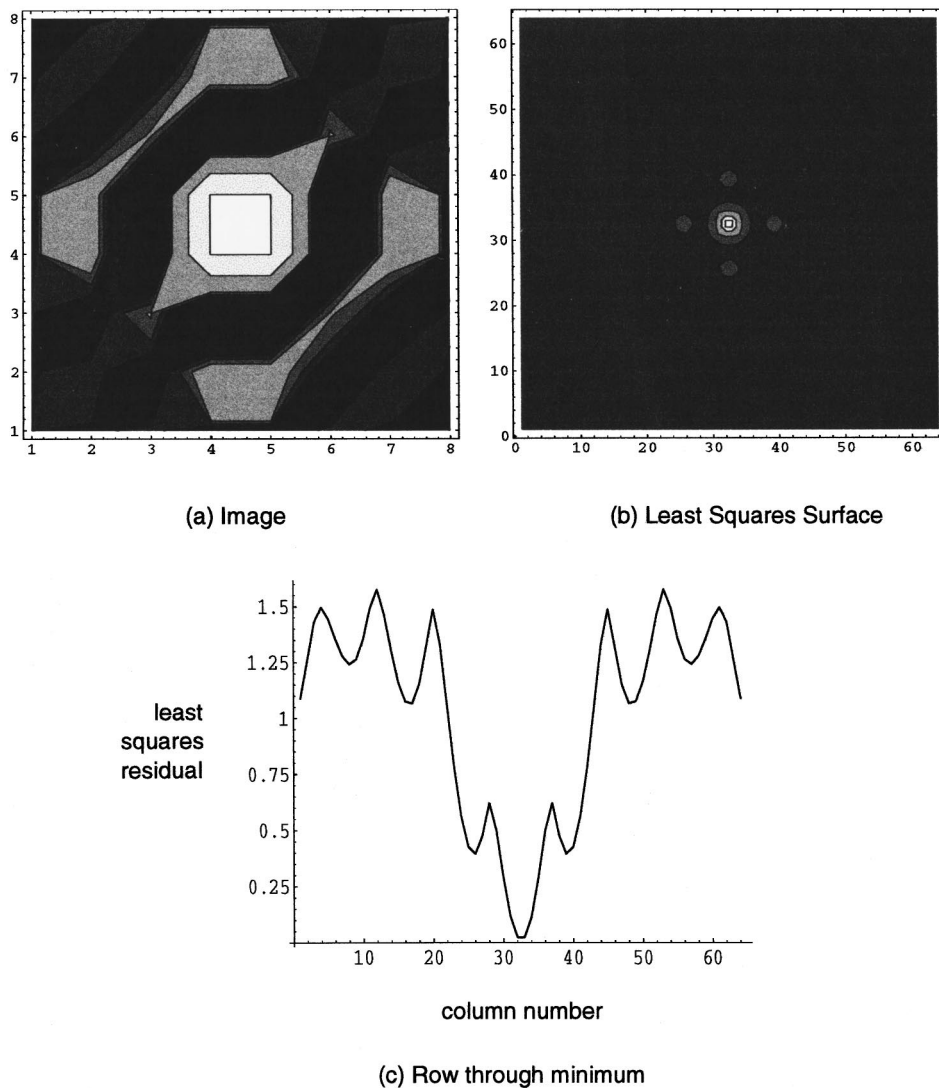


FIG. 4. A graph of the least squares error surface for radial target position (0.0, 0.0). (a) The  $8 \times 8$  image as collected by the sonar. (b) A contour diagram of the reciprocal least squares error surface (contours are 0.99, 0.5, 0.1, 0.05, 0.01 of the maximum value). (c) A one-dimensional slice through the error surface at the minimum error position.

that are being used. As such, it is clear that the ability to localize the object to this resolution is simply due to the fact that the change in position of a target leads to changes in the relative amplitudes of the received pattern of the array.

One surprising feature of the results is that the existence of sidelobes in the beam patterns seems to help in the target localization procedure. This is evidenced by the differences in sensitivity between Figs. 3 and 4. For example, the  $8 \times 8$  data matrix corresponding to the object in the center of both the transmitting and receiving beams, as in Fig. 3, shows extremely low sidelobe levels. This is because reflected energy from the target is practically in the nulls of the received (sinc function) beam patterns. Contrasted with Fig. 4, which samples practically at the local peaks of the sinc function, the minimization surface for the case when the target is in between beams seems to have a steeper minimum than the case when the target reflectance is sampled in the nulls. This means that the capability to localize the target, using this imaging system and the algorithm formulated here, is better when the target is inbetween beams than when the target is in the center of two beams. Clearly the sidelobe levels reduce the ability of the instrument to resolve multiple targets in the presence of clutter; however, as considered here, in the case

of localization, they do not rule out good performance (in high SNR).

One of the merits of this algorithm is that it incorporates all of the system geometry in computing target position. This is a considerable advantage because the system geometry is complicated by the separation between the transmitting and receiving elements. This produces a parallax error when simply assigning the maximal response in the measured data to object position. Having a system model which incorporates the system geometry automatically takes care of this issue. The simulations considered in this paper were constructed so that the parallax issues between the sources and receiver were not a factor.

One of the more interesting issues, in a practical sense, is how well the algorithm will perform in the presence of multiple targets. That is, will the algorithm successfully identify the positions of multiple targets within the field of view of the system? Clearly, if the targets are well resolved in that received patterns from them do not overlap, the performance will be identical to that described here. On the other hand, if there is substantial overlap between reflections from targets, especially weak ones sitting next to strong ones, the strong returns might make the separation of the

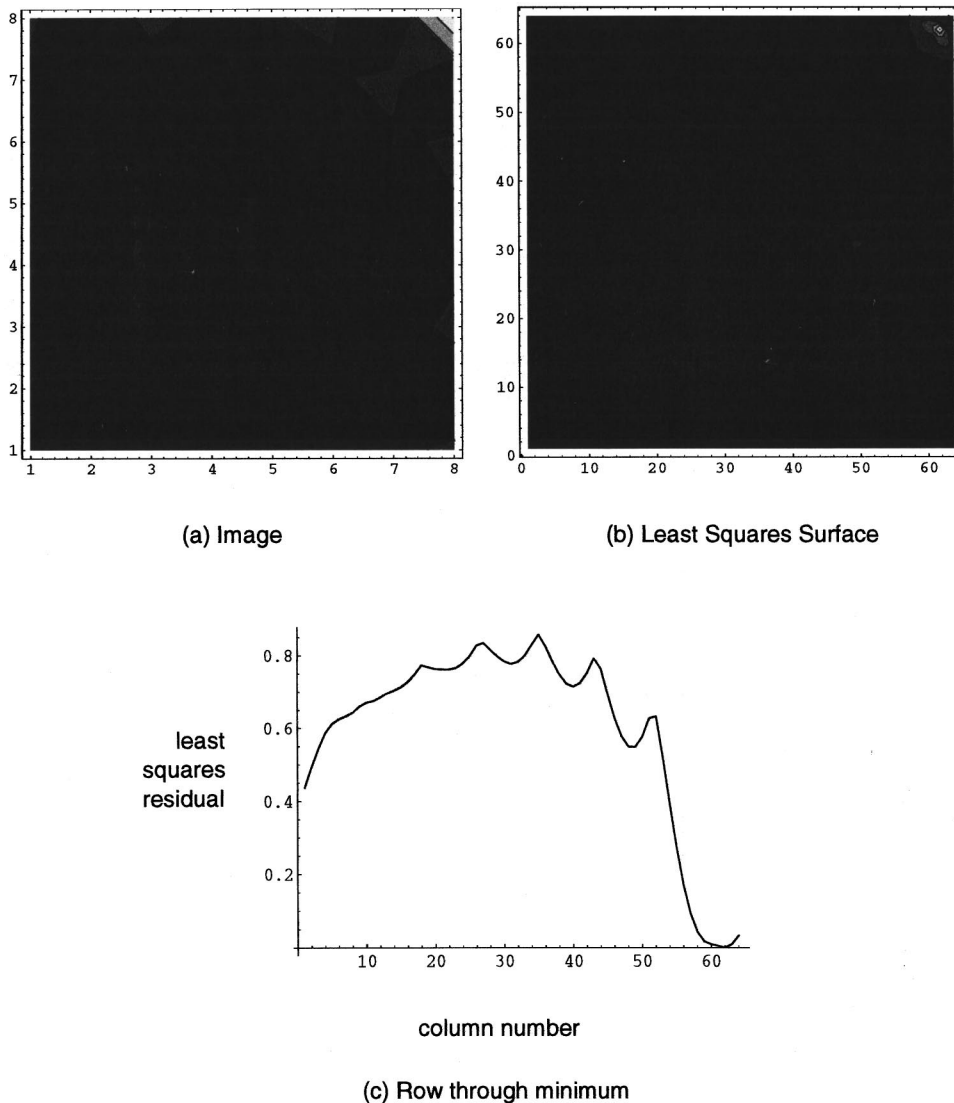


FIG. 5. A graph of the least squares error surface for radial target position (0.13, 0.13). (a) The  $8 \times 8$  image as collected by the sonar. (b) A contour diagram of the reciprocal least squares error surface (contours are 0.99, 0.5, 0.1, 0.05, 0.01 of the maximum value). (c) A one-dimensional slice through the error surface at the minimum error position.

weaker returns difficult, if not impossible for a given signal-to-noise level.

Quantitatively, the problem can be posed as a minimization of the same norm as above, except in this case, the minimization is taken between collected data and the superposition of the targets that are in the field of view of the system:

$$\text{minimize } \sum_{n=1}^N \sum_{i=1}^8 \sum_{j=1}^8 (s_{ijn} - a_{ijn})^2 \text{ for } i=\{1,8\}, j=\{1,8\}, \quad (15)$$

where  $N$  is the total number of targets.

One technique of potential use here is the EM algorithm (Moon, 1996). The algorithm guarantees (for a known number of targets) that the iteratively computed answer will converge to the maximum likelihood estimate for target position. In this procedure, the calculation is divided into both an estimation step and then a maximization step. Unfortunately in this case, the actual number of targets is not known *a priori*. Therefore, an additional step will need to be considered which consists of estimating the number of targets that are in

the field of view of the system. Future research will consider the inclusion of these additional features into the above proposed algorithm.

One way of discriminating against incorrect positional estimates due to multiple targets uses the fact that the presence of multiple targets in the system's field of view at the same range will produce data that do not permit the least squares residual to be refined to as small a value as in the case of a single target. Experiments with real data, which test to see if the least squares residual is greater than an empirically chosen value, have been successful in rejecting locations which are incorrect due to the presence of a number of targets.

The issue of computation time should also be addressed. In the simulations reported in this article, the Mathematica system of computer programs was used. Recently, the algorithm has been implemented in C++ on a Pentium II-266 computer. In order to speed the implementation, a local search algorithm was used which finds the maximum value in the two-dimensional array and does a local search around it of plus or minus one beam. Refined positional estimates for a single target take only several seconds for each of the

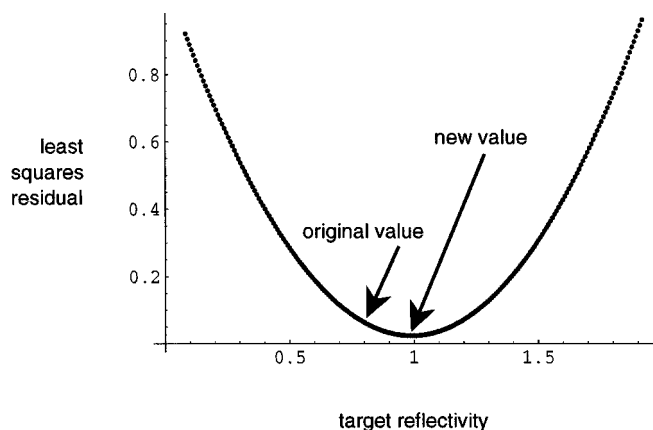


FIG. 6. The dependence of the least squares estimator on the reflectivity for the initial guess and the refined one. The graph shows the refinement procedure's initial guess and the subsequent least squares values.

targets. This has allowed us to use the algorithm to compute the locations of several thousand targets from experimental data.

Lastly, the effect of noise on the performance of the algorithm should be considered. In the simulations performed here, noise was not added to the data. On the other hand, inspection of the least squares surfaces should allow some prediction of the effect of noise on the performance of the algorithm. So, for example, the presence of noise with a given pdf (probability density function) will result in a set of positional estimates with a resultant pdf.

In conclusion, it has been demonstrated that the accurate estimation of both position and target strength for the targets in the field of view of the FishTV system can be accomplished via a procedure which minimizes a least squares norm between the predictions of a model and the observed data. The performance of the minimization procedure has been verified through computer simulations which demon-

strate for the noiseless case that (1) the true answer achieves a global minimum for the data, and that (2) the target strength of the reflector can be estimated via a knowledge of the estimated position. Future research on the problem will consider both the effects of noise and the presence of multiple unresolved targets in the field of view of the system.

- Baggeroer, A. B., Kuperman, W. A., and Schmidt, H. (1988). "Matched field processing: source localization in correlated noise as an optimum parameter estimation problem," *J. Acoust. Soc. Am.* **83**, 571–587.
- Barange, M., Hampton, I., and Soule, M. (1996). "Empirical determination of *in situ* target strengths of three loosely aggregated pelagic fish species," *ICES J. Mar. Sci.* **53**, 225–232.
- Frenkel, L., and Feder, M. (1995). "Recursive estimate-maximize (EM) algorithms for time varying parameters with applications to multiple target tracking," in *1995 International Conference on Acoustics, Speech, and Signal Processing, Detroit, MI, 9–12 May 1995* (IEEE, New York, 1995), pp. 2068–2071, Vol. 3.
- Goodman, J. (1985). *Statistical Optics* (Wiley, New York).
- Jaffe, J. S., Ohman, M. D., and DeRobertis, A. (1998). "OASIS in the sea: Measurement of the acoustic reflectivity of zooplankton with concurrent optical imaging," *Deep-Sea Res.* **45**, 1239–1243.
- Jaffe, J. S., Reuss, E., McGehee, D., and Chandran, G. (1995). "FTV, a sonar for tracking macrozooplankton in 3-dimensions," *Deep-Sea Res.* **42**, 1495–1512.
- Johnson, Don H., and Dudgeon, D. E. (1993). *Array Signal Processing: Concepts and Techniques* (Prentice-Hall, Englewood Cliffs, NJ).
- Kay, Steven M. (1993). *Fundamentals of Statistical Signal Processing: Estimation Theory* (Prentice-Hall, Englewood Cliffs, NJ).
- McGehee, D., and Jaffe, J. S. (1996). "Three-dimensional swimming behavior of individual zooplankters: Observations using the acoustical imaging system FishTV," *ICES J. Mar. Sci.* **53**, 363–369.
- Moon, T. K. (1996). "The expectation-maximization algorithm," *IEEE Signal Process. Mag.* **13**, 47–60.
- Papazoglou, M., and Krolik, J. L. (1996). "High resolution adaptive beamforming for three-dimensional acoustic imaging of zooplankton," *J. Acoust. Soc. Am.* **100**, 3621–3630.
- Papoulis, A. (1984). *Probability, Random Variable and Stochastic Processes*, 2nd ed. (McGraw-Hill, New York).
- Steinberg, B. D. (1976). *Principles of Aperture and Array System Design* (Wiley, New York).
- Urick, R. J. (1983). *Principles of Underwater Sound* (McGraw-Hill, New York).

# Iterative time reversal in the ocean

H. C. Song, W. A. Kuperman, and W. S. Hodgkiss

*Marine Physical Laboratory, Scripps Institution of Oceanography, University of California, San Diego,  
La Jolla, California 92093-0701*

T. Akal and C. Ferla

*SACLANT Undersea Research Center, 19138 La Spezia, Italy*

(Received 8 May 1998; accepted for publication 22 February 1999)

The iterative time-reversal process focusing on the strongest scatterer in a multitarget medium has been described theoretically in terms of eigenvalues and eigenvectors of a time-reversal operator  $\mathbf{K}^*\mathbf{K}$  in ultrasonics [Prada *et al.*, *J. Acoust. Soc. Am.* **97**, 62–71 (1995)]. In this paper, we extend the concept of iterative time-reversal to waveguide propagation in the ocean. For a single target, the iterative time-reversal process results in a minor improvement in spatial focusing. However, data from a recent experiment in the Mediterranean Sea [Kuperman *et al.*, *J. Acoust. Soc. Am.* **103**, 25–40 (1998)] illustrates the importance of the waveguide and source transducer characteristics even in the single target case. When the ocean contains several reflectors, iterative time-reversal focuses on the target corresponding to the largest eigenvalue of the time-reversal operator, which depends not only on the reflectivity of the targets, but also on the complex propagation effects between the targets and time-reversal mirror. Analysis of the experimental data for a single target and simulation results with multiple targets in the ocean are presented. © 1999 Acoustical Society of America. [S0001-4966(99)02306-1]

PACS numbers: 43.30.Yj, 43.30.Vh, 43.30.Wi, 43.30.Re [DLB]

## INTRODUCTION

Over the past several years, acoustic time-reversal mirrors (TRM) have been studied extensively in medical ultrasound,<sup>1–3</sup> nondestructive testing,<sup>4,5</sup> and ocean acoustics.<sup>6–7</sup> A good overview of TRM is provided in the recent paper by Fink.<sup>8</sup> Unlike an ordinary mirror that produces the virtual image of an acoustic object, the TRM produces a real acoustic image of the probe source (PS) by converting a divergent wave emitted from the acoustic PS into a convergent wave focusing on the PS. A TRM can be realized by a source-receive array. The incident signal is received, time reversed, and retransmitted from an array of sources collocated with the receivers.

When the medium contains several reflectors, the time-reversal process can be iterated in order to focus on the most reflective one, as demonstrated in ultrasonic laboratory acoustic experiments.<sup>9–11</sup> The theory of the iterative time-reversal mirror has been presented by Prada *et al.*<sup>11</sup> in terms of eigenvalues and eigenvectors of the time-reversal operator  $\mathbf{K}^*(f)\mathbf{K}(f)$ , where  $f$  is frequency,  $*$  denotes complex conjugation, and  $\mathbf{K}(f)$  is the transfer matrix of the array transducers ensonifying a time-invariant scattering medium. Prada *et al.* further extended the iterative time-reversal process to focus on a specific target by decomposition of the time-reversal operator called the DORT method.<sup>12</sup>

Iterative time reversal in the ocean has been demonstrated recently in an experiment conducted in the Mediterranean Sea.<sup>7</sup> In this paper, we revisit the experimental results for a single target and investigate further iterative time reversal in the ocean for multiple targets. In our analysis, boundary reverberation is not considered explicitly since we assume the waveguide boundaries are smooth.

In Sec. I, we review the theory of the iterative time-

reversal mirror developed in free-space ultrasonics. Section II describes the experimental results with a single target in the ocean. Section III presents simulation results in the ocean with two targets. Finally, conclusions are given in Sec. IV.

## I. ITERATIVE TIME REVERSAL

The theory of iterative time reversal has already been presented in Ref. 11. We review briefly the basics of iterative time reversal assuming pointlike targets and single scattering, and then derive the intensity of the sound field over the iterations. The pointlike scatterers are small compared to the wavelength and have a spherical response to the incident pressure field.

### A. Overview of theory

Assume that the array of transducers for the TRM consists of  $M$  similar elements and that the medium contains  $d$  pointlike scatterers with reflection coefficients  $C_1(f)$ ,  $C_2(f)$ , ...,  $C_d(f)$ . In order to express the received signals as a function of those transmitted, we define for each pair of transducers an interelement impulse response  $k_{lm}(t)$ , from element  $l$  to element  $m$  (Fig. 1). This impulse response includes all of the propagation effects through the medium under investigation as well as the acousto-electrical responses of the two elements. Let  $a_t(t)$  and  $a_r(t)$  be the transducers' acousto-electrical response in transmission and in reception with Fourier transforms  $A_t(f)$  and  $A_r(f)$ , respectively. For each frequency component  $f$ , a transfer matrix is defined as  $\mathbf{K}(f) = (K_{lm}(f))_{1 \leq l, m \leq M}$ , where  $K_{lm}(f)$  is a Fourier transform of  $k_{lm}(t)$ .

Assuming pointlike targets and single scattering,<sup>11</sup> the transfer matrix  $\mathbf{K}$  can then be written as

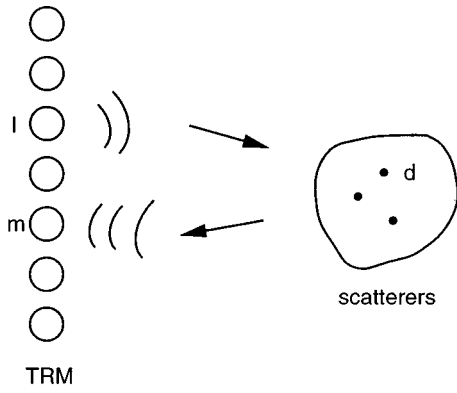


FIG. 1. Interelement impulse response  $k_{im}$ .

$$\mathbf{K}(f) = A_t(f)A_r(f) \sum_{i=1}^d C_i(f) \mathbf{H}_i(f) \mathbf{H}_i^T(f), \quad (1)$$

where  $\mathbf{H}_i = [H_{i1}, H_{i2}, \dots, H_{iM}]^T$  is a column vector of transfer functions between the  $i$ th scatterer and  $M$  elements of the TRM, and  $T$  denotes a transpose operation. Note that  $\mathbf{K}(f)$  is symmetrical due to reciprocity. For convenience, we drop the frequency dependence  $f$  and suppress the acousto-electrical responses  $A_t$  and  $A_r$  in most of the following analysis, since they can be absorbed into each of the transfer function  $\mathbf{H}_i$  as a scale factor, except in instances where the transducer characteristics must be considered.

Once we measure a transfer matrix  $\mathbf{K}$ , an iterative TRM can be described theoretically in terms of eigenvalues and eigenvectors of a Hermitian matrix  $\mathbf{K}^* \mathbf{K}$  referred to as a time-reversal operator in Refs. 10–12. Here, we confine our interests to the case of ideal separation of the reflectors<sup>11</sup> by assuming the orthogonality of transfer functions

$$\mathbf{H}_i^T \mathbf{H}_j^* = 0 \quad \text{for } i \neq j. \quad (2)$$

This assumption is valid, for example, for reflectors at the same range but different depths in an ideal waveguide due to the ‘‘closure’’ or completeness property of the orthonormal modal functions if the source–receive array (SRA) spans most of the water column and adequately samples most of the modes so that the orthogonality condition is satisfied.<sup>7</sup> In practice, however, there are only a finite number of propagating modes in an oceanic waveguide so that Eq. (2) is satisfied only approximately, resulting in depth resolution of  $D/N$  where  $D$  is the water depth and  $N$  is the number of propagating modes.<sup>13</sup>

In case of ideal separations, the Hermitian operator can be decomposed into

$$\mathbf{K}^* \mathbf{K} = \sum_{i=1}^d |C_i|^2 |\mathbf{H}_i|^4 \frac{\mathbf{H}_i^* \mathbf{H}_i^T}{|\mathbf{H}_i|^2}, \quad (3)$$

where the eigenvectors are  $\mathbf{H}_i^*$  and the corresponding eigenvalues  $\lambda_i$  including the acousto-electrical responses are

$$\lambda_i = |A_t A_r|^2 |C_i|^2 |\mathbf{H}_i|^4. \quad (4)$$

Note that each eigenvector of the operator  $\mathbf{H}_i^*$  is associated with one of the pointlike scatterers, exactly the vector signal after time reversal for one scatterer considered separately.

The eigenvalues are positive because the operator  $\mathbf{K}^* \mathbf{K}$  is Hermitian. The eigenvalues depend both on the reflectivity  $C_i$  and on the propagation effects  $|\mathbf{H}_i|$  which could be more significant in a complex oceanic waveguide due to the 4th power as compared to the 2nd power in  $C_i$  as revealed in the following sections. The impact of the acousto-electrical responses of the transducers on the iterative time-reversal mirror also will be addressed in Sec. II.

## B. Intensity at the $n$ th iteration

In this section, we derive the intensity of the field in terms of eigenvalues of the time-reversal operator  $\mathbf{K}^* \mathbf{K}$  for ideally resolved scatterers. It is shown in Ref. 11 that the pressure field resulting from an iterative process converges differently toward odd and even limits. In ultrasonics, the effect of focusing has been displayed in the vicinity of targets by the maximum magnitude of the pressure field. The probe pulse for ultrasonic laboratory experiments typically is a half cycle of the carrier frequency around 3.5 MHz.<sup>8</sup> However, in our ocean acoustic experiment where a 22-cycle pure-tone pulse with carrier frequency of 445 Hz was used,<sup>7</sup> we can better represent the focusing effect in terms of energy over a segment of data rather than the pressure itself.

Let  $\mathbf{P}[0]$  be the complex pressure vector received by each scatterer position after initial ensonification of the field by an input vector on the TRM,  $\mathbf{E}[0]$ , such that  $\mathbf{P}[0] = A_r \mathbf{H}^T \mathbf{E}[0]$  where  $\mathbf{H} = [\mathbf{H}_1, \mathbf{H}_2, \dots, \mathbf{H}_d]$  is an  $M \times d$  transfer function matrix. The pressure field of the scatterers at the iteration number  $n$ ,  $\mathbf{P}[n]$ , can then be related to the pressure field at the iteration number  $n-1$ ,  $\mathbf{P}[n-1]$ , as follows (see Fig. 1 in Ref. 12):

$$\mathbf{P}[n] = \mathbf{H}^T (\mathbf{H} \mathbf{C} \mathbf{P}[n-1])^* = \mathbf{H}^T \mathbf{H}^* \mathbf{C}^* \mathbf{P}^*[n-1], \quad (5)$$

where  $\mathbf{C} = \text{diag}(C_1, C_2, \dots, C_d)$  is a diagonal matrix due to the single scattering assumption. Since the columns of  $\mathbf{H}$  are orthogonal to each other for ideal separations of the reflectors as shown in Eq. (2), the intensity of the field by  $i$ th scatterer at iteration number  $n$  simply reduces to

$$I_i[n] = |C_i|^2 |\mathbf{H}_i|^4 I_i[n-1] = \lambda_i I_i[n-1] = \lambda_i^n I_i[0], \quad (6)$$

where  $I_i[0] = |P_i[0]|^2$ . For broadband pulse propagation in a time-reversal mirror, a total energy can be obtained using Parseval’s relation over the frequency band of the signal,

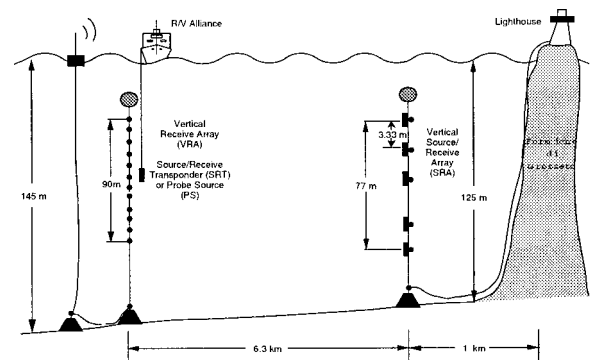


FIG. 2. Experimental setup of April 1996 phase conjugation experiment.



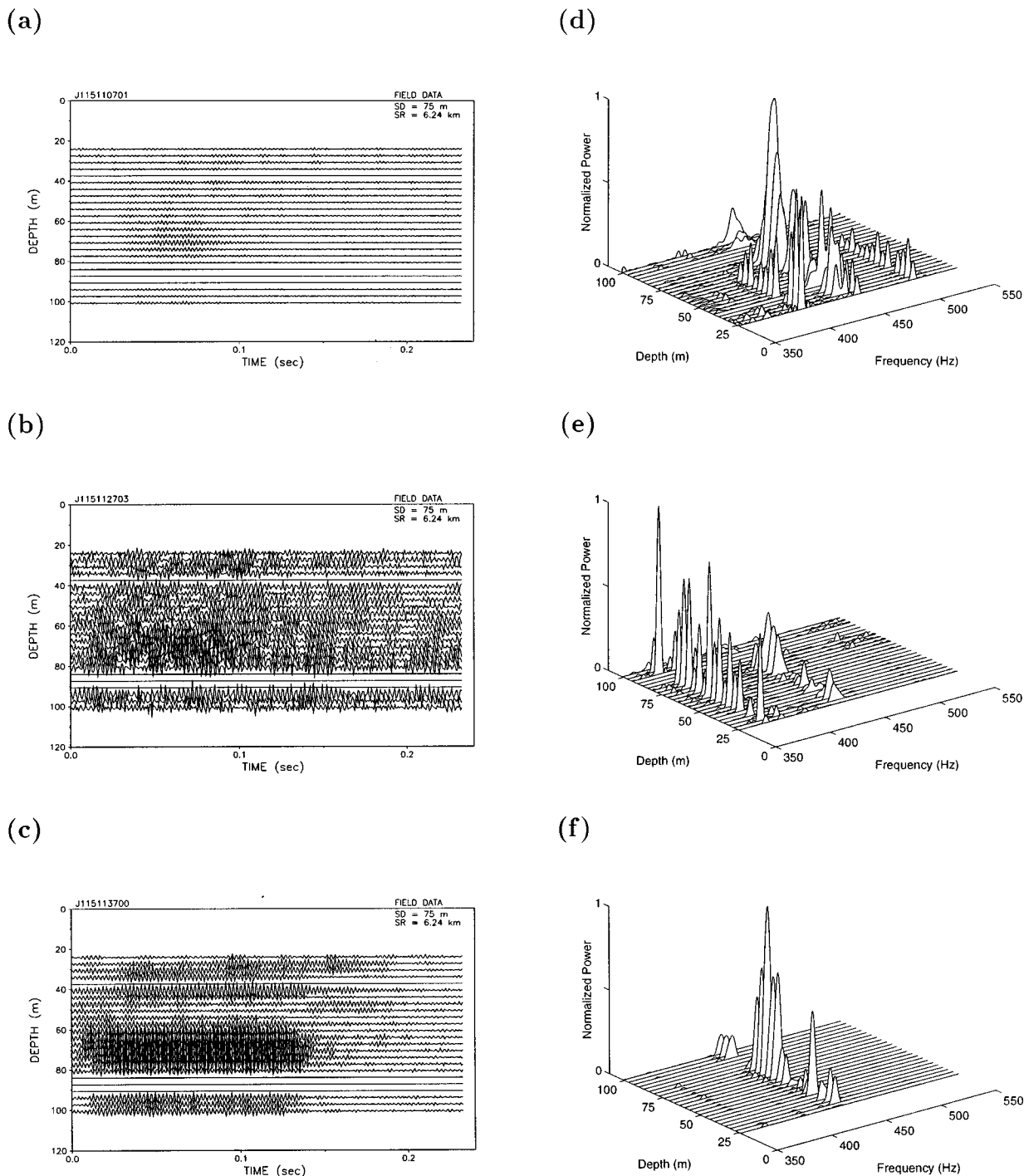


FIG. 3. Experimental results from iterative time reversal received on the SRA for a simulated scatterer (SRT) located at 75 m depth and 6.24 km range from the SRA. (a)–(c) display the pulse data received on the SRA at iterations #1, #11, and #16, respectively. (d)–(f) are the corresponding normalized spectra. Note that four elements of the SRA were dead during the experiment. Typically, there are three spectral-noise components around 410, 460, and 510 Hz, as shown in (d). The dominant noise component around 370 Hz in (e) is due to a small freighter passing 1 km away from the SRA.

$$E_i[n] = \int \lambda_i^n(f) I_i[0](f) df. \quad (7)$$

## II. ITERATIVE PHASE CONJUGATION IN THE OCEAN

Phase conjugation as applied to underwater acoustics has been explored theoretically<sup>6</sup> and demonstrated experimentally in the ocean.<sup>7</sup> Here, we describe the implementation of iterative phase conjugation using the geometry of the

TRM experiment shown schematically in Fig. 2 and then analyze the experimental results. More details on the theory and experiment are given in Ref. 7.

### A. Experimental setup: Acoustic Ping-Pong

The April 1996 TRM experiment utilized a vertical source–receive array (SRA) spanning 77 m of a 125-m water column with 20 sources and receivers and a single source/receive transponder (SRT) (echo repeater) or probe source

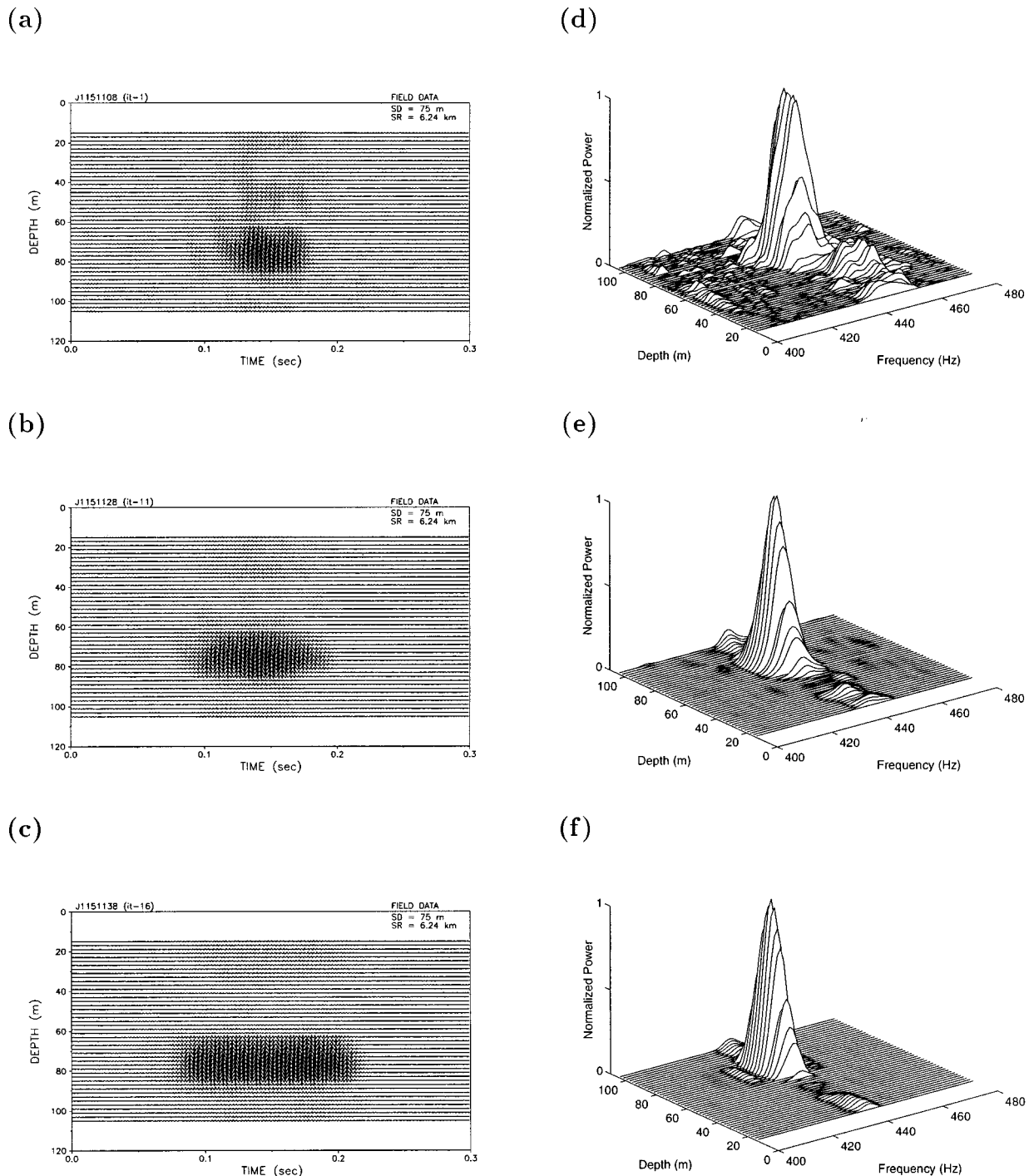


FIG. 4. Experimental results from iterative time reversal received on the VRA for a simulated scatterer (SRT) located at 75 m depth and 6.24 km range from the SRA. (a)–(c) display the pulse data received on the VRA from the time-reversed transmissions of the pulses shown in Fig. 3(a)–(c) at iterations #1, #11, and #16, respectively. (d)–(f) are the corresponding normalized spectra.

(PS) collocated in range with another vertical array (VRA) of 46 elements spanning 90 m of a 145-m water column located 6.3 km away from the SRA. Phase conjugation was implemented by transmitting a 50-ms pulse with center frequency of 445 Hz from the PS to the SRA, digitizing the received signal, and retransmitting the time-reversed signal from all sources of the SRA. The transmitted signal was then received at the VRA. An assortment of runs was made to examine the structure of the focal point region and the temporal stability of the process. These results are reported in Ref. 7.

For the iterative phase conjugation experiment, the probe source with collocated receiver (an element of the VRA) now acts as a simulated scatterer (SRT). The iterative time-reversal process is initiated by transmitting a 50-ms pulse with equal amplitudes on all elements from the SRA to ensonify the waveguide. The transmission is captured at 75 m depth by the SRT and retransmitted (echoed) back to the SRA. The SRA time reverses the received signal and retransmits it from all the SRA sources back to the VRA. This acoustic Ping-Pong process was repeated many times over an

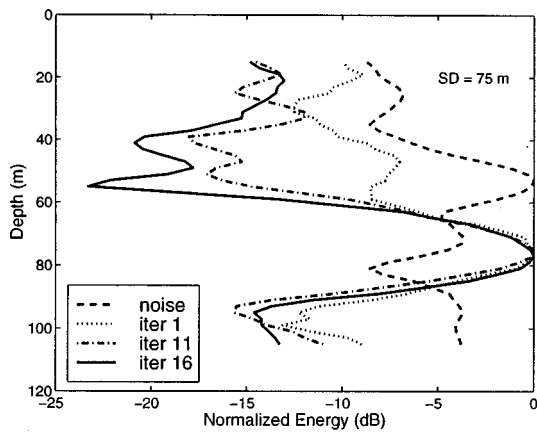


FIG. 5. Acoustic Ping-Pong results between the SRA and a transponder at 75 m depth and 6.24 km range. Displayed is the energy over a 0.3-s window on the VRA as a function of depth for iterations #1, #11, and #16. The dashed line indicates the energy of the simulated backpropagation of the noise shown in Fig. 8(b).

hour with 2 min between each round trip. Here, we report on the experimental data obtained on JD 115 (24 April 1996) from 11:06 to 11:38 Z for a total of 17 successive iterations.

## B. Experimental results

Figure 3 shows the pulse data and corresponding spectra received on the SRA for iterations #1, #11, and #16. The simulated scatterer (SRT) was located at 75 m depth and 6.24 km range. Here, we refer to iteration #1 as the first time-reversal operation after initial ensonification. The data are a combination of signal and noise. A 233-ms segment of data was digitized and time reversed for retransmission to the VRA at each iteration. Note in the time-series displays that the signal level increases over the iterations due to an iterative loop gain being greater than 1. Also note in Fig. 3(b) that the signal is hardly visible due to shipping noise around 370 Hz. Typically, there are three spectral-noise components around 410, 460, and 510 Hz, as shown in Fig. 3(d). As will be seen, these noise components turn out to be an important factor in the analysis of the experimental data.

Figure 4 shows the time-series data and corresponding spectra as received on the VRA from the time-reversed transmission of the pulses shown in Fig. 3. The time series are normalized with respect to the maximum amplitude as opposed to Fig. 3, where no normalization has been applied. Two observations can be made. At the first iteration shown in Fig. 4(a), we clearly see a vertical spatial focusing of  $\pm 15$  m around the 75-m SRT depth as well as a temporal compression to 50 ms (the original pulse length) plus the time spread of the sound channel due to the initial ensonification by the SRA which is not significant, as will be seen in Fig. 11(a). Note, however, that the time series duration expands significantly over the sequence of iterations. Expansion of the signal duration is confirmed by the more narrow bandwidth of the corresponding spectra over the sequence of iterations on the right side of Fig. 4. Second, Fig. 4(b) shows very good spatial focusing in the presence of the dominant shipping noise around 370 Hz, as shown in Fig. 3(b).

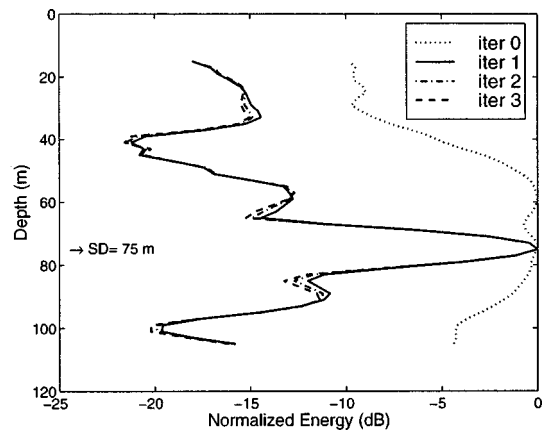


FIG. 6. Simulation results for a single target at 75 m depth. The energy over a 0.3-s window as a function of depth on the VRA for iterations #1 to #3 shows no significant improvement in focusing over the iteration process. The dotted line indicates the ensonified field resulting from transmitting a 50-ms pulse with equal amplitudes on all elements of the SRA. Note that the field decreases more rapidly towards the surface compared to the bottom, due to the downward-refracting sound-speed profile employed in the model.

An alternative way to display the VRA data is to plot the energy over a 0.3-s window as a function of depth and iteration. Due to the sidelobe suppression between 40–60 m above the SRT depth of 75 m, Fig. 5 appears to indicate an increase in focusing over the sequence of iterations. However, the iterative simulation results for a single target in Fig. 6 suggest only a minor improvement in spatial focusing. The environmental model used for the simulations throughout this paper is displayed in Fig. 7. Note that Fig. 6 also indicates the degree to which orthogonality of the transfer functions as described in Eq. (2) is satisfied.

The apparent improvement in spatial focusing is due to the noise components observed on the SRA (especially around 460 Hz, close to the carrier frequency of 445 Hz). As mentioned earlier, typically there are three noise components around 410, 460, and 510 Hz embedded in the observed data on the SRA. A typical noise-alone time series observed on the SRA is shown in Fig. 8(b), along with its corresponding normalized spectrum in Fig. 8(d). The result of backpropagating the noise-alone time series to the VRA in simulation

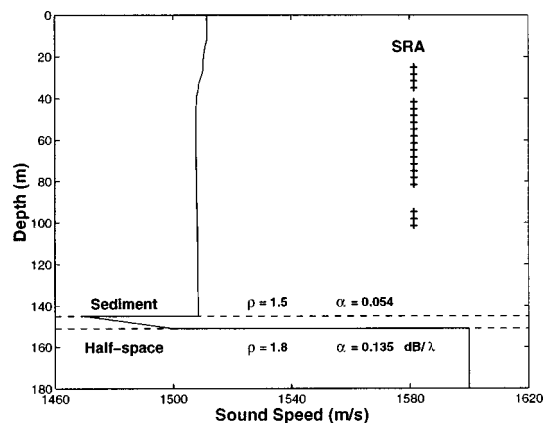


FIG. 7. Sound-speed profile used in the simulations along with the bottom geoacoustic parameters. The depths of the SRA elements are also indicated.

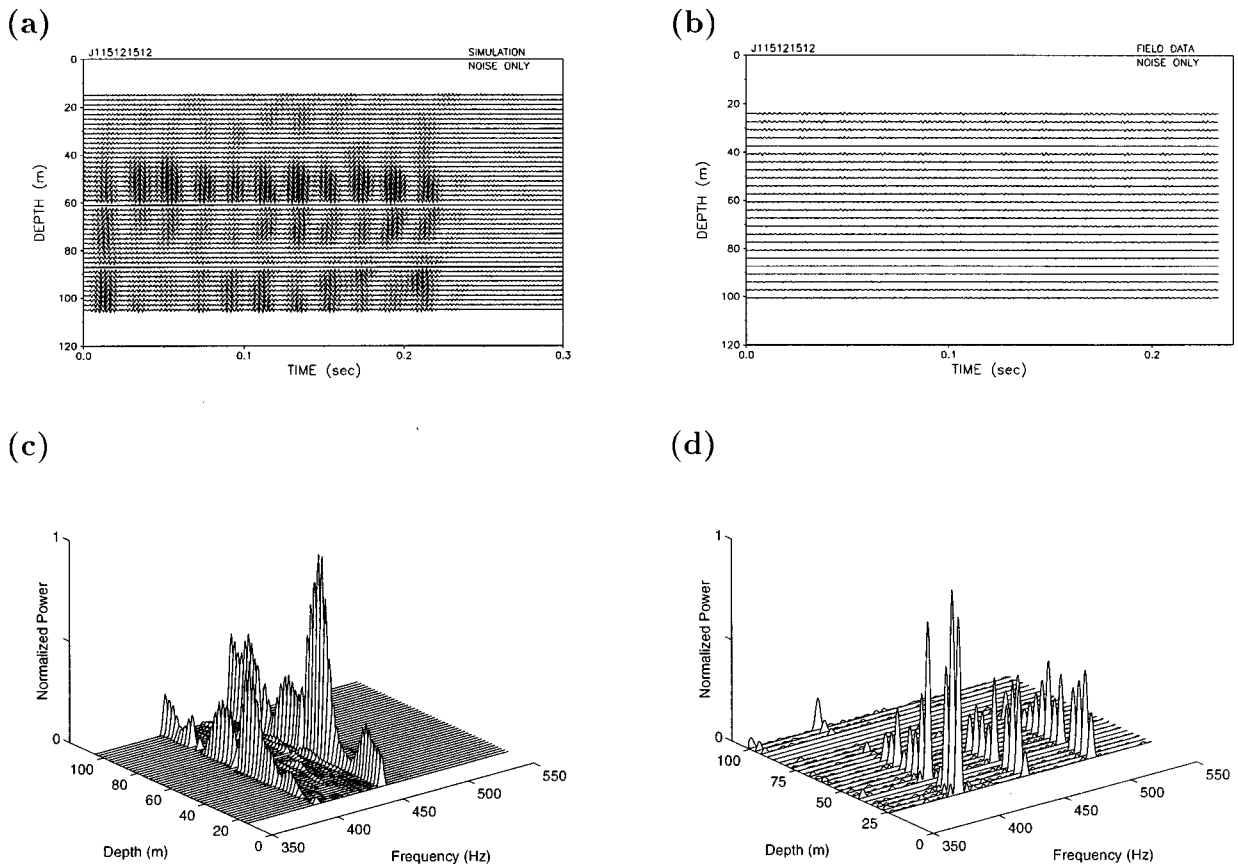


FIG. 8. Simulated backpropagation (a) and corresponding spectrum (c) using noise-alone data observed on the SRA shown in (b). The three noise components are shown in (d) around 410, 460, and 510 Hz. Note that the spectral content in (c) extends only up to 480 Hz because the higher-frequency component is attenuated substantially by the SRA transducer characteristics.

and the corresponding spectrum are displayed in Fig. 8(a) and (c), respectively.

Note that the backpropagated noise contributes substantially to the field between 40 and 60 m depth, as is also shown by the dashed line in Fig. 5. At the beginning of the iteration process, this backpropagated noise field degrades the apparent spatial focusing significantly, as shown in Fig. 5. Over the sequence of iterations, however, the noise level becomes negligible compared to the signal level. The signal level increases due to the iterative loop gain being greater than 1 (see Fig. 3). Thus, after a number of iterations we attain a spatial focusing that would be similar to the spatial focusing without noise, as shown in Fig. 6.

Aside from the noise issues, Fig. 9 shows the normalized spectra (dB) at the target depth of 75 m for iterations #1, #11, and #16 selected from (d)–(f) in Fig. 4. It is interesting to note a 5-Hz shift in the center frequency from 445 to 440 Hz over the sequence of iterations as well as the more narrow bandwidth described earlier.

As derived in Sec. IB, the intensity of the sound field at a single target after the  $n$ th iteration is  $I_1[n] = \lambda^n(f) I_1[0]$ , with the initial ensonification of  $I_1[0] = |A_r|^2 |\mathbf{H}_1^T \mathbf{E}[0]|^2$ . The initial input vector  $\mathbf{E}[0]$  is equal to  $S(f) \mathbf{I}$ , where  $S(f)$  is the source spectrum of the pulse initially transmitted from the SRA and  $\mathbf{I}$  is a unit vector. The expression for the eigenvalue of the time-reversal operator  $\lambda$  is given by Eq. (4), which depends upon the SRA source transducer functions  $A_t(f)$  and  $A_r(f)$ , the medium propagation characteristics  $\mathbf{H}_1(f)$ ,

and the target reflectivity  $C_1$ . In the experimental data, the SRT (echo repeater) transducer was identical to those used in the source array (SRA). Accordingly,  $C_1$  in Eq. (4) needs to be multiplied by an additional factor of  $A_t(f)$ . On the other hand, the receiver hydrophones of the SRA have an almost flat frequency response over the frequency band involved [i.e.,  $A_r(f) = \text{const}$ ].

Figure 10 compares the characteristics of the three functions involved:  $|\mathbf{H}(f)|$ ,  $|A_t(f)|$ , and  $I[0]$ . Note the fourth

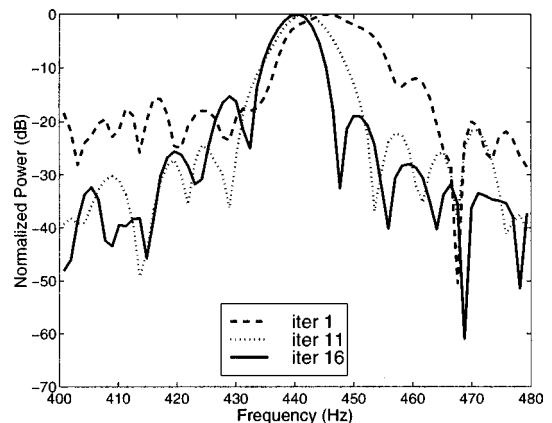


FIG. 9. Normalized spectra (dB) at the target depth of 75 m for iterations #1, #11, and #16 from (d)–(f) in Fig. 4. Note the more narrow bandwidth and shift in the center frequency from 445 to 440 Hz over the sequence of iterations.

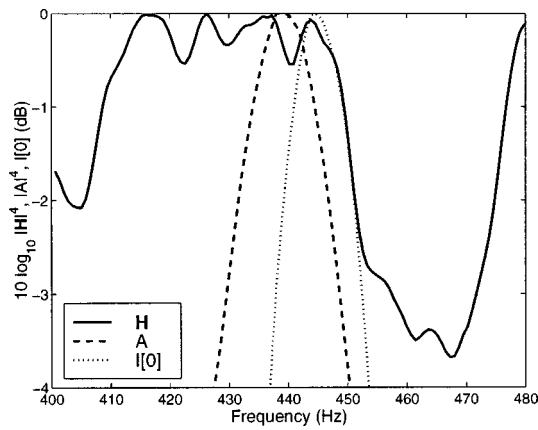


FIG. 10. The medium propagation characteristics  $|\mathbf{H}(f)|$  (solid) between the target at 75 m depth and the SRA, the source transducer characteristics  $|A_s(f)|$  (dashed), and the initial ensonification at the 75-m target depth  $I[0](f)$  (dotted) determined mainly by the source signal spectrum  $|S(f)|$ .

power of  $|\mathbf{H}(f)|$  and  $|A_s(f)|$  as compared to the initial ensonification  $I[0]$  to emphasize their contribution at each iteration. As shown by the dotted line,  $I[0]$  looks similar to the source spectrum  $|S(f)|$ , whose center frequency is 445 Hz. The dashed line displays a source transducer function  $|A_s(f)|$ , which has a resonance at approximately 440 Hz and a 3-dB bandwidth of approximately 40 Hz. The magnitude of the complex transfer function vector  $|\mathbf{H}(f)|$  shows a band-pass characteristic with a cutoff beyond 450 Hz. Multiplying these transfer functions at the SRA and SRT repetitively over the sequence of iterations results in a narrower bandwidth of the signal as well as a 5-Hz shift in the center frequency. This frequency shift should be distinguished from a Doppler shift due to either source or receiver motion.

### III. MULTIPLE TARGETS IN THE OCEAN

In this section, we investigate the case of multiple targets in the ocean through simulations. For simplicity, we consider two pointlike targets at the same range (6.3 km), but different depths of 40 and 75 m, denoted by T1 and T2, respectively. Without loss of generality, we assume a flat frequency response of the transducers. Noise is not included in this simulation.

First, consider two targets with equal strength, i.e.,  $C_1 = C_2$ . As in the single-target case, we ensonify the acoustic waveguide by transmitting 50-ms pure-tone pulse with equal amplitude from all sources of the SRA to the VRA. Figure 11(a) shows the initial ensonified field  $I[0]$  at the VRA, which is not homogeneous with depth as opposed to the free-field case. As noted earlier, pulse spreading is not significant mainly due to the narrow bandwidth of the pulse employed. Note also that we have much weaker energy near the surface due to the downward-refracting sound-speed profile included in the environmental model.

Figure 11(b) and (c) show results after the first and third iterations, respectively. Note that target T1 at the shallower depth is barely visible after the first iteration and disappears almost completely after the third iteration. This observation can be explained from Eqs. (6) and (7). In our example, the ratio of the ensonified field  $I_1[0]/I_2[0] \cong 0.24$  and the ratio

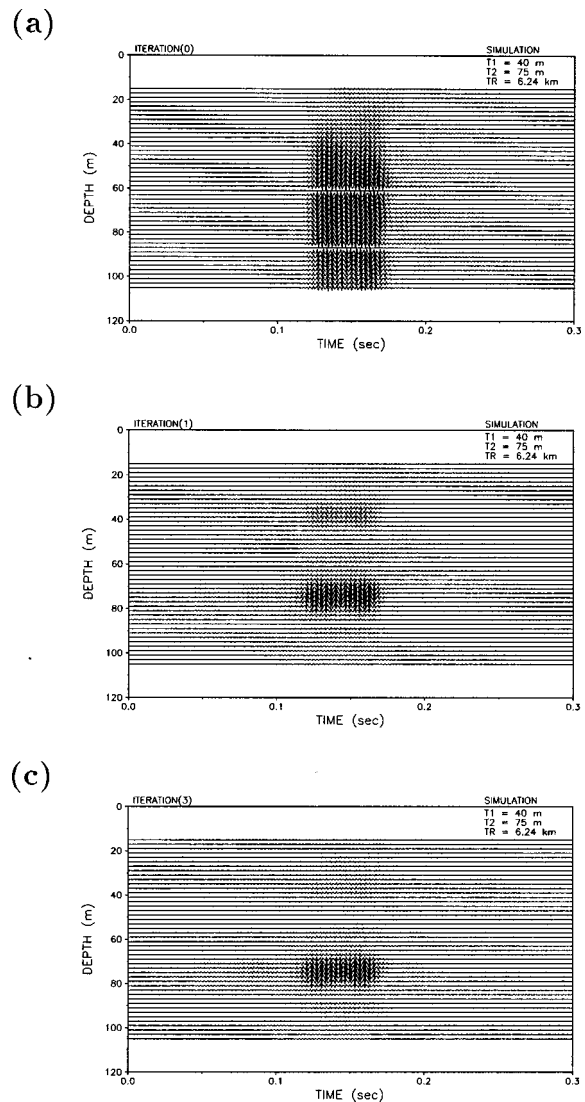


FIG. 11. Simulation results for two pointlike targets with equal reflectivities at depths of 40 and 75 m. (a) Ensonified field on the VRA by transmitting a 50-ms pulse with equal amplitudes from all sources of the SRA. (b) Results after the first iteration. (c) Results after the third iteration. Note that even after the first iteration, the target T1 at shallower depth is hardly visible.

of eigenvalues  $\lambda_1/\lambda_2 = |\mathbf{H}_1|^4/|\mathbf{H}_2|^4 \cong 0.44$  at the center frequency of 445 Hz for targets with the same reflectivity. As shown in Fig. 12, after one iteration the intensity of T1 drops almost 7–8 dB with respect to the level of T2 when integrated over the frequency band.

This example demonstrates that an iterative time-reversal process in an underwater waveguide does not necessarily focus on the most reflective target. The “effective” reflectivity includes the propagation effects between the targets and the SRA. The target with the largest effective reflectivity is the one which corresponds to the largest eigenvalue of the time-reversal operator.

The next question then is, under what condition can we focus on the shallow target T1. Since the propagation effects are determined by the relative position of the targets with respect to the SRA, the target strength of T1 must be strong enough to compensate the smaller  $|\mathbf{H}_1|$  as compared to  $|\mathbf{H}_2|$  such that  $\lambda_1/\lambda_2 > 1$ . We already have computed that the ratio

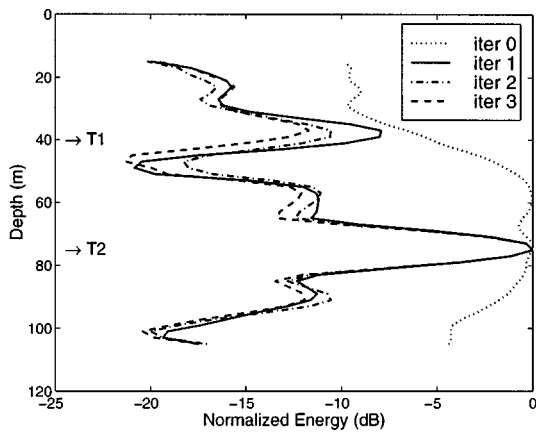


FIG. 12. Simulation results for two targets with equal reflectivities at depths of 40 and 75 m. The energy over a 0.3-s window as a function of depth on the VRA is plotted for iterations 0 to 3. The dotted line indicates the initial ensonified field.

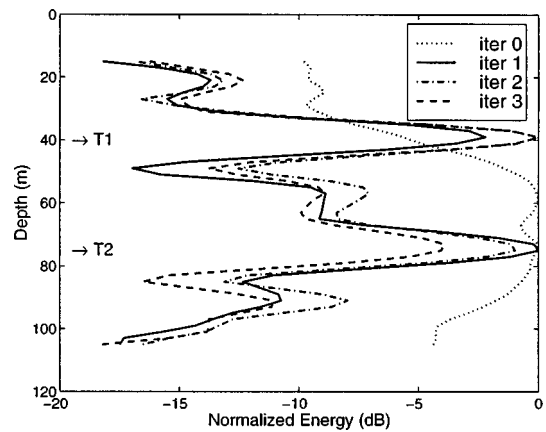


FIG. 14. Simulation results for two targets with unequal reflectivities at depths of 40 and 75 m. The energy over a 0.3-s window as a function of depth on the VRA is plotted for iterations 0 to 3. The dotted line indicates the initial ensonified field.

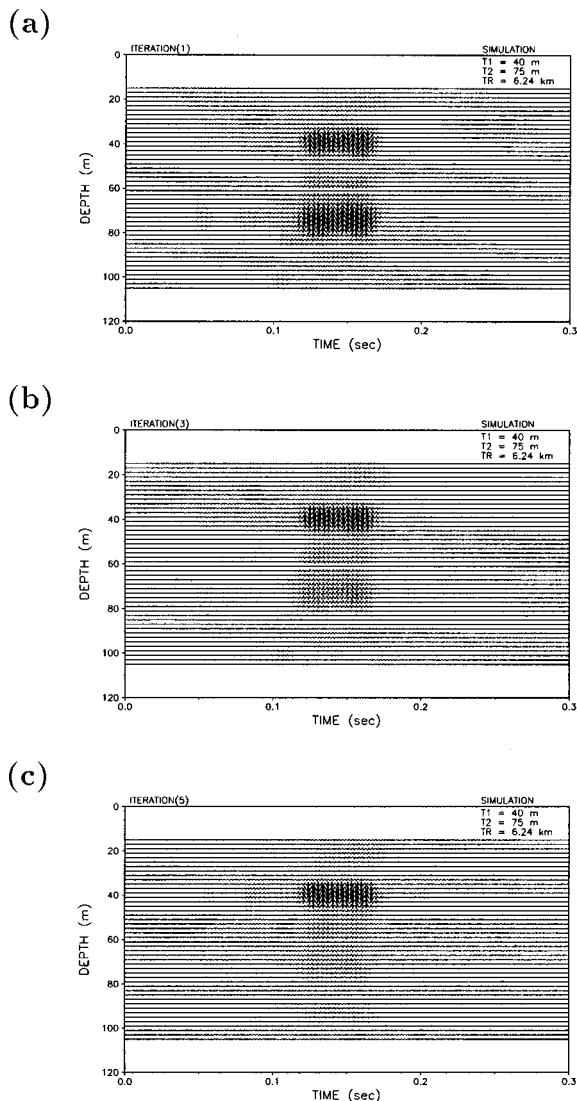


FIG. 13. Simulation results for two pointlike targets with unequal reflectivities at depths of 40 and 75 m. The ratio of reflectivities is  $C_1/C_2=2$ . (a), (b), and (c) are the results after iterations #1, #3, and #5, respectively. We see two targets after the first iteration. After five iterations, T1 dominates the field at the VRA.

of the eigenvalues is  $\lambda_1/\lambda_2 \approx 0.44$  for the same reflectivities. Then, from Eq. (4) the target strength of T1 must be increased to greater than  $1/\sqrt{0.44} \approx 1.5$ . Figure 13 shows simulation results of iterative time reversal when the target strength at s.d.=40 m is twice that at s.d.=75 m, i.e.,  $C_1/C_2=2$ . After the first iteration, the field strength at 75 m depth is still greater than that at 40 m, as shown in Fig. 14. However, successive iterations yields the stronger field at s.d.=40 m. By iteration 5, we clearly see that the field focuses at s.d.=40 m.

#### IV. CONCLUSIONS

The concept of an iterative time-reversal mirror has been extended to waveguide propagation in the ocean. For a single target, the iterative time-reversal process results in a minor improvement in spatial focusing. However, analysis of data from a recent experiment in the Mediterranean Sea did illustrate the importance of the waveguide and source transducer characteristics even in the single transducer case. Their combined effects through repeated multiplication over the iteration process resulted in narrowing of the pulse bandwidth along with a 5-Hz shift in the center frequency.

When we have multiple targets in free space, iterative time reversal selects out the most reflective target which corresponds to the largest eigenvalue of the time-reversal operator. In the waveguide case with smooth boundaries, the eigenvalues of the time-reversal operator are a function of the target reflectivities as well as the complex propagation characteristics of the medium. As shown in Eq. (4), the propagation characteristics are very important in an oceanic waveguide. Simulations with two targets at the same range but different depths were used to illustrate both of these effects on focusing.

#### ACKNOWLEDGMENTS

This research was supported by the Office of Naval Research Code 321US, Contract Nos. N00014-96-I-0065 and N00014-97-D-0003.

- <sup>1</sup>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).
- <sup>2</sup>M. Fink, "Time-reversal mirrors," *J. Phys. D* **26**, 1330–1350 (1993).
- <sup>3</sup>C. Dorme and M. Fink, "Focusing in transmit–receive mode through inhomogeneous media: The time-reversal matched filter approach," *J. Acoust. Soc. Am.* **98**, 1155–1162 (1995).
- <sup>4</sup>J. L. Thomas, P. Roux, and M. Fink, "Inverse scattering analysis with an acoustic time-reversal mirror," *Phys. Rev. Lett.* **72**, 637–640 (1994).
- <sup>5</sup>N. Chakroun, M. Fink, and F. Wu, "Time reversal processing in ultrasonic nondestructive testing," *IEEE Trans. Ultrason. Ferroelectr. Freq. Control* **42**, 1087–1098 (1995).
- <sup>6</sup>D. R. Jackson and D. R. Dowling, "Phase conjugation in underwater acoustics," *J. Acoust. Soc. Am.* **89**, 171–181 (1991).
- <sup>7</sup>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).
- <sup>8</sup>Mathias Fink, "Time-reversed acoustics," *Phys. Today* **50**, 34–40 (1997).
- <sup>9</sup>C. Prada, F. Wu, and M. Fink, "The iterative time reversal mirror: A solution to self-focusing in the pulse echo mode," *J. Acoust. Soc. Am.* **90**, 1119–1129 (1991).
- <sup>10</sup>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).
- <sup>11</sup>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).
- <sup>12</sup>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).
- <sup>13</sup>E. C. Shang, "Source depth estimation in waveguides," *J. Acoust. Soc. Am.* **77**, 1413–1418 (1985).

# A test of basin-scale acoustic thermometry using a large-aperture vertical array at 3250-km range in the eastern North Pacific Ocean

Peter F. Worcester, Bruce D. Cornuelle, Matthew A. Dzieciuch, and Walter H. Munk  
*Scripps Institution of Oceanography, University of California at San Diego, La Jolla, California 92093*

Bruce M. Howe, James A. Mercer, and Robert C. Spindel  
*Applied Physics Laboratory, University of Washington, Seattle, Washington 98105*

John A. Colosi  
*Woods Hole Oceanographic Institution, Woods Hole, Massachusetts 02543*

Kurt Metzger and Theodore G. Birdsall  
*Communications and Signal Processing Laboratory, Electrical Engineering and Computer Science Department, University of Michigan, Ann Arbor, Michigan 48109*

Arthur B. Baggeroer  
*Massachusetts Institute of Technology, Cambridge, Massachusetts 02139*

(Received 12 November 1997; revised 14 January 1999; accepted 5 March 1999)

Broadband acoustic signals were transmitted during November 1994 from a 75-Hz source suspended near the depth of the sound-channel axis to a 700-m long vertical receiving array approximately 3250 km distant in the eastern North Pacific Ocean. The early part of the arrival pattern consists of raylike wave fronts that are resolvable, identifiable, and stable. The later part of the arrival pattern does not contain identifiable raylike arrivals, due to scattering from internal-wave-induced sound-speed fluctuations. The observed ray travel times differ from ray predictions based on the sound-speed field constructed using nearly concurrent temperature and salinity measurements by more than *a priori* variability estimates, suggesting that the equation used to compute sound speed requires refinement. The range-averaged ocean sound speed can be determined with an uncertainty of about 0.05 m/s from the observed ray travel times together with the time at which the near-axial acoustic reception ends, used as a surrogate for the group delay of adiabatic mode 1. The change in temperature over six days can be estimated with an uncertainty of about 0.006 °C. The sensitivity of the travel times to ocean variability is concentrated near the ocean surface and at the corresponding conjugate depths, because all of the resolved ray arrivals have upper turning depths within a few hundred meters of the surface. © 1999 Acoustical Society of America. [S0001-4966(99)04506-3]

PACS numbers: 43.30.Pc, 43.30.Cq [SAC-B]

## INTRODUCTION

Broadband acoustic signals were transmitted during November 1994 from a 75-Hz source to a 700-m-long vertical line array of hydrophones approximately 3250 km distant in the eastern North Pacific Ocean as part of the Acoustic Engineering Test (AET) of the Acoustic Thermometry of Ocean Climate (ATOC) project. The overall goals were (i) to determine the accuracy with which gyre- and basin-scale ocean temperature and heat content variability can be measured using the methods of ocean acoustic tomography, (ii) to determine the vertical resolution that can be obtained at multimegahertz range in a single vertical slice containing an acoustic source and receiver, and (iii) to understand the influences that smaller-scale processes like internal waves and mesoscale eddies have on the signals. This paper focuses on the accuracy and resolution with which gyre-scale temperature and heat content can be measured using acoustic methods. A companion paper by Colosi *et al.* (1999) focuses on

internal-wave effects and the potential predictability of these effects using analytic acoustic fluctuation theories based on internal-wave dominance.

A number of previous experiments have firmly established the accuracy and resolution with which tomographic methods can measure ocean temperature and current at ranges of up to about 1000 km (Munk *et al.*, 1995). There have been only a few previous experiments in which broadband acoustic signals have been transmitted over multimegahertz ranges, however, and these experiments have left a number of issues unresolved, as discussed below.

In an experiment conducted intermittently from 1983-1989, a broadband 133-Hz source mounted on the bottom at 183-m depth near Kaneohe, Hawaii, transmitted roughly 3.7 Mm to a bottom-mounted horizontal array at 1433-m depth offshore of California [see Spiesberger and Tappert (1996), and the papers referenced therein]. Interpretation of the received signals was complicated because all of the energy reflected one or more times from the slope near Oahu before becoming trapped in the sound channel. No individual, re-



solved ray arrivals were observed. Rather, five distinct and relatively stable arrivals found near the beginning of complex receptions lasting several seconds were shown to be ray arrival *clusters*, each containing many predicted ray arrivals. All five ray clusters consisted of steep rays inclined at about 15° to the horizontal at the depth of the sound-channel axis. The accuracy of the interpretation of the observed travel-time variability of the ray clusters purely in terms of ocean temperature changes is somewhat uncertain. As ocean conditions change, the locations at which the rays interact with the bottom will change. The bathymetry at the locations ensounded by the rays is not well known, and so the extent to which changes in path length of the bottom-reflected rays might be contributing to the observed travel-time changes is also uncertain.

A second experiment conducted in 1987 showed that individual ray arrivals associated with very steep ray paths are resolvable, identifiable, and stable at 3-4-Mm ranges, when bottom interactions are unimportant. In this experiment, transmissions from a broadband 250-Hz source moored near the depth of the sound-channel axis in the central North Pacific Ocean (Dushaw *et al.*, 1993) to a bottom-mounted horizontal array in deep water about 3 Mm distant were analyzed in detail (Spiesberger and Metzger, 1991; Spiesberger *et al.*, 1994). All of the measured arrivals for this transmission path corresponded to rays that either reflected from the surface or had upper turning depths within a few tens of meters of the surface. As a result the vertical resolution that might have been achievable using a receiver located closer to the depth of the sound-channel axis could not be determined. Interpretation of the absolute travel times in this experiment was difficult because the locations of the source and receiver were not precisely known and because no measurements were made of the sound-speed field between the source and receiver at the time of the experiment.

Finally, the Heard Island Feasibility Test (HIFT) in January 1991 demonstrated that low-frequency coded transmissions with a center frequency of 57 Hz could be coherently processed to yield adequate signal-to-noise ratios (SNRs) at ranges of up to 18 Mm (Munk *et al.*, 1994). None of the HIFT receptions yielded arrivals that were resolvable and identifiable with specific rays or modes, even at ranges as short as 5 Mm. Given that the previous experiments discussed above had yielded either ray clusters or individual ray arrivals that were identifiable at 3-4-Mm ranges, the inability to identify specific arrivals at about 5-Mm range in the HIFT case was probably due in part to the fact that all the signals traversed the complex oceanographic structure associated with the Antarctic Circumpolar Front. The difficulties encountered in interpreting the observed receptions were exacerbated by the fact that the sources were suspended from a moving ship, so that the geometry was continually changing and the long-term stability of the transmissions could not be determined.

The ATOC Acoustic Engineering Test differed from these earlier experiments in a number of important ways. Both the 75-Hz source and the receiver were located near the depth of the sound-channel axis at deep-ocean locations to avoid bottom interactions. The receiver was a 700-m-long

vertical line array so that the vertical structure of the received signals could be measured. One implication is that the vertical arrival angles of any raylike arrivals could be determined and used to confirm the ray identifications. Further, the array was designed to be able to spatially resolve the low order acoustic modes, so that modal travel times could be measured and modal scattering studied. The locations of the source and receiver were precisely determined using GPS navigation, so that the absolute range is known to within a few meters. Finally, the temperature and salinity fields along the transmission path connecting the source and receiver were measured during the experiment with closely spaced expendable bathythermograph (XBT) and expendable conductivity-temperature-depth (XCTD) casts, so that the upper-ocean sound-speed field can be accurately computed.

We are unaware of any previous broadband measurements that have been made at ranges in excess of 1000 km with a vertical receiving array. A long vertical array was deployed off Monterey, California, during HIFT at a range of 17 Mm to measure the modal content of the received signals, but low SNRs precluded broadband processing (Baggeroer *et al.*, 1994). Narrowband analyses showed a surprisingly rich modal population. The experiment that most closely resembles the one described here occurred during July 1989, when broadband signals were transmitted from a moored 250-Hz source to a 3-km-long vertical receiving array 1000 km distant in the north central Pacific Ocean (Howe *et al.*, 1991; Duda *et al.*, 1992; Cornuelle *et al.*, 1993; Worcester *et al.*, 1994). This experiment showed that energy confined near the sound-channel axis is significantly scattered by small-scale oceanic variability, reducing the vertical resolution that can be obtained using tomographic methods (Colosi *et al.*, 1994).

The ATOC Acoustic Engineering Test is described in Sec. I of this paper. The measured broadband acoustic arrival pattern at 3250-km range and 75-Hz center frequency is then compared in Sec. II with the arrival pattern predicted using the nearly concurrent temperature and salinity measurements. The goal is to determine if the two are quantitatively consistent as a function of travel time and depth to within the uncertainty imposed by our necessarily incomplete knowledge of the sound-speed field. A full and complete understanding of the forward problem is the *sine qua non* of any inverse problem. The observed ray travel times are found to differ from ray predictions by more than *a priori* variability estimates, suggesting that the equation used to compute sound speed from temperature, salinity, and pressure requires refinement. One of the key issues is the extent to which near-axial acoustic energy is scattered by internal-wave-induced sound-speed fluctuations. It is conceivable that acoustic scattering from small-scale oceanic variability will set a maximum range beyond which it is impossible to extract meaningful information on the temperature of the intervening ocean from acoustic travel-time data. Linear inverse methods are used in Sec. III to estimate the accuracy and vertical resolution with which the sound-speed field (and therefore temperature and heat content) in the vertical plane containing the source and receiver can be determined at mul-

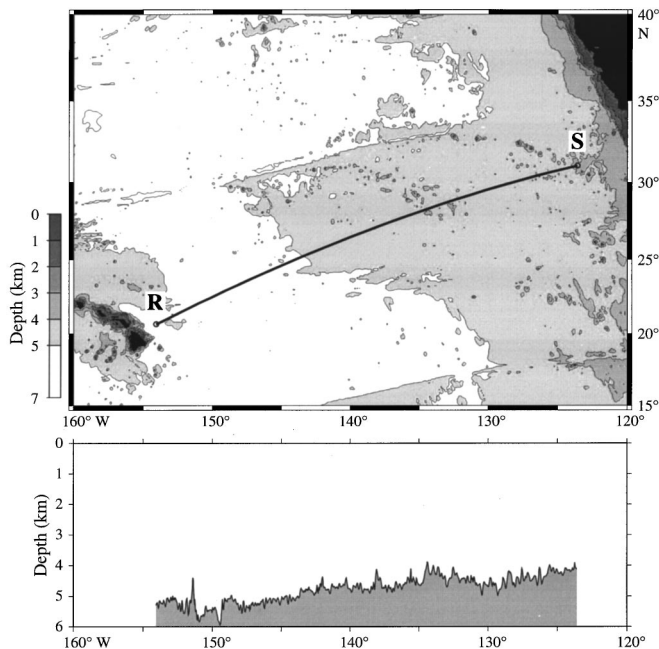


FIG. 1. Acoustic path from the 75-Hz acoustic source **S** suspended from R/P FLIP to the vertical receiving array **R** moored east of Hawaii. The path is superimposed on ETOPO-5 bathymetry (National Geophysical Data Center, 1988). Neither the ETOPO-5 bathymetry nor an improved estimate of the bathymetry using satellite altimetry data (bottom) shows significant features along the acoustic path. The improved bathymetry was derived by combining conventional sounding data with depths estimated from the gravity anomaly field computed from satellite altimetry data (Smith and Sandwell, 1994; Sandwell and Smith, 1997; Smith and Sandwell, 1997).

timegameter ranges. Finally, in Sec. IV we discuss the implications of these results.

## I. THE EXPERIMENT

The acoustic source was suspended at 652-m depth, near the sound-channel axis, from the floating instrument platform R/P FLIP, which was moored in the eastern North Pacific Ocean at  $31^{\circ} 2.050' N$ ,  $123^{\circ} 35.420' W$ , in water more than 4000 m deep (Fig. 1). R/P FLIP was in a trimoor to minimize source motion. The transmissions were received on two autonomous vertical line array (AVLA) receivers, as well as on numerous U.S. Navy receivers in the North Pacific Ocean and on a deep sonobuoy receiver near New Zealand. The results presented here were obtained from the AVLA moored at  $20^{\circ} 39.040' N$ ,  $154^{\circ} 04.640' W$ , just east of Hawaii, in water about 5312 m deep. The source and receiver positions were determined to within a few meters using differential GPS navigation. The range was 3 252 382 m (WGS-84). The refracted geodesic, which includes the effects of horizontal gradients in sound speed (see Munk *et al.*, 1995), is separated from the unrefracted geodesic by less than 2 km along the entire path and is longer than the unrefracted geodesic by only about 3 m. The transmission path does not cross any major oceanographic fronts or other features. There are no significant bathymetric features along the path (Fig. 1).

The instrumentation and signal processing are described in detail by the ATOC Instrumentation Group (1995). Only a brief summary is given here.

The source transmitted a phase-modulated signal with a center frequency of 75 Hz. The phase modulation was encoded using a linear maximal-length shift-register sequence containing 1023 digits. The source transducer has less bandwidth than desired but more drive capability than needed. The input waveform used to drive the source was therefore adjusted to give an output signal containing phase-modulated digits with two cycles of carrier each, even though the inherent bandwidth of the source was significantly narrower than  $Q=2$  (37.5 Hz). Each digit was then 26.667 ms in duration, and each sequence period was 27.2800 s long. Fifty-four transmissions were made, consisting of a mix of 10-, 20-, and 40-min transmissions, with 2 to 4 h between transmissions. The source level was 260 W (195 dB *re* 1  $\mu$ Pa at 1 m).

The vertical receiving array consisted of 20 hydrophones at 35-m spacing (1.75 wavelengths at 75 Hz) between 900- and 1600-m depth. (A second 20-hydrophone subarray, between 200- and 900-m depth, failed.) The incoming hydrophone signals were amplified, bandpass filtered, and sampled using 16-bit analog-to-digital converters at a 300-Hz rate.

Transmissions were made for seven days during November 1994, beginning on yearday 321 (17 November 1994) and ending on yearday 327 (23 November 1994). After return to shore, the acoustic data were complex demodulated and the phase modulation was removed. Forty periods (1091.20 s) of the received signal were recorded for the 20-min transmissions, but only 28 periods (763.840 s) were coherently processed to form the receptions shown here. The SNR of the ray arrivals did not increase for longer averaging times, indicating that the signal duration exceeded the signal coherence time. Doppler processing was required owing to the relative velocities of the moored source and receiver. Even though the relative velocities were small (1-2 cm/s), the long coherent processing time made the relative motion during transmissions significant. A standard Doppler search was performed for each reception to determine the relative velocity (assumed constant) that yielded the maximum output signal level.

The position of the source and the position and shape of the AVLA were measured by long-baseline acoustic navigation systems, using acoustic transponders on the seafloor. The source's position was determined to within about 1 m rms, using acoustic interrogators on the source package and on R/P FLIP. Source displacements as large as 1500 m occurred, with typical velocities of 1-2 cm/s (Fig. 2). The absolute positions of the elements in the vertical receiving array were determined to within about 1.5-m rms, using an acoustic interrogator at the center of the array. The navigation signals were received on six of the same hydrophones used to receive the 75-Hz signals, distributed over the array aperture, as well as on the interrogation transducer. The relative positions of the hydrophones were determined to within about 0.2 m rms. Array displacements in excess of 100 m were observed, as well as significant array curvature (Figs. 2 and 3).

Surprisingly, the relative velocities determined from the measured source and receiver locations did not always agree with those determined from the Doppler processing. The discrepancy is due to acoustic travel-time (phase) changes pro-

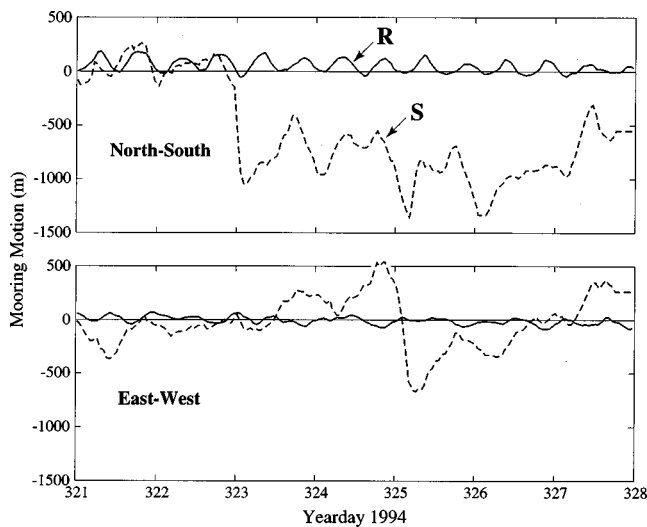


FIG. 2. North-south and east-west displacements of the source **S** (dashed) and of the center of the vertical receiving array **R** (solid) versus time.

duced during a transmission by the change in the barotropic tidal current along the acoustic path over the 20-min transmission interval. (Constant velocities, even if range dependent, do not produce a Doppler shift.) Tidally induced phase changes can be comparable to those caused by relative motions of the source and receiver. Even if both the source and receiver were truly fixed, a tidal correction would be required, either by Doppler processing or by a tidal prediction which allows for the finite duration of the transmission.

The temperature and salinity fields along the acoustic path were directly measured during the experiment. XBT (T-7) casts to 760-m depth were made at approximately 30-km intervals. XCTD casts to 1000-m depth were made at approximately 300-km intervals, in conjunction with every tenth XBT cast. The XBT/XCTD survey required seven days, from yearday 325 (21 November 1994) through yearday 331 (27 November 1994), and thus overlapped with the end of the transmission period. The survey began at the receiver and proceeded toward the source.

## II. THE FORWARD PROBLEM: PREDICTABILITY OF ACOUSTIC PROPAGATION

The combination of a vertical receiving array, nearly concurrent environmental measurements, and broadband signals designed to measure acoustic travel times with a precision of a few milliseconds makes it possible to test quantitatively whether the measured arrival pattern is consistent with

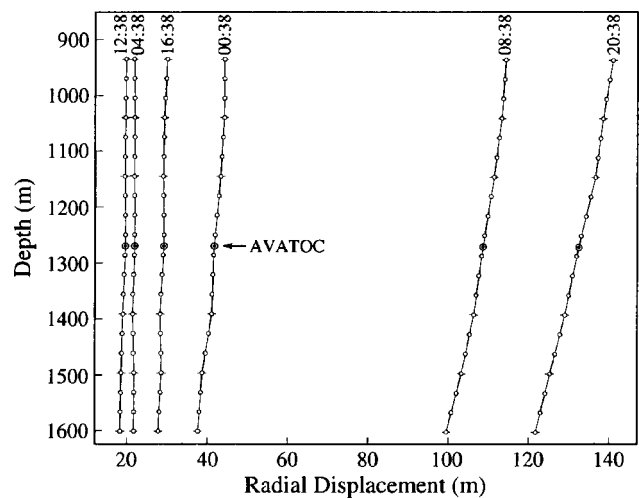


FIG. 3. Radial displacement of the Hawaii AVLA on yearday 326, 1994. All times are UTC.

predictions to within measurement uncertainty. Plotting acoustic intensity as a function of travel time and hydrophone depth clearly reveals acoustic wave fronts sweeping across the array (Fig. 4). Wave fronts can be seen prior to a travel time of about 2195 s, after which the arrival pattern becomes quite complex.

### A. Sound-speed field

The depth of the sound-channel axis gradually deepens as one proceeds from the source toward the receiving array (Fig. 5). Sound speeds were derived from climatological temperature (Levitus and Boyer, 1994) and salinity (Levitus *et al.*, 1994) data for November extracted from the World Ocean Atlas 1994. This climatology will be referred to hereafter as the WOA94 climatology, with the understanding that the climatology for November is used, rather than the annual average. Sound speeds were computed using the sound-speed equation of Del Grosso (1974). [Millero and Li (1994) provide a corrected version of the sound-speed equation of Chen and Millero (1977) that is equivalent for the parameter range of interest here and could equally well have been used.] The historical sound-channel axis is at about 650-m depth at the source and reaches about 800-m depth at the receiver, although the minima are not sharply defined. The source is therefore approximately on axis, and the top of the receiving array is slightly below axial depth.

Sound speeds in the upper 760 m at the time of the experiment were derived from the XBT and XCTD data. The

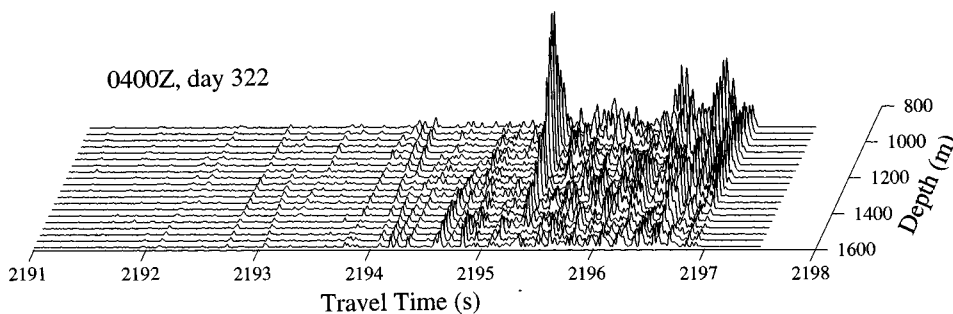


FIG. 4. Acoustic intensity as a function of travel time and hydrophone depth for the transmission at 0400 UTC on yearday 322.

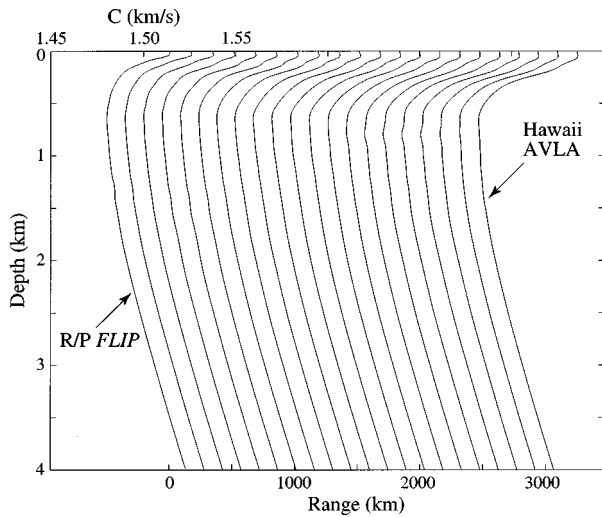


FIG. 5. Sound-speed profiles as a function of range between the source and receiving array, derived from the WOA94 climatology for November. Profiles are plotted every 162 km, corresponding to a sound-speed offset of 10 m/s.

salinity values corresponding to the XBT temperatures were estimated using the temperature-salinity relation from the nearest XCTD cast. Sound speeds below 760-m depth were calculated from the WOA94 climatology, as in Fig. 5. The upper ocean perturbations derived from the XBT/XCTD data were smoothed into the WOA94 sound-speed field below 760-m depth using a 200-m vertical  $e$ -folding scale.

A smooth, range-dependent sound-speed field was constructed from a combination of the XBT/XCTD measurements and the historical data by objective mapping (Fig. 6). The objective mapping procedure provides an estimate of the uncertainties in the smooth field due to limited coverage and measurement noise. These uncertainties can be converted into uncertainties for the travel-time predictions made from the objectively mapped field by using the linearized forward problem (Dushaw *et al.*, 1993; Cornuelle *et al.*, 1993). The objective mapping procedure also largely removes internal-wave variability from the sound-speed maps by smoothing in the horizontal and the vertical with scales appropriate to ocean mesoscale variability.

To construct the maps, the sound-speed field was expressed as a range-dependent perturbation field relative to a range-dependent reference field obtained from the WOA94 climatology. The sound-speed perturbation was assumed to

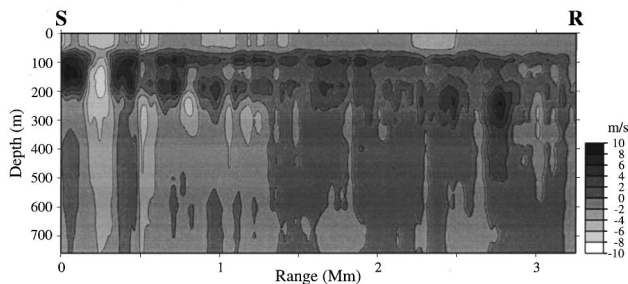


FIG. 6. Sound-speed perturbation in the upper 760 m computed by subtracting sound speeds derived from the WOA94 climatology from objectively mapped sound speeds derived from the XBT/XCTD data. The contour interval is 2 m/s. The source **S** is on the left and the receiver **R** is on the right.

be a random variable with zero mean and a known covariance, separable into an inhomogeneous vertical covariance and a homogeneous, depth-independent horizontal covariance with 70-km  $e$ -folding scale.

The sound-speed perturbation field  $\Delta C(x, z)$  was discretized as a sum of products of horizontal and vertical functions,

$$\Delta C(x, z) = \sum_{j=1}^M \sum_{k=1}^K a_{jk} F_j(z) G_k(x), \quad (1)$$

where  $a_{jk}$  are the model parameters. The  $F_j(z)$  were chosen to be the eigenvectors of the vertical covariance of the sound-speed perturbations as a function of depth as estimated from the XBT data,

$$\langle \Delta C(z_1) \Delta C(z_2) \rangle = \sum_{j=1}^M F_j(z_1) \lambda_j^2 F_j(z_2), \quad (2)$$

where the  $\lambda_j$  are the eigenvalues, i.e., the  $F_j(z)$  are the empirical orthogonal functions (EOFs) for the XBT data set. The first six EOFs were used, which account for 93% of the observed XBT variance. Retaining additional EOFs only gradually reduces the residual variance while increasing the number of model parameters and the small-scale structure in the mapped field.

Truncated Fourier series were used for the horizontal parametrization. The wave number spectrum was specified to be white for wavelengths longer than 300 km, and proportional to wavelength (inverse wave number) squared for shorter wavelengths. This spectrum gives a horizontal covariance  $e$ -folding scale of 70 km but includes allowances for a strong mean component.

Sound speeds computed from the XBT/XCTD measurements,  $C(x_i, z_i)$ , are related to the model parameters by

$$\begin{aligned} \Delta C_i(x_i, z_i) &\equiv C(x_i, z_i) - C_0(x_i, z_i) \\ &= \sum_{j=1}^M \sum_{k=1}^K a_{jk} F_j(z_i) G_k(x_i) + r_i, \end{aligned} \quad (3)$$

where the  $\Delta C_i$  are the sound-speed perturbations at the measurement locations relative to a reference sound-speed field  $C_0(x_i, z_i)$ , selected here to be the WOA94 climatology. The measurement noise  $r_i$  is due to instrument error (which varies with depth and is typically correlated because of fall-rate error) and unmodeled phenomena (model error), including internal waves. The coefficients  $a_{jk}$  are computed by objective mapping, i.e., by a tapered, weighted least-squares fit that minimizes the differences between the  $\Delta C_i$  computed from the XBT data and those computed from the ocean model.

The sound-speed perturbation field shows significant eddy activity near the source, with peak perturbations of up to about 10 m/s (Fig. 6). The eddy activity is weaker over most of the remaining path. A more-or-less uniform surface mixed layer with a depth of about 100 m is present over almost the entire path.

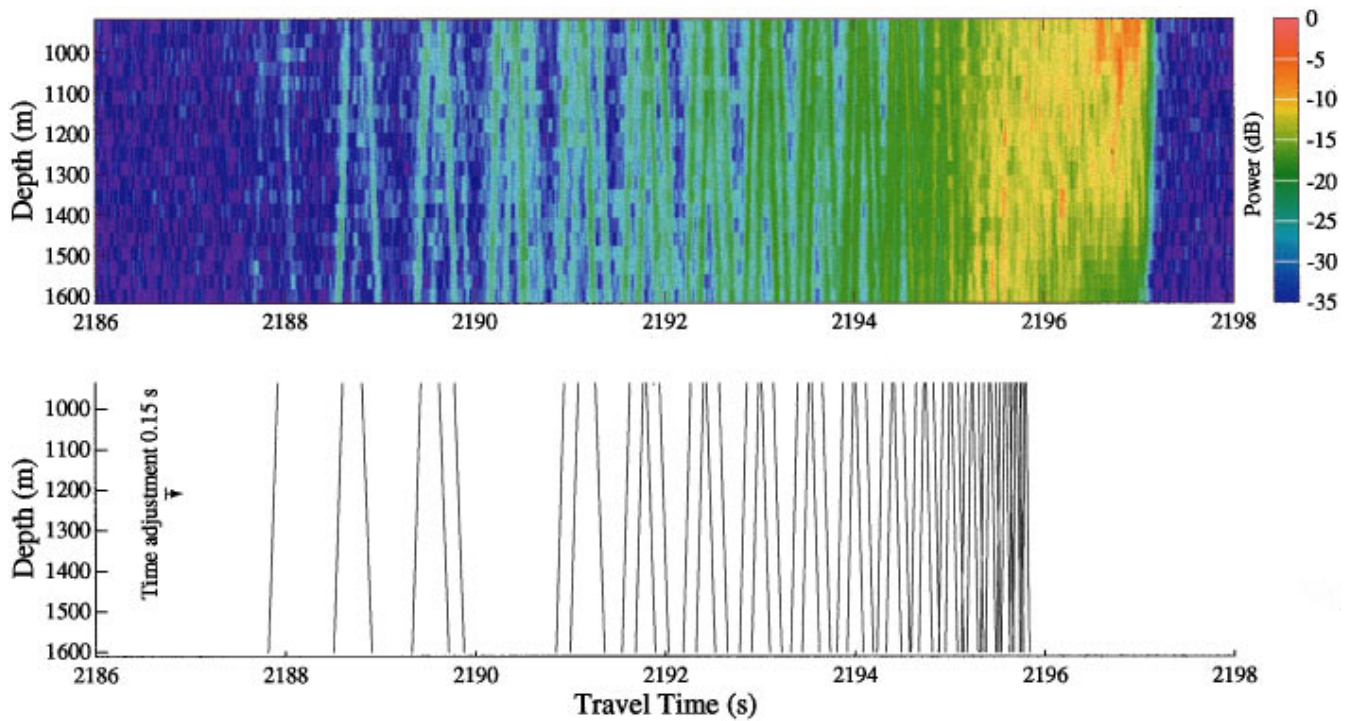


FIG. 7. (Top) Measured daily average acoustic intensity as a function of travel time and hydrophone depth. The measured pattern is an incoherent average of all of the receptions on yearday 326, 1994. Acoustic intensity relative to the peak intensity in the daily average is plotted in decibels. (Bottom) Predicted travel times versus hydrophone depth computed using a range-dependent ray tracing algorithm and the sound-speed field derived from the WOA94 climatology. Only the wave fronts predicted by ray theory prior to about 2196 s are shown. The predicted travel times have been delayed by 0.15 s to align the measured and predicted arrival patterns.

## B. Geometric optics and ray identification

Raylike arrivals, in the form of wave fronts sweeping across the array, are clearly evident in the measured data prior to a travel time of about 2195 s, as noted above. These arrivals are both resolvable and stable over the duration of the experiment. It will be shown in this section (*i*) that these arrivals can be unambiguously identified with specific acoustic rays using any of several ocean climatologies, without the need to use the XBT/XCTD data, and (*ii*) that the measured travel times are offset from the travel times predicted using the sound-speed field constructed from the XBT/XCTD measurements by a surprisingly large amount.

Using the sound-speed field constructed from the WOA94 climatology, range-dependent ray tracing over the measured range between the source and receiver gives predicted wave fronts that match those observed (Fig. 7). The range-dependent ray code used is that of Colosi; the predicted travel times have been checked against those from RAY (Bowlin *et al.*, 1993). The ray tracing was done in Cartesian coordinates, after appropriately transforming the depth coordinate and sound-speed profiles from spherical coordinates (Munk *et al.*, 1995). The radius of curvature to be used in the transformation is not obvious, as the true radius of curvature is a function of position on a spheroidal earth. Changing the assumed radius of curvature by up to 50 km changes the computed travel times for the geometry of this experiment by less than 3 ms, however, so the precise choice of the radius of curvature is not critical.

The measured and predicted travel times do not agree

exactly, of course, at least in part because the ocean sound-speed field at the time of the acoustic transmissions is not the same as that computed from the WOA94 climatology. The predicted travel times shown in Fig. 7 have been delayed by 0.15 s to approximately align the measured and predicted arrival patterns. The similarity of the measured and predicted arrival patterns makes the identification of the measured wave fronts with particular ray paths straightforward. At least in this case climatological data are adequate for ray identification. The essential information to be used in the inversions in Sec. III is the slight mismatch between the measured and predicted travel times. The significance of the overall shift between the two patterns will be discussed at length in that section.

The identification can be confirmed by comparing the measured and predicted vertical arrival angles of the rays (Fig. 8). The measured arrival angles were estimated from the data with a conventional plane-wave beamformer, using the entire 700-m array aperture. The predicted arrival angles were obtained from the same range-dependent ray trace used to construct Fig. 7. After shifting the predicted travel times 0.15 s later, as in Fig. 7, the agreement between the measured and predicted ray arrival times and angles is excellent. Each wave front can therefore be labeled with an identifier  $\pm n$ , where + (−) indicates a ray that initially travels upward (downward) at the source and has a total of  $n$  upper and lower turning points between the source and receiver. None of the identified rays interacts with the seafloor.

Given the large number of resolved wave fronts in the

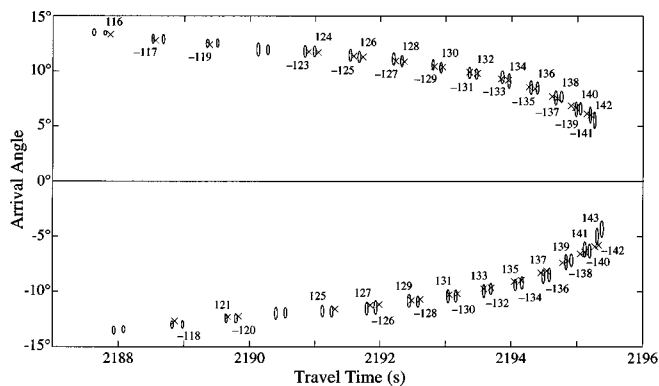


FIG. 8. Measured (ellipses) and predicted (crosses) vertical arrival angles versus travel time for the raylike arrivals. The ellipses are centered on the time-mean travel times of the ray arrivals for the entire experiment. The sizes of the ellipses show the rms variations in angle and time about the means. The predicted arrivals were computed using the WOA94 climatology and were delayed by 0.15 s to align the measured and predicted arrival patterns, as in Fig. 7. The arrivals are labeled with their ray identifiers.

measured arrival pattern, with similar vertical arrival angles, the ray identification might appear to be ambiguous. To test whether this is indeed the case, the rms difference between the measured and predicted travel-time patterns was computed as a function of the time shift applied to the predicted travel-time pattern (Fig. 9). To make this calculation, time series of the measured travel times for each ray arrival were constructed from the beamformer output (Fig. 10). The time-means of these series were then used to compute the rms travel-time difference. The rms difference has a clear global minimum for a time shift of 0.15 s, indicating that the measured and predicted time-mean patterns are most closely aligned when the predicted pattern is shifted later by this amount. (This is the shift used to align the measured and predicted arrival patterns in Figs. 7 and 8.) Sound speeds from the WOA94 climatology were used to construct Fig. 9. Similar results (Table I) were obtained using sound speeds computed from the earlier 1982 Levitus climatology (Levitus, 1982) and a climatology recently developed by Men-

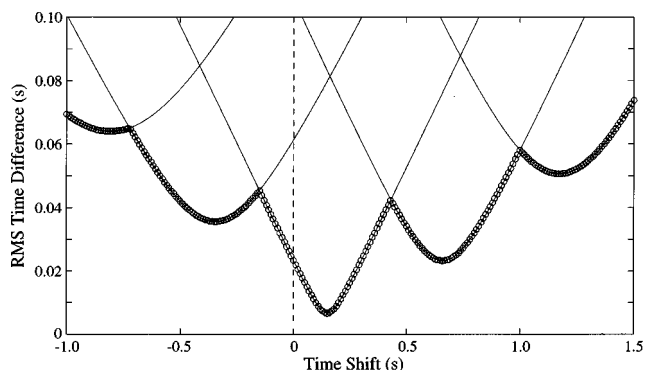


FIG. 9. The rms difference between the time means of the measured travel times for each ray arrival, as determined from the beamformer output, and the predicted travel times, versus the time shift applied to the predicted travel times to align the two patterns. The rms travel-time difference is normalized by the number of identified rays. The predicted arrivals were computed using the WOA94 climatology, as in Figs. 7 and 8. A positive time shift means that the predicted travel times are less than the measured travel times.

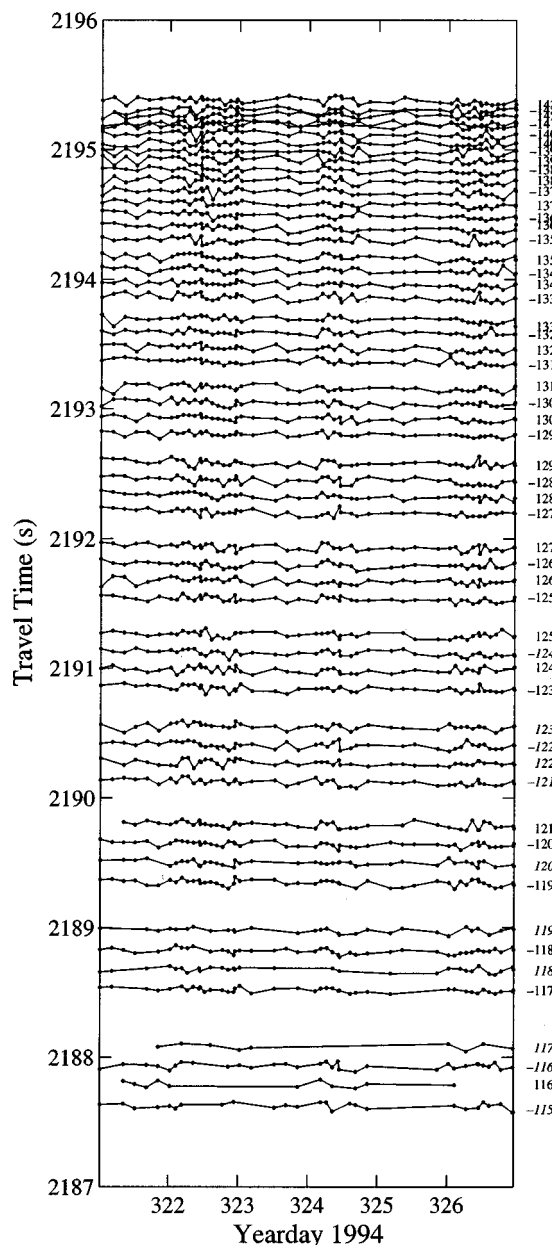


FIG. 10. Acoustic travel times for the raylike arrivals, as determined from the plane-wave beamformer output, versus yearday. The beamformer output is for an effective depth of 1268 m, the midpoint of the nominal array location. The travel times have been corrected for motion of the source and receiving array, as well as for clock drifts in the receiver. The source clock was synchronized to UTC using a GPS receiver. Ray identifiers are given on the right. Ray identifiers in italics indicate rays for which the ray trace code used did not succeed in locating the eigenray due to computational difficulties, although the identifications are nonetheless unambiguous.

emenlis *et al.* (1997). In all cases, the minimum is unique, so the ray identification is unambiguous.

The structure of the rms time difference as a function of time shift can be understood by noting that both the measured and predicted arrivals come in groups of four, as is evident in Fig. 7 (see, e.g., Munk *et al.*, 1995). Any ambiguity will therefore be in associating measured and predicted groups of four arrivals. The secondary minima in Fig. 9 occur when groups of four in the predicted pattern are associated with incorrect groups of four in the measured pattern.

Given an unambiguous ray identification, the discrep-

TABLE I. Comparison of the ray travel times predicted for various sound-speed fields with the time means of the measured travel times for each ray arrival, as determined from the beamformer output. The time shift that gives the best alignment between the predicted and measured travel times, the rms difference between the two for this time shift (normalized by the number of identified rays), and the number of identified rays are given. A positive time shift means that the measured travel times are greater than the predicted travel times.

	Time shift, (s)	rms, (s)	No. of identified rays
Levitus (1982)	0.01	0.0067	48
WOA94	0.15	0.0063	47
Menemenlis	0.01	0.0114	52
WOA94+XBT	0.22	0.0049	49

ancy between the measured and predicted ray arrival times can be quantified. Figure 11 shows the travel-time perturbation  $\Delta\tau$  computed by subtracting the predicted travel times from the time-mean measured travel times, as derived from the plane-wave beamformer output. Predictions were made using both the sound-speed field derived from the WOA94 climatology and the objectively mapped sound-speed field derived from the XBT/XCTD data. The travel-time perturbations are plotted as a function of the minimum upper and maximum lower turning point depths. (The upper and lower turning point depths are functions of position for range-dependent propagation.) This choice of ordinate aids in interpreting the data, because the rays are most sensitive to the ocean at the turning point depths. Figure 11 shows that the measured travel times are greater than the predicted travel times computed using the WOA94 climatology by about 0.15 s on average, which is why delaying the predicted travel times by this amount in Figs. 7 and 8 approximately aligns the two patterns. Positive travel-time perturbations for all

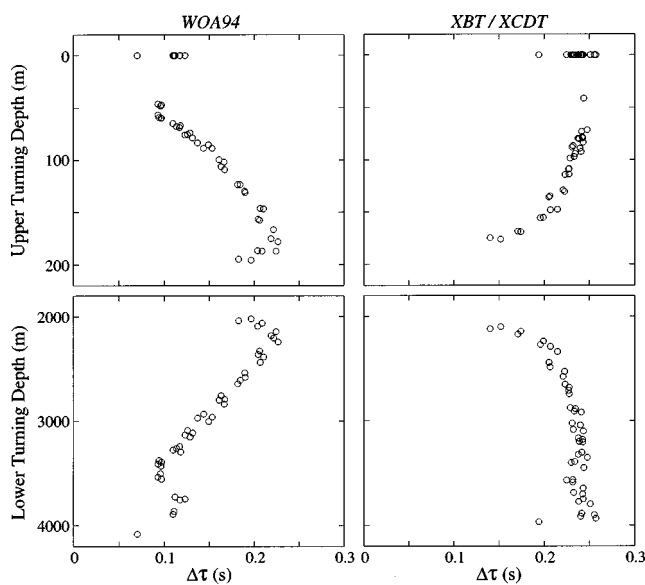


FIG. 11. Travel-time perturbations computed by subtracting the predicted travel times from the time-mean measured travel times, as derived from the beamformer output, plotted as a function of minimum upper turning depth (top) and maximum lower turning depth (bottom). The predictions were made using the sound-speed field derived from the WOA94 climatology (left) and the objectively mapped sound-speed field derived from the XBT/XCTD data combined with the WOA94 climatology (right).

rays that turn in a particular depth range suggest that sound speeds in that depth range are lower than those used to make the predictions. Because each ray turns above and below the sound-channel axis at depths with equal sound speeds (conjugate depths), the location of the sound-speed perturbation responsible for the travel-time perturbation is fundamentally ambiguous (Munk and Wunsch, 1982).

Using the time-mean travel times for the comparison assumes that, aside from internal waves and tides, changes in the sound-speed field over the duration of the experiment are small. The six-day period during which acoustic data were collected is short compared to the several-week periods typical of mesoscale variability, so this assumption is plausible. The acoustic travel times show a small, but nonzero, trend over the duration of the experiment (Fig. 10), supporting the assumption of a largely unchanging subinertial sound-speed field, at least for the purpose of identifying the ray arrivals.

Explicit error bars can be computed for the travel-time offsets relative to predictions made using the XBT/XCTD data. The error bars are a combination of the uncertainties in the measured and predicted travel times.

The computed uncertainty in the predicted travel times arises from uncertainties in the objectively mapped sound-speed field used to make the predictions (owing to the imperfect sampling of the field), to the uncertainty in the equation used to compute sound speed from temperature and salinity ( $\pm 0.05$  m/s rms), and to the uncertainty in source-receiver range. The precision with which the source (1 m rms) and receiver (1.5 m rms) locations were determined using differential GPS navigation leaves an uncertainty of only about 2 m rms in the range, which is negligible compared to the other sources of uncertainty. The uncertainties in the predicted travel times were calculated using the output error covariance provided by the mapping procedure and the linearized forward problem matrix, which converts ocean parameters to travel-time perturbations. Because the objective map covered only the upper ocean, uncertainties in the deep sound-speed field were estimated from historical data.

The uncertainties in the time-mean measured travel times are a combination of the error in determining the peak arrival time due to ambient noise, internal-wave-induced travel-time fluctuations, tidally induced travel-time fluctuations, and the small travel-time trend observed over the course of the acoustic measurements. Travel-time fluctuations at tidal frequencies were fit and removed prior to forming the mean. After energy at tidal frequencies is removed, the remaining high-frequency travel-time fluctuations about a linear trend are found to be 11-19 ms rms, due primarily to internal waves. [See Colosi *et al.* (1999) for a comparison of the observed travel-time variances to predictions made assuming that the fluctuations are due to internal-wave-induced scattering.] These high-frequency travel-time fluctuations are suppressed in the mean. The trend will be discussed further in Sec. III. In this case the dominant sources of error are the travel-time trend and the uncertainty in the predicted travel times.

Surprisingly, the measured travel times are significantly offset from predictions made using the XBT/XCTD data (Fig. 11). The average time shift required to align the arrival

pattern predicted using the XBT/XCTD data and the measured arrival pattern is 0.22 s, which is significantly larger than the 0.15 s needed to align the predictions made using the WOA94 climatology with the measurements (Table I). The offset is about three times the expected rms uncertainty due to mapping errors and climatological uncertainties, although the expected rms variability of range-averaged temperature and salinity below 760-m depth is difficult to quantify. The rms difference between the measured and predicted arrival times for the optimum time shift is smaller when using the XBT/XCTD data, however, which means that the relative spacing within the arrival pattern is better predicted using the nearly concurrent environmental data (Table I). This improvement is reflected in Fig. 11, in that the travel-time offsets computed using the XBT/XCTD data are relatively independent of ray turning depth. We will return to this issue in Sec. III.

### C. Surface interactions

The discussion to this point has focused on a comparison of the ray arrivals observed prior to a travel time of about 2195 s to predictions made using geometric optics. In this section, predictions made using geometric optics are compared with predictions made using a broadband adiabatic mode code to assess the adequacy of the geometric optics approximation at the long ranges and low frequencies of interest. The adiabatic mode approximation includes more complete physics than geometric optics and thus should more faithfully model reality, at least in cases for which the adiabatic approximation is valid. Boden *et al.* (1991) made similar comparisons and reported the ray approximation to be adequate for estimating travel times out to 4000-km range for frequencies above 100 Hz.

Predictions were made for a source depth of 700 m, a range of 3000 km, and a range-independent sound-speed profile typical of that found in the northeast Pacific Ocean (Colosi *et al.*, 1994). The adiabatic predictions were made for 120 modes, using the code of Dzieciuch (1993). Frequency domain predictions for a center frequency of 75 Hz and a  $-3$ -dB full bandwidth of 30 Hz were Fourier synthesized to form broadband arrival patterns in the time domain.

Comparison of calculated ray travel times with those derived from the adiabatic prediction shows that the two agree within a few milliseconds (at 3000-km range) for rays that reflect off the surface and for rays with upper turning depths well below the surface (Fig. 12). Rays with upper turning depths within a few wavelengths of the surface ( $\lambda = 20$  m at 75 Hz) arrive up to 40 ms earlier in the geometric optics approximation than in the adiabatic mode approximation, however. Rays refracting near the surface, but that do not actually reflect off it, are unaffected by its presence. Modes, however, have evanescent tails that are affected by the presence of the pressure-release boundary condition at the surface, even when the modal turning depth is below the surface, slightly reducing the group speeds.

This effect is somewhat larger than the precision with which travel times can be measured, and therefore constitutes a potentially important bias for ray travel-time inversions at long ranges. For the simple range-independent ge-

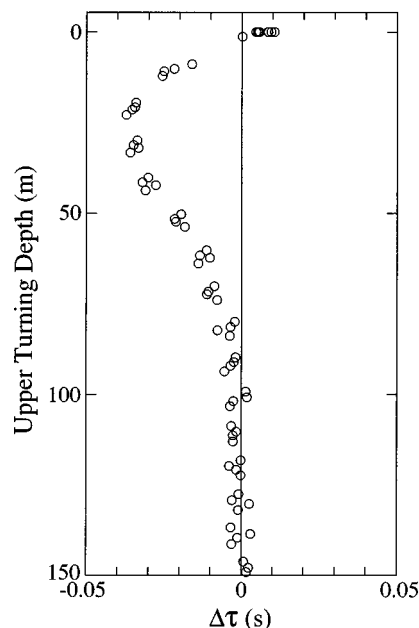


FIG. 12. Travel-time bias due to nongeometric effects for rays turning near the surface. Ray travel times minus adiabatic-mode travel times are plotted as a function of ray upper turning depth. To estimate the bias the predictions were made for a source depth of 700 m and a range of 3000 km, using a range-independent profile typical of the northeast Pacific Ocean.

ometry for which the test calculations were made, the rays turning within a few wavelengths of the surface have approximately 50 upper turning points, so the bias is roughly 1 ms per surface interaction.

The effect is expected to be relatively small for the measurements described here, however, compared to travel-time uncertainty. Surface interactions for the observed rays are confined to approximately the first 500 km of the path, because the depth of the sound-channel axis increases from the source toward the receiver, giving at most 12 surface interactions per ray, or about 12 ms bias. This small bias will be neglected when the travel times are used to estimate the sound-speed field in Sec. III. In addition, few rays have minimum upper turning depths in the upper 50 m, where the effect is most significant (Fig. 11).

This effect is not new (Murphy and Davis, 1974), but has not previously been accounted for in long-range acoustic thermometry. Boden *et al.* (1991) comment that although they found the mean differences between the ray and normal mode travel times to be small, this was not entirely due to the goodness of fit between the ray and normal-mode solutions. Their ray travel times were consistently greater than their normal-mode travel times for surface-reflected rays ( $\pm 13.5^\circ$  to  $\pm 16.5^\circ$ ), but the ray travel times were consistently less than the normal-mode travel times for near-surface ( $\pm 11.5^\circ$  to  $\pm 13.5^\circ$ ) refracted arrivals. The sign of the difference for the near-surface refracted arrivals is consistent with that found here and suggests that this effect may have been important in their simulations as well, although they did not discuss it.



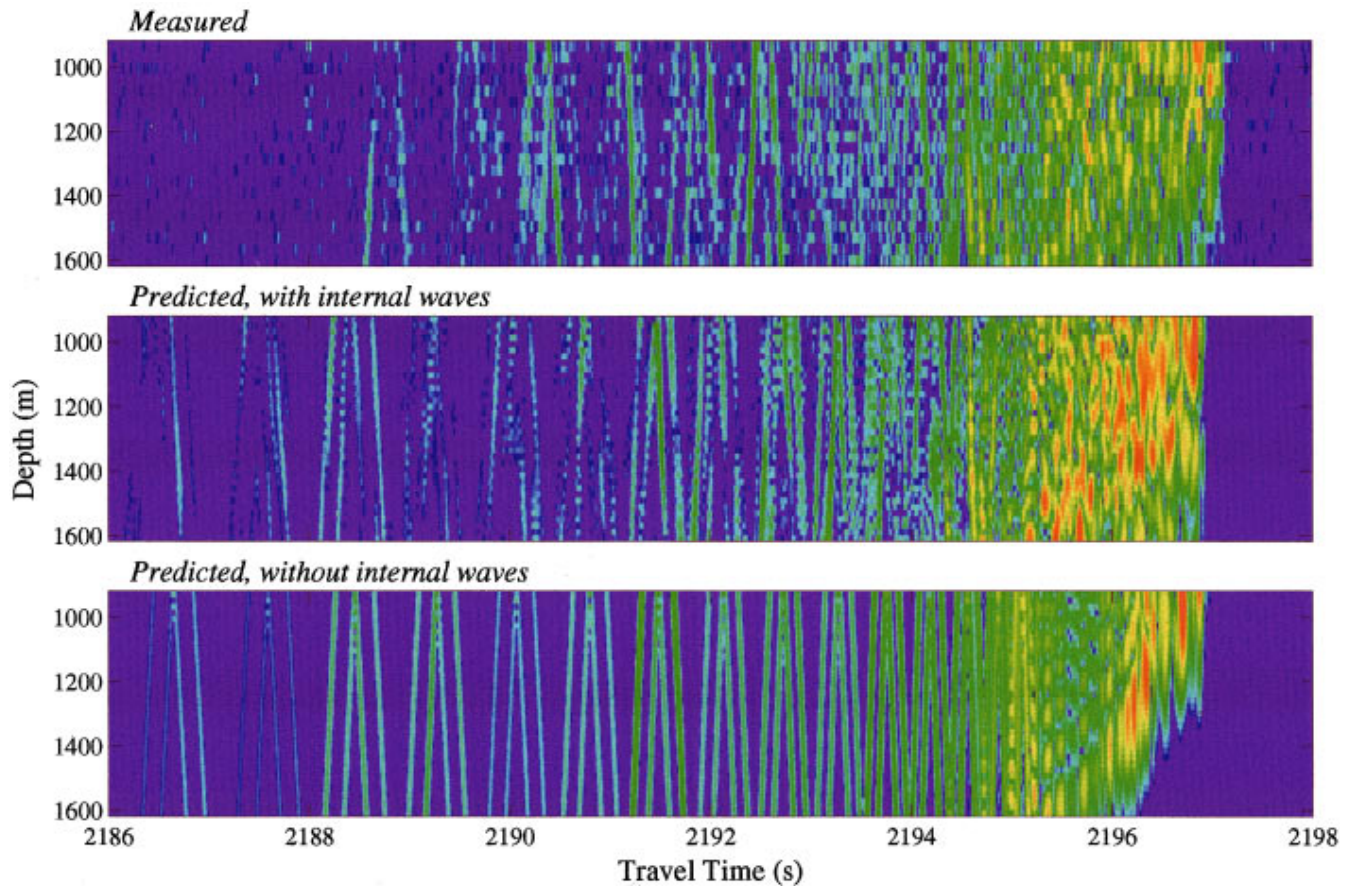


FIG. 13. Acoustic intensity as a function of travel time and hydrophone depth for one transmission (top) compared to broadband parabolic equation predictions with (middle) and without (bottom) simulated internal waves. Acoustic intensity relative to the peak intensity is plotted in decibels. The background range-dependent sound-speed field used in the simulations is shown in Fig. 6.

#### D. Internal-wave scattering and the pulse termination

Discreet ray arrivals are no longer evident in the measured arrival pattern for travel times greater than about 2195 s (Figs. 4 and 7). This behavior is similar to that previously observed at 1000-km range and 250-Hz center frequency (Worcester *et al.*, 1994). In that case Colosi *et al.* (1994) used broadband parabolic equation simulations to show that the observed scattering of the near-axial acoustic energy was consistent with what would be expected for propagation through an internal-wave field containing approximately one-half the energy in the standard Garrett–Munk internal-wave spectrum.

Similar broadband parabolic-equation simulations were made for the geometry of this experiment. Frequency domain predictions for a center frequency of 75 Hz and a  $-3$ -dB full bandwidth of 30 Hz were Fourier synthesized to form broadband arrival patterns in the time domain, as in the previous section. Predictions made for the sound-speed field constructed using the XBT/XCTD data (Fig. 6) give pulse terminations (i.e., the times at which the acoustic reception ends) much earlier than observed for hydrophones well below the sound-channel axis (Fig. 13). Adding a simulated internal-wave field with the standard Garrett–Munk internal-wave energy level gives a predicted arrival pattern that does not contain discreet ray arrivals after a travel time of about 2195 s and that has pulse terminations nearly independent of

depth over the 900- to 1600-m-deep array aperture, as observed. Modal decompositions of similar parabolic-equation simulations at multimegameter ranges show that the propagation of the near-axial acoustic energy is generally not adiabatic, but rather that there is strong modal coupling due to the presence of internal waves (Colosi and Flatté, 1996; Colosi, 1997).

The strong scattering of near-axial energy by internal waves makes it difficult to extract useful observables for the energy arriving after about 2195 s. Rays can no longer be separated from one another, even using the vertical array to provide resolution in both time and angle. Although modal arrivals can be separated using the vertical array, internal-wave-induced modal coupling means that energy contained in a given mode at the receiver has propagated in a variety of modes, at a variety of group velocities, en route from the source to the receiver, making the interpretation of the observed group delays fundamentally ambiguous. Further, the scattered modes show significantly more time spread at the receiver than would be expected considering only modal frequency dispersion across the band of interest in the absence of internal-wave scattering, making accurate determination of the group delays difficult.

The lowest-order modes are found in the numerical simulations to be more nearly adiabatic than higher-order modes (Colosi and Flatté, 1996), however, suggesting that

their group delays (travel times) might be useful in inversions to determine the sound-speed field, increasing the vertical resolution of the resulting estimates.

The data from the vertical array can be used to estimate the amplitude of mode 1 as a function of time using least-squares techniques, even though the array aperture is at best marginal for mode isolation using conventional spatial filtering techniques (Sutton *et al.*, 1994). Unfortunately, the time history of even the lowest-order mode amplitude estimated by least squares is quite complex. Using the peak of the mode 1 amplitude to define the modal group delay gives an rms variation about the mean of 178 ms, which is much greater than the rms of about 20 ms for the observed ray travel times. This large rms variation suggests either that mode coupling due to small-scale ocean structure is more important than the simulations might suggest, or that the energy from other modes is leaking into the solution for the mode 1 amplitude because of the suboptimum array geometry. The group delay of mode 1 is thus difficult to determine.

The pulse termination at the most axial hydrophone is relatively unambiguous, however, and might be viewed as a surrogate for the group delay of adiabatic mode 1. The rms variation about a linear trend fit to the pulse termination times is 22.9 ms, which is much less than the rms variation of mode 1 as determined above and is comparable to the rms variation of the observed ray travel times. [See Colosi *et al.* (1999) for a comparison of the observed travel-time variance to a prediction made assuming that the fluctuations are due to internal-wave-induced scattering.] Even in the presence of mode coupling, no energy can travel more slowly than the energy that propagates the entire distance between source and receiver in mode 1. To the extent that any acoustic energy remains in mode 1 over the entire path, the pulse termination will reflect the travel time of that energy.

The time-mean measured travel time of the pulse termination is 398 ms greater than predicted using the WOA94 climatology, but only 75 ms greater than predicted using the objectively mapped XBT/XCTD data. Surprisingly, the use of the nearly concurrent XBT/XCTD data improves the agreement between the measured and predicted group delay for mode 1 much more significantly than their use improved the agreement between measured and predicted ray travel times, as discussed in Sec. II B. This is presumably due to a combination of effects. The midpoint of the acoustic measurements occurred on yearday 324, while the midpoint of the XBT/XCTD survey occurred four days later on yearday 328. The upper ocean can change much more rapidly than the deeper ocean, so the XBT/XCTD data can be expected to represent the state of the near-axial ocean somewhat more accurately than the state of the upper ocean at the time of the acoustic measurements. Further, and probably more important, mode 1 is sensitive to the error in near-axial sound speed integrated over the entire range. This error is reduced using the XBT/XCTD data. In contrast, the rays are sensitive both to the upper-ocean sound-speed field sampled by the XBT/XCTD measurements and to the deep sound-speed field about which the XBT/XCTD data provide no information.

### III. THE INVERSE PROBLEM

In this section both the time mean and the trend in the sound-speed field between the source and receiver are estimated. The range-independent perturbation needed to bring the measured and predicted travel times into consistency is estimated first, using least squares inverse methods similar to those used to map the sound-speed field in Sec. II. The data consist of the time-mean travel times of the resolved rays and of the pulse termination at the most axial hydrophone, used as a surrogate for the group delay of mode 1. The linear trends in sound speed, temperature, and heat content over the six-day experiment are estimated next, using as data the linear trends in the travel times of the resolved rays and the pulse termination.

#### A. Inverting for the mean state

As was noted in Sec. II, the time-mean travel times are significantly offset from predictions based on the range-dependent sound-speed field constructed using the XBT/XCTD measurements. It is in principle impossible to construct a consistent inverse estimate for the sound-speed field between the source and receiver without increasing one or more of the *a priori* uncertainties assumed in Sec. II. The inverse estimates presented here assume a greater uncertainty in the sound-speed field below 1000-m depth than was previously estimated from historical temperature and salinity data. The notion is not that the range-averaged deep ocean temperature and salinity fields differ substantially from historically observed values, but that the equation used to compute sound speed from temperature and salinity may require further refinement at high pressure. Unfortunately, no direct measurements of the deep temperature and salinity fields were made along the acoustic path at the time of the acoustic transmissions. The estimated sound-speed perturbation below 1000-m depth therefore represents a combination of a correction to the sound-speed equation, real temperature and, to a much lesser extent, salinity perturbations, and the possible effects of travel-time biases.

To first order, the ray travel-time perturbations  $\Delta\tau_i$  are weighted averages of the sound-speed perturbations  $\Delta C(z)$  integrated along the unperturbed ray paths  $\Gamma_i$ :

$$\Delta\tau_i = - \int_{\Gamma_i} \frac{ds}{C_0^2(x,z)} \Delta C(z), \quad (4)$$

where ocean currents have been neglected and  $C_0(x,z)$  is the selected reference sound-speed field. The corresponding sensitivity of the group delay of adiabatic mode 1 to the sound-speed perturbation was derived numerically.

#### 1. Reference state

The range-dependent sound-speed field  $C_0(x,z)$  constructed in Sec. II A by combining the XBT/XCTD data with the WOA94 climatology was used as the reference state (Fig. 6).

TABLE II. Vertex locations for the vertical parametrization used in the inversions.

Depth range (m)	Vertex spacing (m)
0–200	20
200–600	40
600–1400	80
1400–4000	200
4000–6000	500

## 2. Sound-speed perturbation model

The sound-speed perturbation field  $\Delta C(z)$  was discretized (“modeled”) as a sum of range-independent vertical functions (“modes”):

$$\Delta C(z) = \sum_{j=1}^M a_j F_j^C(z), \quad (5)$$

where  $a_j$  are the model parameters, which were assumed to be uncorrelated. In anticipation of the next section, the vertical sound-speed modes were derived by first representing the temperature perturbation field as a sum of range-independent vertical functions

$$\Delta T(z) = \sum_{j=1}^M a_j F_j^T(z). \quad (6)$$

The XBT EOFs used previously are not expected to represent arbitrary range-averaged offsets, and thus are not a good choice for upper ocean functions in this application. The historical data base is too limited a sample to provide reliable estimates of the vertical covariance function expected for range-averaged temperature perturbations. Instead, linear splines (piecewise linear segments) were used, with the knot spacing adjusted in accord with rough estimates of the vertical length scales,

$$\begin{aligned} F_j^T(z) &= 0, & z < z_{j-1}, \\ F_j^T(z) &= \frac{(z - z_{j-1})}{(z_j - z_{j-1})}, & z_{j-1} \leq z \leq z_j, \\ F_j^T(z) &= \frac{(z_{j+1} - z)}{(z_{j+1} - z_j)}, & z_j < z \leq z_{j+1}, \\ F_j^T(z) &= 0, & z_{j+1} < z. \end{aligned} \quad (7)$$

Forty-eight vertical functions were used, with vertex spacings that increase with increasing depth, reflecting our prejudice that vertical scales become larger at greater depths (Table II). A range-independent model was used in this case in part because little range-dependent information is expected from the travel times and in part for simplicity.

Sound-speed and heat-content modes were constructed for each temperature mode by treating the temperature modes as perturbations to a reference temperature profile constructed by averaging the WOA94 climatology along the acoustic path. Each temperature perturbation (with a maximum amplitude of  $0.1^\circ\text{C}$  for linearity) was added to the reference temperature profile, and sound speed and heat content were calculated for the perturbed profile using the range-averaged salinity from the WOA94 climatology. Heat con-

tent was integrated upward from 4200-m depth, which was chosen because this depth is below the deepest ray turning point,

$$H(z) = \rho C_p \int_{-D}^z T(z') dz', \quad (8)$$

where  $H$  has units of energy per unit area,  $\rho = 1024 \text{ kg/m}^3$  is the density of seawater,  $C_p = 3996 \text{ J/(kg }^\circ\text{C)}$  is the specific heat of seawater,  $T(z)$  is the range-averaged temperature profile, and  $D = 4200 \text{ m}$ . The reference sound speed or heat content was then subtracted from the perturbed profile, as appropriate, and the resultant perturbation used as the sound speed or heat content analog of the vertical temperature mode.

Only the sound-speed perturbation was computed for the mean state, so this rather elaborate procedure was not necessary. But sound-speed, temperature, and heat content perturbations will be computed in the next section for the trend. This procedure uses a single set of model parameters to provide consistent estimates for all three perturbations, assuming no salinity perturbations.

Finally, the expected range-average temperature variability as a function of depth, relative to the reference state constructed using the XBT/XCTD data, is estimated to be  $0.12^\circ\text{C}$  rms at the surface, decreasing nearly linearly to  $0.05^\circ\text{C}$  rms at 1000-m depth. The variability was assumed to remain constant at  $0.05^\circ\text{C}$  rms below 1000 m. A number of considerations went into this estimate.

Above 760-m depth the dominant source of temperature uncertainty in the range-average field is the four-day separation between the midpoint of the acoustic measurements and the midpoint of the XBT/XCTD measurements. A rough guess at the possible magnitude of the temperature change was made by assuming that the range-averaged temperature could change  $1^\circ\text{C}$  in 30 days (consistent with possible mesoscale variability), giving a change of  $0.12^\circ\text{C}$  in the 3.6 days between the midpoints of the acoustic and XBT/XCTD measurements. The XBT/XCTD measurements themselves provide a rather tight constraint on the quasi-instantaneous range-average field. The combination of measurement error and internal-wave-induced temperature fluctuations is estimated to give a rms temperature uncertainty of about  $0.1^\circ\text{C}$  for a single measurement. In the range average over 111 profiles this uncertainty is reduced by a factor of approximately 10 to about  $0.010^\circ\text{C}$ . The bias of the XBT temperatures is not expected to exceed  $0.020^\circ\text{C}$ .

At depths below those for which there are XBT/XCTD data, the rms temperature uncertainty in the range average deduced from historical data is rather small. At 3000-m depth, for example, the rms temperature variability is about  $0.025^\circ\text{C}$ . If one assumes a horizontal scale of 100 km, the uncertainty in the range average over the 3250-km-long path is then only about  $0.004^\circ\text{C}$ . The estimated rms uncertainty in the equation used to convert temperature and salinity to sound speed is  $0.05 \text{ m/s}$  (Del Grosso, 1974; Spiesberger, 1993; Dushaw *et al.*, 1993; Meinen and Watts, 1997), corresponding to a rms temperature uncertainty of about  $0.011^\circ\text{C}$ . Because the separation between the source and receiver was not measured nearly as precisely in previous

field tests of the sound-speed equation as in this experiment, however, it is possible that an error in the sound-speed equation was previously incorrectly attributed to an error in the range. The best previous field test of the sound-speed equation had range uncertainties of about 70 m, for example (Dushaw *et al.*, 1993). To account for the uncertainty in the sound-speed equation, the *a priori* rms temperature uncertainty was increased to 0.05 °C below 1000-m depth, which allowed consistent inverse estimates.

### 3. Data

The data are the travel-time perturbations  $\Delta\tau_i$  computed by subtracting the travel times predicted using the XBT/XCTD data from the time-mean measured travel times (Fig. 11). The travel-time uncertainties are a combination of measurement error due to ambient noise, internal-wave-induced travel-time fluctuations, and travel-time fluctuations due to mesoscale variability unresolved by the XBT/XCTD data. The high-frequency travel-time fluctuations about linear trends fit to the ray data are 11–19 ms rms (Colosi *et al.*, 1999). The uncertainty in the time-mean ray travel times is then about 3 ms rms. More significant is the travel-time uncertainty associated with mesoscale range dependence that is not resolved by the XBT/XCTD data and not included in the sound-speed perturbation model. If we assume that travel-time variability between rays with nearly the same lower turning points must be due to unresolved horizontal variability, Fig. 11 suggests that mesoscale-induced travel-time fluctuations are less than 20 ms rms. The uncertainty in the time-mean ray travel times was therefore rather conservatively taken to be 20 ms rms. Finally, the uncertainty in the mean pulse termination travel time is taken to be 100 ms rms, considerably larger than the measured variability of about 23 ms rms, to reflect the uncertainty in interpreting the pulse termination as a surrogate for adiabatic mode 1 and to allow for any internal-wave-induced group delay bias.

Any travel-time biases have been ignored in the inversions, other than to increase the travel-time uncertainties. Assuming a standard Garrett–Munk internal-wave spectrum Colosi and Flatté (1996) found biases of about  $-10$  ms at 3-Mm range from their simulations for a case similar, but not identical, to the one examined here. A negative bias means that the ray is predicted to arrive earlier (have a shorter travel time) in the presence of internal waves than in their absence. Negative biases would slightly increase the discrepancy in Fig. 11 because the predictions would have even shorter travel times, making the measured minus predicted travel times more positive. Colosi *et al.* (1999) find that the observed travel-time variances for the experiment described here are consistent with predictions assuming that the internal-wave field has one-half of the standard Garrett–Munk energy level. Using this energy level would reduce the biases in the simulations from  $-10$  to  $-5$  ms. Further, the observed pulse spreads are only 0–5 ms (Colosi *et al.*, 1999). The pulse spread and bias are expected to be of the same order, and so this result is consistent with the biases found in the simulations. The biases of the rays are therefore comparable to or less than the assumed random errors. Nonetheless,

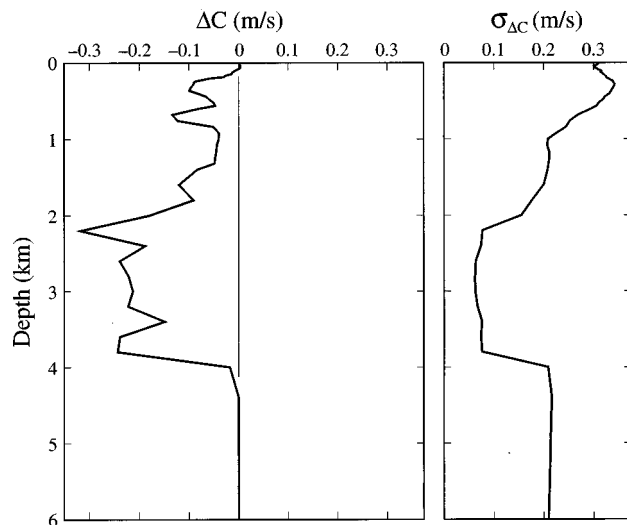


FIG. 14. Range-independent sound-speed perturbation and the associated rms uncertainty as a function of depth.

the biases may still affect the inverse estimates if all of the ray travel times have similar, systematic biases.

The group delays of the lowest-order normal modes are found to be more biased by internal-wave-induced scattering than are the travel times of the steep rays. Colosi and Flatté (1996) found from their simulations that the bias for acoustic mode 1 grows as the square of the range and is about  $-45$  ms at 3-Mm range for the sound-speed profile they used and for the standard Garrett–Munk internal-wave energy level. Using one-half of the standard Garrett–Munk energy level in the simulations would reduce the bias from  $-45$  to  $-22.5$  ms. This bias is substantially smaller than the group delay uncertainty of 100 ms assumed here.

### 4. Results

Finally, inverting the time-mean travel-time perturbations in Fig. 11 for the  $a_j$  gives the mean sound-speed perturbation profile shown in Fig. 14. The linear travel-time residuals are consistent with the *a priori* assumptions (Table III). Retracing rays through the perturbed sound-speed field  $C_0(x, z) + \Delta C(z)$  gives travel times consistent with the measurements, indicating that the reference state selected is sufficiently close to the true state for the inversions to be linear. No iterations of the inverse were done.

The estimated sound-speed perturbations are small above about 1000-m depth, reflecting the constraints imposed by the XBT/XCTD data. The estimated perturbation is

TABLE III. Unweighted travel-time perturbations (rms) for the rays and acoustic modes 1–20 before (data) and after (residuals) inversion. The data are travel-time perturbations relative to the reference state. The residuals are linear estimates of the travel-time perturbations relative to the perturbed sound-speed profile. The values for acoustic mode 1 are shown separately. Only mode 1 was used in the inversion.

	Data (s)	Residuals (s)
Ray travel times (rms)	0.226	0.028
Modal group delays (rms)	0.283	0.228
Mode 1 group delay	0.075	$-0.020$

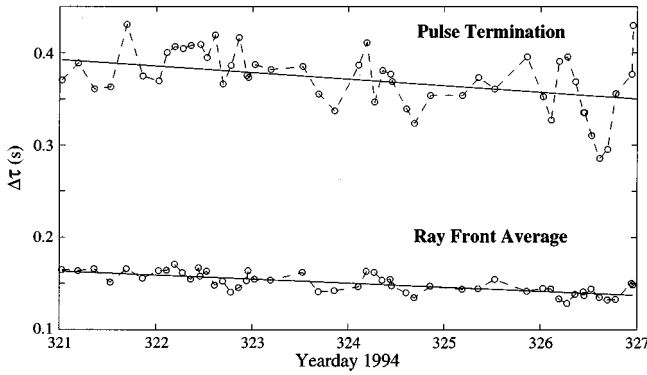


FIG. 15. Travel-time perturbations computed by subtracting the predicted travel times for the WOA94 climatology from the measured travel times, as derived from the pulse termination (top) and the beamformer output for the identified ray wave fronts (bottom). The lower curve is an average over all of the identified wave fronts. A straight-line fit gives a trend of  $-7.1$  ms/day for the pulse termination and  $-4.4$  ms/day for the ray front average.

largest, about  $-0.2 \pm 0.05$  m/s, between 2000 and 4000 m. If the sound-speed perturbation were due entirely to an error in the sound-speed equation, the correction would be three to four times the *a priori* uncertainty in the equation. (Theoretical corrections to the sound-speed equation due to variations in the relative composition of salts in the ocean and due to the dependence of sound speed on acoustic frequency are smaller than 0.05 m/s (Dushaw *et al.*, 1993)). If the sound-speed perturbation were entirely due to temperature, the corresponding temperature perturbation would be  $-0.044 \pm 0.011$  °C, which is 10-11 times the *a priori* range-average temperature uncertainty. Both effects may well be important, although it seems probable that the dominant effect is an error in the sound-speed equation. Accurate deep hydrographic data taken during the time of the acoustic transmissions would be needed to resolve definitively the issue.

One important conclusion from these results is that once the sound-speed equation is better determined, the absolute range-average temperature at depth can be determined acoustically with an estimated uncertainty of about 10 millidegrees, neglecting any effects due to salinity.

## B. Inverting for the trend

In the application to climate studies, the important consideration is the precision with which temperature *changes*, rather than the absolute temperature, can be measured. The measured travel times of both the ray wave fronts and the pulse termination slowly decreased over the six days of the experiment (Fig. 15).

To first order, the time derivative of the travel-time perturbation for ray  $i$  is

$$\frac{d}{dt}(\Delta \tau_i) = - \int_{\Gamma_i} \frac{ds}{C_0^2(x,z)} \frac{d}{dt}[\Delta C(z)]. \quad (9)$$

The sensitivity of the time derivatives of the travel-time perturbations to the time derivatives of the sound-speed perturbations is identical to that of the travel-time perturbations themselves to the sound-speed perturbations.

## 1. Reference state

The range-dependent, time-independent, sound-speed field  $C_0(x,z)$  constructed by combining the XBT/XCTD data with the WOA94 climatology was the reference state, as in the preceding section.

## 2. Sound-speed trend perturbation model

The time derivative of the sound-speed perturbation field  $d[\Delta C(z)]/dt$  was modeled as a sum of the same range-independent vertical modes as in the preceding section, so the vertical modes used previously now represent the time rate of change of the field,

$$\frac{d}{dt}[\Delta T(z)] = \sum_{j=1}^M b_j F_j^T(z), \quad (10)$$

and

$$\frac{d}{dt}[\Delta C(z)] = \sum_{j=1}^M b_j F_j^C(z), \quad (11)$$

where  $b_j$  are the model parameters.

The rms variability in the temperature trend was taken very arbitrarily to be  $0.030$  °C/day at the surface, corresponding to a trend of  $1$  °C/month, dropping nearly linearly to  $0.002$  °C/day rms at 1000-m depth and  $0.001$  °C/day rms at 6000-m depth.

## 3. Data

The data were the linear travel-time trends  $d(\Delta \tau_i)/dt$  computed from least-squares fits to the travel-time series (Fig. 15). The ray-mean travel times of all of the resolved ray arrivals decreased by  $4.4$  ms/day, with a rms variation of  $5.9$  ms rms about the linear trend. The trends from separate linear fits to each ray-travel-time series were used as data in the inverse. The travel time of the pulse termination decreased somewhat more rapidly, by  $7.1$  ms/day, with a rms variation of  $22.9$  ms about the linear trend, as reported previously. The uncertainty in the individual linear trends was taken to be  $3$  ms/day rms, independent of ray path, due to the observed scatter around the individual trends.

## 4. Results

Finally, inverting the travel-time trends for the  $b_j$  gives the sound-speed perturbation trend shown in Fig. 16. The temperature and heat-content perturbation trends are computed from the same  $b_j$  because of the way in which the sound-speed perturbation trend model was developed. Sound speed, temperature, and heat content are all increasing, leading to shorter travel times. The sound-speed and temperature trends both show positive perturbations above 1000-m depth and between 2000- and 4000-m depth. The uncertainty in the temperature estimate is approximately  $0.001$  °C/day below 1000-m depth, increasing to about  $0.030$  °C/day at the surface (i.e., the original uncertainty at the surface). The total temperature change over the six-day period can therefore be estimated with an uncertainty of about  $0.006$  °C below 1000-m depth, corresponding to a sound-speed uncertainty of about  $0.03$  m/s. Not unexpectedly, the total sound-speed

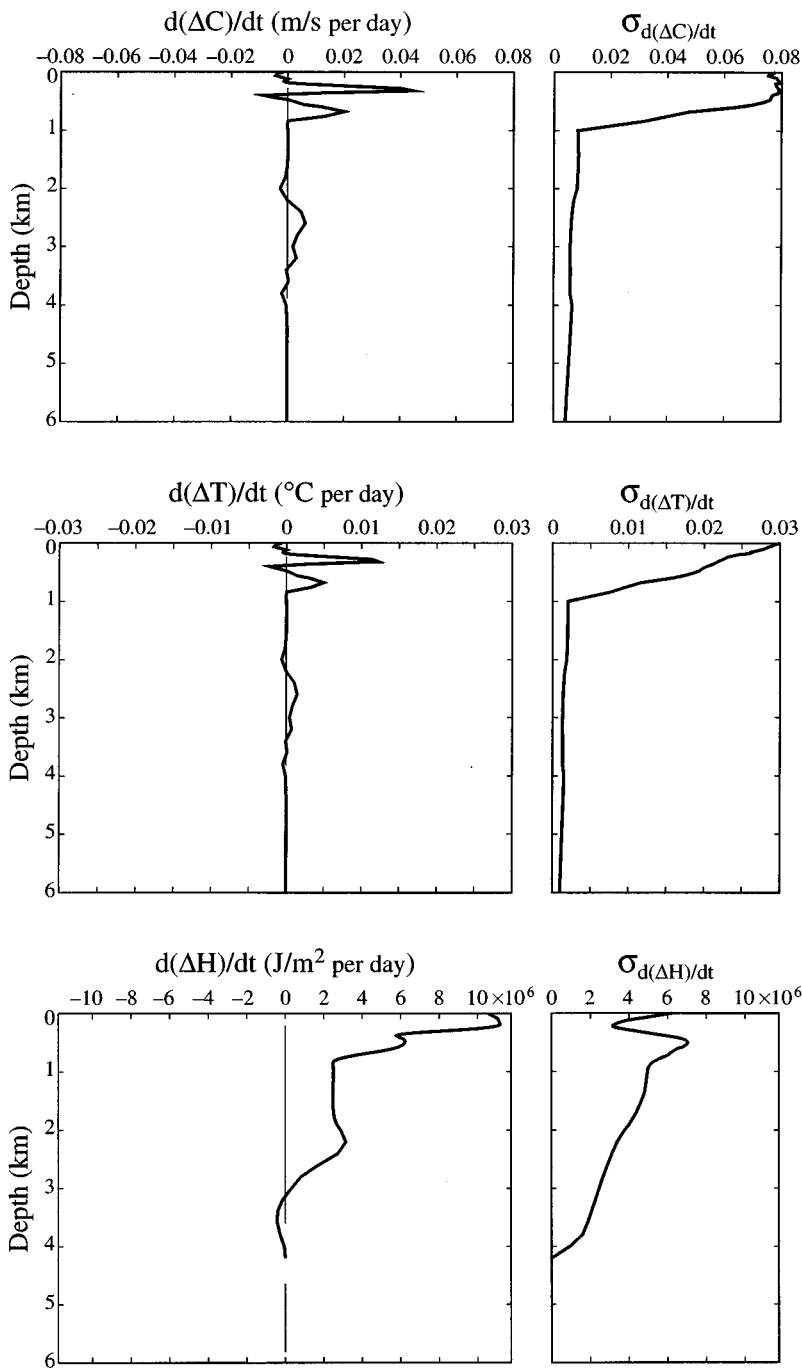


FIG. 16. Rate of change of sound speed, temperature, and heat content and the associated rms uncertainties as a function of depth over the duration of the experiment, as inferred from the linear trends shown in Fig. 15.

change over the six days can be estimated with an uncertainty roughly comparable to that with which the absolute sound-speed change could be determined in the previous section.

The correlated perturbations near the surface and between 2000 and 4000 m are a reflection of up-down ambiguity (Munk and Wunsch, 1982). Sound-speed perturbations above and below the sound-channel axis affect the ray travel times similarly, so the inverse estimate distributes the total variance over depth in accord with the *a priori* sound-speed trend uncertainty profile. This effect was less evident in the estimate of the mean perturbation because the XBT/XCTD data constrained the inverse estimate above the axis. There is no similar constraint on the trend. The inverse estimate uncertainty above and below the axis is correlated, however, so

the uncertainty in the depth integral of heat content is less. The reduction in the uncertainty of the heat content trend just beneath the surface, with a minimum at about 100-m depth, occurs where the integral includes both the upper and lower turning points of most of the ray paths.

#### IV. DISCUSSION

The ATOC Acoustic Engineering Test provides a stringent test of our ability to predict absolute travel times because the XBT/XCTD data constrain the sound-speed field in the upper ocean, where maximum variability is expected, and because the range between the source and receiver is known to within a few meters. Somewhat surprisingly, the measured and predicted absolute travel times are found to be

offset by more than is consistent with *a priori* ocean variability estimates derived from climatology. The discrepancy is probably due to an error in the sound-speed equation at high pressures, as simulations and observations of pulse time spread,  $\tau_0$ , strongly suggest that the biases in the measured travel times are minimal. The absolute ray arrival travel times can be inverted to give estimates of sound speed with an uncertainty of about 0.05 m/s below the sound-channel axis, which is about 20% of the *a priori* variability. The uncertainty in the equation used to convert temperature and salinity to sound speed makes conversion of these absolute sound-speed estimates to temperature uncertain. More important, the trend in temperature with time can be estimated with an uncertainty of about 0.001 °C/day below the sound-channel axis, which is about one-half of the *a priori* variability of 0.002 °C/day.

The vertical resolution of the estimates is limited by internal-wave-induced scattering for this acoustic path. Ray arrivals from paths with minimum upper turning depths above about 200 m were resolvable, identifiable, and stable. Rays with upper turning depths deeper than about 200 m were scattered by internal-wave-induced sound-speed fluctuations. Inversions using ray travel times together with the time of the near-axial pulse termination as a surrogate for the group delay of adiabatic mode 1 therefore have limited vertical resolution, with information from the ray arrivals concentrated above 200-m depth and at the corresponding conjugate depths of 2000 to 4000 m and information from the pulse termination concentrated near axial depth. Other data suggest that this result is strongly dependent on the sound-speed profile, however, and that measurements in other geographic regions can be expected to have different resolution properties. Acoustic receptions at about 1000-km range in the northwestern Atlantic Ocean (Spiesberger *et al.*, 1980), for example, contain resolvable, identifiable, and stable ray arrivals with a much wider range of upper turning point depths than are observed at similar range in the northeastern Pacific Ocean (Dushaw *et al.*, 1993; Worcester *et al.*, 1994).

Our conclusion is that it is feasible to use acoustic transmissions to at least 3-Mm range to make high-precision measurements of range-averaged profiles of ocean sound speed and of changes in ocean temperature and heat content. Subsequent to the ATOC Acoustic Engineering Test, the ATOC project has in fact used acoustic sources and receivers in the Northeast Pacific Ocean to obtain travel-time data for many ray paths. These data have been combined with satellite altimeter measurements and a numerical ocean general circulation model using data assimilation methods to produce improved estimates of the ocean state, including heat content and fluxes (ATOC Consortium, 1998).

## ACKNOWLEDGMENTS

Many people contributed to the fielding and analysis of this experiment, including S. Abbott, L. Day, K. Hardy, D. Horwitt, A. Marshall, D. Peckham, A. Rivera, and R. Truesdale from the Scripps Institution of Oceanography (SIO); J. Kemp, E. Scheer, and K. von der Heydt from the Woods Hole Oceanographic Institution (WHOI); and S. Anderson, R. Drever, F. Karig, S. Leach, V. Miller, L. Olson, D. Red-

daway, R. Ryan, R. Stein, J. Watson, and S. Weslander from the Applied Physics Laboratory at the University of Washington. The crew of R/P FLIP and the support personnel at the Marine Physical Laboratory of the Scripps Institution of Oceanography were exceptionally helpful. The crews of M/V Recovery One, R/V Moana Wave, and CSS John P. Tully all contributed to the success of the experiment. This work was supported largely by the Strategic Environmental Research and Development Program through Defense Advanced Research Projects Agency (DARPA) Grant No. MDA972-93-1-0003. Additional support was provided at SIO by the Office of Naval Research (ONR) through Grant No. N00014-97-1-0258. J. Colosi wishes to acknowledge support from an ONR Young Investigator Award, from the J. Lamar Worzel Assistant Scientist Fund, and from the Penzance Endowed Fund in support of scientific staff at WHOI.

- ATOC Consortium (1998). Baggeroer, A. B., Birdsall, T. G., Clark, C., Colosi, J. A., Cornuelle, B. D., Costa, D., Dushaw, B. D., Dzieciuch, M., Forbes, A. M. G., Hill, C., Howe, B. M., Marshall, J., Menemenlis, D., Mercer, J. A., Metzger, K., Munk, W., Spindel, R. C., Stammer, D., Worcester, P. F., and Wunsch, C., "Ocean climate change: Comparison of acoustic tomography, satellite altimetry, and modeling," *Science* **281**, 1327–1332.
- ATOC Instrumentation Group (1995). Howe, B. M., Anderson, S. G., Baggeroer, A. B., Colosi, J. A., Hardy, K. R., Horwitt, D., Karig, F. W., Leach, S., Mercer, J. A., Metzger, K., Jr., Olson, L. O., Peckham, D. A., Reddaway, D. A., Ryan, R. R., Stein, R. P., von der Heydt, K., Watson, J. D., Weslander, S. L., and Worcester, P. F. "Instrumentation for the Acoustic Thermometry of Ocean Climate (ATOC) prototype Pacific Ocean network," *OCEANS'95 Conference Proceedings*, San Diego, CA, 9–12 October 1995, pp. 1483–1500.
- Baggeroer, A. B., Sperry, B., Lashkari, K., Chiu, C.-S., Miller, J. H., Mikhailevsky, P. N., and von der Heydt, K. (1994). "Vertical array receptions of the Heard Island transmissions," *J. Acoust. Soc. Am.* **96**, 2395–2413.
- Boden, L., Bowlin, J. B., and Spiesberger, J. L. (1991). "Time domain analysis of normal mode, parabolic, and ray solutions of the wave equation," *J. Acoust. Soc. Am.* **90**, 954–958.
- Bowlin, J. B., Spiesberger, J. L., Duda, T. F., and Freitag, L. (1993). "Ocean acoustical ray-tracing software RAY," Woods Hole Oceanographic Tech. Rep., WHOI-93-10.
- Chen, C.-T., and Millero, F. J. (1977). "Speed of sound in seawater at high pressures," *J. Acoust. Soc. Am.* **62**, 1129–1135.
- Colosi, J. A. (1997). "Random media effects in basin-scale acoustic transmissions," *Proceedings of the 1997 Aha Hulikoā Winter Workshop*, pp. 157–166.
- Colosi, J. A., and Flatté, S. M. (1996). "Mode coupling by internal waves for multimegahertz acoustic propagation in the ocean," *J. Acoust. Soc. Am.* **100**, 3607–3620.
- Colosi, J. A., Flatté, S. M., and Bracher, C. (1994). "Internal-wave effects on 1000-km oceanic acoustic pulse propagation: Simulation and comparison with experiment," *J. Acoust. Soc. Am.* **96**, 452–468.
- Colosi, J. A., Scheer, E. K., Flatté, S. M., Cornuelle, B. D., Dzieciuch, M. A., Munk, W. H., Worcester, P. F., Howe, B. M., Mercer, J. A., Spindel, R. C., Metzger, K., Birdsall, T. G., and Baggeroer, A. B. (1999). "Comparisons of measured and predicted acoustic fluctuations for a 3250-km propagation experiment in the eastern North Pacific Ocean," *J. Acoust. Soc. Am.* **105**, 3202–3218.
- Cornuelle, B. D., Worcester, P. F., Hildebrand, J. A., Hodgkiss, Jr., W. S., Duda, T. F., Boyd, J., Howe, B. M., Mercer, J. A., and Spindel, R. C. (1993). "Ocean acoustic tomography at 1000-km range using wave fronts measured with a large-aperture vertical array," *J. Geophys. Res.* **98**, 16365–16377.
- Del Grosso, V. A. (1974). "New equation for the speed of sound in natural waters (with comparisons to other equations)," *J. Acoust. Soc. Am.* **56**, 1084–1091.
- Duda, T. F., Flatté, S. M., Colosi, J. A., Cornuelle, B. D., Hildebrand, J. A., Hodgkiss, Jr., W. S., Worcester, P. F., Howe, B. M., Mercer, J. A., and Spindel, R. C. (1992). "Measured wave-front fluctuations in 1000-km

- pulse propagation in the Pacific Ocean," J. Acoust. Soc. Am. **92**, 939–955.
- Dushaw, B. D., Worcester, P. F., Cornuelle, B. D., and Howe, B. M. (1993). "On equations for the speed of sound in seawater," J. Acoust. Soc. Am. **93**, 255–275.
- Dzienciuch, M. A. (1993). "A spectral method for calculation of ocean acoustic normal modes using Chebyshev polynomials," J. Acoust. Soc. Am. **94**, 1785.
- Howe, B. M., Mercer, J. M., Spindel, R. C., Worcester, P. F., Hildebrand, J. A., Hodgkiss, Jr., W. S., Duda, T. F., and Flatté, S. M. (1991). "SLICE89: A single slice tomography experiment," in *Ocean Variability and Acoustic Propagation*, Proceedings of the Workshop on Ocean Variability and Acoustic Propagation, La Spezia, Italy, 4–8 June 1990, edited by J. Potter and A. Warn Varnas (Kluwer Academic, Dordrecht), pp. 81–86.
- Levitus, S. (1982). "Climatological Atlas of the World Ocean," NOAA Professional Paper 13.
- Levitus, S., and Boyer, T. P. (1994). *World Ocean Atlas 1994, Volume 4: Temperature*, NOAA Atlas NESDIS 4.
- Levitus, S., Burgett, R., and Boyer, T. P. (1994). *World Ocean Atlas 1994, Volume 3: Salinity*, NOAA Atlas NESDIS 3.
- Meinen, C. S., and Watts, D. R. (1997). "Further evidence that the sound-speed algorithm of Del Grosso is more accurate than that of Chen and Millero," J. Acoust. Soc. Am. **102**, 2058–2062.
- Menemenlis, D., Fieguth, P., Wunsch, C., and Willsky, A. (1997). "Adaptation of a fast optimal interpolation algorithm to the mapping of oceanographic data," J. Geophys. Res. **102**, 10573–10584.
- Millero, F. J., and Li, X. (1994). "Comments on 'On equations for the speed of sound in seawater' [J. Acoust. Soc. Am. **93**, 255–275 (1993)]," J. Acoust. Soc. Am. **95**, 2757–2759.
- Munk, W., and Wunsch, C. (1982). "Up-down resolution in ocean acoustic tomography," *Deep-Sea Res.* **29**, 1415–1436.
- Munk, W., Worcester, P., and Wunsch, C. (1995). *Ocean Acoustic Tomography* (Cambridge U.P., Cambridge).
- Munk, W. H., Spindel, R. C., Baggeroer, A. B., and Birdsall, T. G. (1994). "The Heard Island Feasibility Test," J. Acoust. Soc. Am. **96**, 2330–2342.
- Murphy, E. L., and Davis, J. A. (1974). "Modified ray theory for bounded media," J. Acoust. Soc. Am. **56**, 1747–1760.
- National Geophysical Data Center (1988). "ETOPO-5 bathymetry/topography data," Data Announcement 88-MGG-02, Natl. Oceanic and Atmos. Admin., U.S. Department of Commerce, Boulder, CO.
- Sandwell, D. T., and Smith, W. H. F. (1997). "Marine gravity from Geosat and ERS-1 altimetry," J. Geophys. Res. **102**, 10039–10054.
- Smith, W. H. F., and Sandwell, D. T. (1994). "Bathymetric prediction from dense satellite altimetry and sparse shipboard bathymetry," J. Geophys. Res. **99**, 21803–21824.
- Smith, W. H. F., and Sandwell, D. T. (1997). "Global sea floor topography from satellite altimetry and ship depth soundings," *Science* **277**, 1956–1962.
- Spiesberger, J. L. (1993). "Is Del Grosso's sound-speed algorithm correct?" J. Acoust. Soc. Am. **93**, 2235–2237.
- Spiesberger, J. L., and Metzger, K. (1991). "Basin-scale tomography: A new tool for studying weather and climate," J. Geophys. Res. **96**, 4869–4889.
- Spiesberger, J. L., and Tappert, F. D. (1996). "Kaneohe acoustic thermometer further validated with rays over 3700 km and the demise of the idea of axially trapped energy," J. Acoust. Soc. Am. **99**, 173–184.
- Spiesberger, J. L., Spindel, R. C., and Metzger, K. (1980). "Stability and identification of ocean acoustic multipaths," J. Acoust. Soc. Am. **67**, 2011–2017.
- Spiesberger, J. L., Terray, E., and Prada, K. (1994). "Successful ray modeling of acoustic multipaths over a 3000-km section in the Pacific," J. Acoust. Soc. Am. **95**, 3654–3657.
- Sutton, P., Morawitz, W. M. L., Cornuelle, B. D., Masters, G., and Worcester, P. F. (1994). "Incorporation of acoustic normal mode data into tomographic inversions in the Greenland Sea," J. Geophys. Res. **99**, 12487–12502.
- Worcester, P. F., Cornuelle, B. D., Hildebrand, J. A., Hodgkiss, Jr., W. S., Duda, T. F., Boyd, J., Howe, B. M., Mercer, J. A., and Spindel, R. C. (1994). "A comparison of measured and predicted broadband acoustic arrival patterns in travel time-depth coordinates at 1000-km range," J. Acoust. Soc. Am. **95**, 3118–3128.



# Comparisons of measured and predicted acoustic fluctuations for a 3250-km propagation experiment in the eastern North Pacific Ocean

John A. Colosi and Edward K. Scheer

*Woods Hole Oceanographic Institution, Woods Hole, Massachusetts 02543*

Stanley M. Flatté

*Department of Physics, University of California, Santa Cruz, California 95064*

Bruce D. Cornuelle, Matthew A. Dzieciuch, Walter H. Munk, and Peter F. Worcester

*Scipps Institution of Oceanography, La Jolla, California 92093*

Bruce M. Howe, James A. Mercer, and Robert C. Spindel

*Applied Physics Laboratory, University of Washington, Seattle, Washington 98105*

Kurt Metzger and Theodore G. Birdsall

*University of Michigan, Ann Arbor, Michigan 48109*

Arthur B. Baggeroer

*Massachusetts Institute of Technology, Cambridge, Massachusetts 02139*

(Received 9 October 1997; revised 15 January 1999; accepted 5 March 1999)

During the Acoustic Engineering Test (AET) of the Acoustic Thermometry of Ocean Climate (ATOC) program, acoustic signals were transmitted from a broadband source with 75-Hz center frequency to a 700-m-long vertical array of 20 hydrophones at a distance of 3252 km; receptions occurred over a period of six days. Each received pulse showed early identifiable timefronts, followed by about 2 s of highly variable energy. For the identifiable timefronts, observations of travel-time variance, average pulse shape, and the probability density function (PDF) of intensity are presented, and calculations of internal-wave contributions to those fluctuations are compared to the observations. Individual timefronts have rms travel time fluctuations of 11 to 19 ms, with time scales of less than 2 h. The pulse time spreads are between 0 and 5.3 ms rms, which suggest that internal-wave-induced travel-time biases are of the same magnitude. The PDFs of intensity for individual ray arrivals are compared to log-normal and exponential distributions. The observed PDFs are closer to the log-normal distribution, and variances of log intensity are between  $(3.1 \text{ dB})^2$  (with a scintillation index of 0.74) for late-arriving timefronts and  $(2.0 \text{ dB})^2$  (with a scintillation index of 0.2) for the earliest timefronts. Fluctuations of the pulse termination time of the transmissions are observed to be 22 ms rms. The intensity PDF of nonidentified peaks in the pulse crescendo are closer to a log-normal distribution than an exponential distribution, but a Kolmogorov–Smirnov test rejects both distributions. The variance of the nonidentified peaks is  $(3.5 \text{ dB})^2$  and the scintillation index is 0.92. As a group, the observations suggest that the propagation is on the border of the unsaturated and partially saturated regimes. After improving the specification of the ray weighting function, predictions of travel-time variance using the Garrett–Munk (GM) internal-wave spectrum at one-half the reference energy are in good agreement with the observations, and the one-half GM energy level compares well with XBT data taken along the transmission path. Predictions of pulse spread and wave propagation regime are in strong disagreement with the observations. Pulse time spread estimates are nearly two orders of magnitude too large, and  $\Lambda$ – $\Phi$  methods for predicting the wave propagation regime predict full saturation.  
© 1999 Acoustical Society of America. [S0001-4966(99)04606-8]

PACS numbers: 43.30.Re, 43.30.Cq, 43.30.Qd [SAC-B]

## INTRODUCTION

For a week in November of 1994 an Acoustic Engineering Test (AET) for the Acoustic Thermometry of Ocean Climate (ATOC) program was carried out to establish the precision with which basin-scale ocean temperatures can be measured, to determine what vertical resolution might be obtained in a single vertical slice connecting the source and the

receiver, and to understand the influences smaller scale processes like internal waves and mesoscale eddies may have on the signals (see the preceding article<sup>1</sup>). In this paper the focus is on internal-wave effects and the potential predictability of these effects using analytic acoustic fluctuation theories based on internal-wave dominance.

The theory of wave propagation in random media (WPRM) for weak fluctuations was summarized in the 1960s

in the monograph by Tatarskii<sup>2</sup> on optical wave-propagation through homogeneous isotropic turbulence. Much of the early work in ocean acoustics involved unsuccessful efforts to graft the Tatarskii theory, with its assumptions of homogeneous isotropic fluctuations, into the ocean environment. With the realization in the early 1970s that the fluctuations in ocean sound speed were dominated by internal waves which were neither homogeneous nor isotropic<sup>3,4</sup> a significant step forward was made in predicting acoustic fluctuations. Munk and Zachariasen's predictions of the variances of acoustic phase and log intensity from weak fluctuation theory were within a factor of 2 of the observations.<sup>4</sup> In addition to the improvement in understanding the ocean sound-speed fluctuation field, a key theoretical breakthrough was the application of path integral techniques pioneered by Dashen, Flatté, and co-workers<sup>5-9</sup> which led to the formulation of analytical expressions for quantities like pulse spread, travel-time bias, and coherences as a function of vertical, temporal, and horizontal separations. Much of this early work is summarized in a monograph by Flatté *et al.*<sup>10</sup> and later in a review article by Flatté.<sup>11</sup>

Comparisons of measurements with the theory of Flatté, Dashen, and co-workers have been made mostly for short-range and high-acoustic frequency and in general reasonable agreement has been shown.<sup>7,10,12-16</sup> However, due to the short ranges of these experiments, the internal wave effect was generally small and difficult to quantify. We present observations and predictions for the 3250-km AET transmission experiment in the Eastern North Pacific. The AET data serves as a good testing ground for the Flatté and Dashen theory; a vertical array was used so that wavefronts could be unambiguously identified with specific ray paths and measurements of travel-time fluctuations and pulse spread were made. One drawback to the dataset is the low pulse transmission rate which was as low as one transmission every 4 h and as rapid as one every 2 h. However, 47 pulses were recorded on two vertical line arrays (VLAs), one 3250 km away from the source off the island of Hawaii and a second located 80 km distant. Low-frequency basin-scale transmissions, like those of the AET, represent a regime in wave propagation where the Flatté and Dashen theory has not been tested.

An important WPRM issue is to delineate the general regimes of acoustic wave propagation in the ocean as described by fully saturated, partially saturated, and unsaturated propagation.<sup>10,17,18</sup> This has been done using the phase variance calculated in weak fluctuations,  $\Phi^2$ , and the diffraction parameter,  $\Lambda$ , which is the mean square value of the ratio of the first vertical Fresnel zone radius to the vertical correlation length of sound-speed fluctuations.<sup>2,10</sup> But the calculation of  $\Lambda$  is fundamentally single-frequency because of its use of the Fresnel radius, and the AET transmitted pulses with a 75-Hz center frequency and a 37.5-Hz (3-dB) bandwidth. The AET observations indicate that the propagation is in the unsaturated or barely partially saturated regime, while predictions of  $\Lambda$  and  $\Phi$  suggest full saturation. This result suggests that a fully broadband theory for acoustic fluctuations is needed.

The observation of nearly unsaturated propagation has important implications for ocean acoustic tomography (see

the preceding article in this issue<sup>1</sup>). First, unsaturated propagation implies that sound travels very close to the geometrical optics ray paths, and ray theory is the basis of the tomographic inverse problem. Second, unsaturated propagation implies that travel-time biases from internal waves are small. In fact, the observation of pulse time spreads between 0 and 5.3 ms rms suggests that the bias is of the same order of magnitude.

A second critical issue in WPRM theory is the specification of the effective correlation length ( $L_p$ ) of the sound-speed fluctuations along a geometrical optics ray path.<sup>10</sup> The quantity  $L_p$  gives the sensitivity of the sound field to the sound speed fluctuations as a function of position along a ray path. The quantity  $L_p$  is also important for stochastic tomography of the sound-speed fluctuation field. In this paper a direct numerical approach is taken which shows that  $L_p$  is not as strong a function of ray angle as previous expressions had indicated.<sup>10</sup> Calculations of travel-time variance for identified rays using the direct numerical approach for  $L_p$  and utilizing the Garrett-Munk (GM) internal wave model at one-half the reference energy are in excellent agreement with the observations. The one-half GM level is also consistent with expendable bathythermograph (XBT) data from the experiment, but it must be emphasized that this is not an inverse for the GM energy level.

The general outline of this paper is as follows. In Sec. I we describe the AET conductivity-temperature-depth (CTD) and XBT data which give a rough idea of the depth distribution of internal-wave energy during the experiment. In Sec. I we also present the main acoustical observations of travel-time variance, pulse spread, and intensity probability density function (PDF). In Sec. II we describe our new calculation of  $L_p$  and the important acoustic fluctuation quantities. A brief review of relevant acoustic fluctuation theory is also given in Sec. II. In Sec. III we present the comparisons between observations and theory. Section IV has discussion.

## I. AET DATA

A detailed summary of the AET experiment is given by Worcester *et al.*<sup>1</sup> and a short review is given here.

### A. Experimental review

The ATOC Acoustic Engineering Test (AET) was carried out for a 7-day duration from 17 November to 23 November of 1994. Signals with a center frequency of 75 Hz and bandwidth of 37.5 Hz (3 dB point) were transmitted from a source suspended at 650-m depth from R/P FLIP at a position of 31.03417°N, 123.59033°W near Jasper Seamount (see Fig. 1). The transmission schedule for the AET is described by Worcester *et al.*<sup>1</sup> The signals were received on two vertical line arrays (VLAs), one located at 20.65066°N, 154.00773°W off the Island of Hawaii and another VLA located at 30.24663°N, 123.60817°W near Jasper Seamount. The VLAs and the source positions were monitored using a long baseline navigation system described by the ATOC Instrumentation Group<sup>19</sup> giving a nominal position accuracy of 1 to 2 m. The signals were also received by

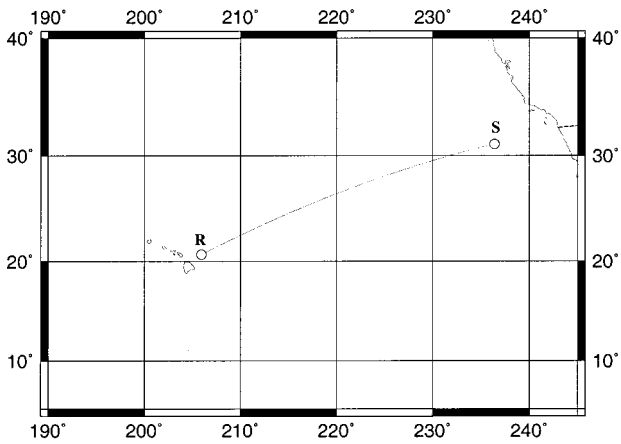


FIG. 1. Geometry of the AET transmissions. The acoustic source (S) suspended from R/P FLIP transmitted to the receiver (R), as well as a second receiver (not shown) only 80 km distant.

several Navy SOSUS arrays located throughout the North Pacific. The focus of this paper will be on the data obtained from the two VLAs at 3252.382 km (Hawaii Island) and 87.325 km (Jasper Seamount) from the source. Both VLAs had 20 hydrophone elements; the Hawaii VLA spanned a depth range of approximately 935 to 1600 m and the Jasper VLA spanned a depth range of approximately 920 to 1585 m. A detailed description of the ATOC instrumentation has been given by the ATOC Instrumentation Group.<sup>19</sup>

## B. Environmental data

Direct measurements of the buoyancy frequency,  $N(z)$ , [ $N^2(z) = -(g/\rho_0)(d\rho/dz)_p$  where subscript  $p$  denotes potential gradient; i.e., total gradient minus the adiabatic gradient] and the potential sound-speed gradient  $(dc/dz)_p$  are important for characterizing the sound-speed fluctuation fields for the experiment. Sound-speed fluctuations in the deep ocean can be expressed in terms of the potential sound-speed gradient and the internal-wave vertical displacement  $\zeta$

$$\langle (\delta c)^2 \rangle = \left( \frac{dc}{dz} \right)_p^2 \langle \zeta^2 \rangle = \left( \frac{dc}{dz} \right)_p^2 \zeta_0^2 N_0 / N(z), \quad (1)$$

where Wentzel–Kramers–Brillouin (WKB) depth scaling has been used for the internal-wave displacement in the final equation which gives a rms displacement  $\zeta_0$  at  $N=N_0$ .

Figure 2 shows profiles of buoyancy-frequency and potential sound-speed gradient calculated from CTDs taken at the Jasper Seamount VLA and the Hawaii VLA at the time of VLA recovery. The Hawaii CTD was taken on 14 February 1995 and the Jasper CTD was taken on 13 January 1995. While there is a 2 to 3 month lag between the CTDs and the transmissions we consider the CTD data to be the best characterization of the buoyancy-frequency and potential sound-speed gradient profiles. These profiles are interpolated in range (some examples are the dashed curves in Fig. 2) for use in the calculation of acoustic fluctuations. Also in Fig. 2 are profiles of  $\langle (\delta c)^2 \rangle$  calculated using Eq. (1) and the buoyancy-frequency and potential-sound-speed-gradient profiles in Fig. 2.

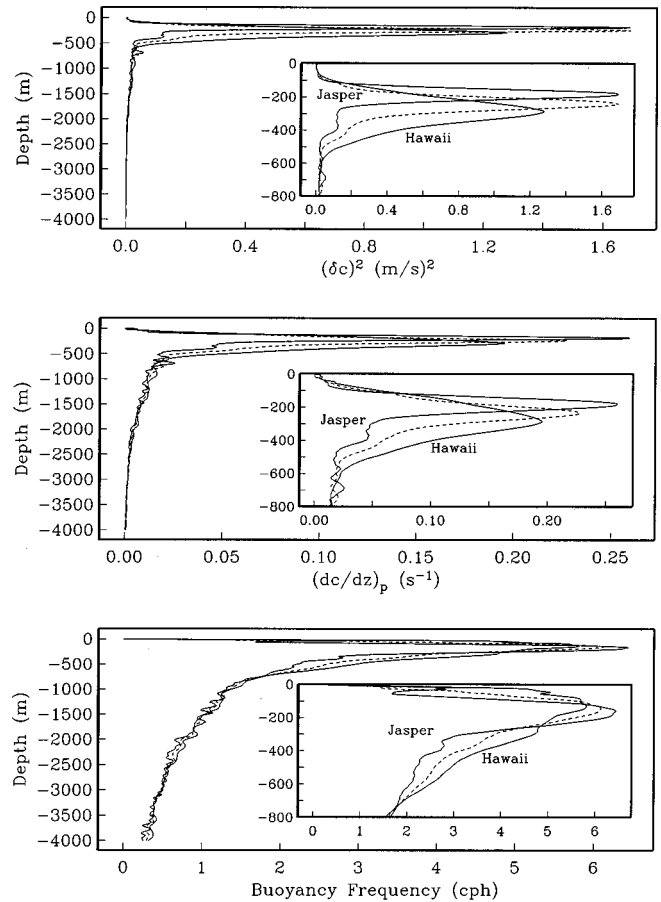


FIG. 2. Profiles of sound speed variance (top), potential sound speed gradient (middle), and buoyancy frequency (bottom) for the Jasper Seamount and the Hawaii VLA. The profiles were derived from CTD casts obtained during VLA recoveries at Jasper Seamount and Hawaii. The dashed curves are interpolated profiles at the range half-way between the source and the receiver. These interpolated profiles are used in the range-dependent calculation of internal-wave fluctuations.

From XBT data taken during the experiment (see Ref. 1) at approximately 30-km range increments between the Jasper Seamount source and the Hawaii VLA, internal displacements have been estimated relative to the Levitus 1994 climatology as

$$\zeta_{xbr}(x, z) = \frac{(\Theta_{xbr}(x, z) - \Theta_{lev94}(x, z))}{(d\Theta_{lev94}/dz)_p(x, z)}, \quad (2)$$

where  $\Theta$  is temperature. To calculate internal-wave displacement variance as a function of depth the range average of  $\zeta$  at each depth is removed and the variance and 30-km range lagged covariance of the demeaned data are calculated. The difference between the zero lag (variance) and the 30-km lagged covariance is a measure of the fluctuations in  $\zeta$  which are incoherent at 30-km separation; this is the contribution from internal waves. Figure 3 shows the rms internal-wave displacement fluctuations derived from the XBT data and for comparison the WKB depth-scaled displacements from the Garrett–Munk reference level of  $\zeta_0 = 7.3$  m at  $N_0 = 3$  cph and the buoyancy-frequency profile at the center of the transmission path (see Fig. 2). Below 200-m depth the XBT displacements are between 0.4 and 1.0 times the WKB scaled GM level. In the upper ocean the XBT displacements are signifi-

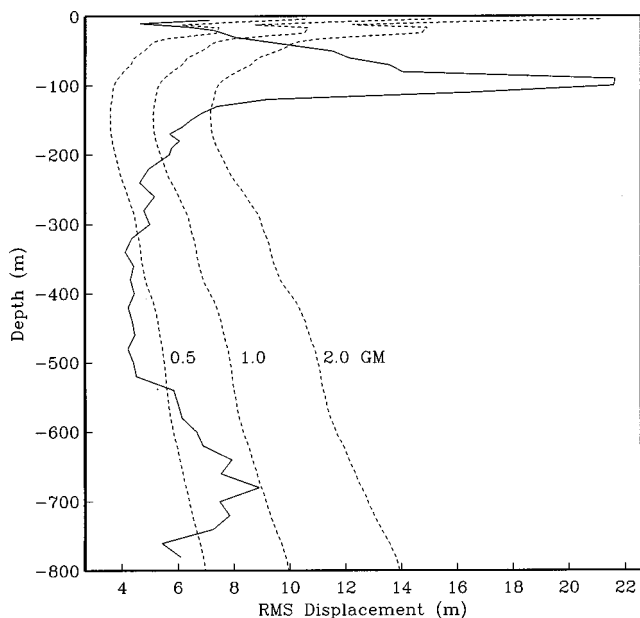


FIG. 3. Profiles of rms vertical displacement derived from the XBT data and Levitus climatology (solid), and calculated using WKB depth scaling at 0.5, 1.0, and 2.0 of the GM reference level (dashed) for the buoyancy profile half-way between the source and receiver (see Fig. 2).

cantly larger than the GM level. The large peak at about 120-m depth corresponds roughly to the bottom of the seasonal mixed layer. The displacement calculations are of questionable validity near the mixed layer due to the rapid changes in the temperature gradient near the mixed-layer interface.

### C. Acoustic data

Figures 4 and 5 show time fronts measured by the Hawaii and Jasper Seamount VLAs. For this paper our primary focus will be on the Hawaii VLA data. The data from the Jasper Seamount VLA will be used to give an estimate of the transmitted pulse shape.

Figure 4 shows the now familiar long-range arrival pattern where early arriving wave fronts are seen sweeping by the receiver array followed by a finale in which no wave

fronts are evident due to internal-wave smearing.<sup>20–23</sup> An analysis of the wave front portion of the recorded signals and the pulse termination time is presented in this paper. We also analyze the complex finale region by calculating the PDF of nonidentified peaks with the understanding that the scintillations in this region give us important information concerning acoustic scattering. The late, highly variable finale arrival energy can be understood in terms of acoustic normal modes<sup>21</sup> or in terms of rays,<sup>24,25</sup> but we do not make an explicit connection to either picture in the data analysis. In Sec. IC 6 a discussion of the late-arriving energy is given in terms of the pulse termination time of the transmissions and the intensity PDF of peaks in the pulse crescendo.

### 1. Arrival selection and identification

The arrival identification has been described by Worcester *et al.*<sup>1</sup> and is summarized here. Individual wave front segments (see Fig. 4) can be unambiguously identified with specific geometrical optics ray paths from any of a number of oceanographic data bases. Therefore we label each wave front with an ID where abs a (ID) is the number of ray turning points and positive/negative IDs are initially pointing towards/away from the ocean surface.

The acoustic arrival times were corrected for source and receiver motion as described by Worcester *et al.*<sup>1</sup> Wave front arrival times for the data were obtained using a geometrical optics wave front prediction as a template which was manually offset to account for a mean travel-time difference and travel-time changes mostly caused by the barotropic tides. Intensity peaks within  $\pm 53$  ms of the geometrical optics wave front were sorted according to the quality function,

$$Q(\text{peak}(t)) = |\Psi(r, t)|^2 \cos(\pi(t - t_p)/W), \quad (3)$$

where  $|\Psi(r, t)|^2$  is the peak intensity,  $t$  is the time of the candidate peak,  $t_p$  is the template ray arrival time, and  $W$  is the window size, 106 ms. This algorithm considers both the intensity of the peak and the distance of the peak from the template, thereby giving more weight to high intensity peaks and peaks closer to the template. For each wave front at time  $t$  and for every ID and depth  $z$  the peak with the highest  $Q$

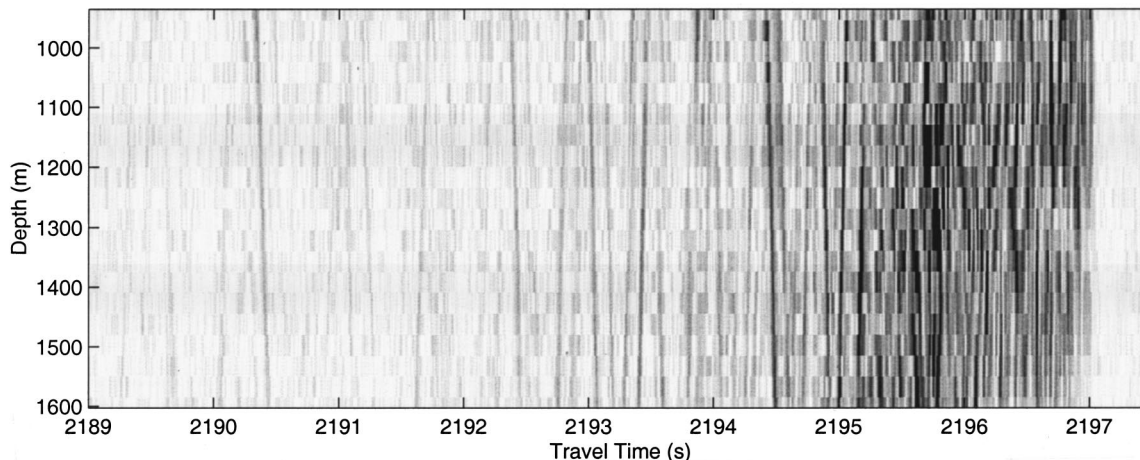


FIG. 4. A measured time front at the AET Hawaii VLA from the transmission started at yearday 326 22:00 h. The gray-scale indicates acoustic amplitude level.

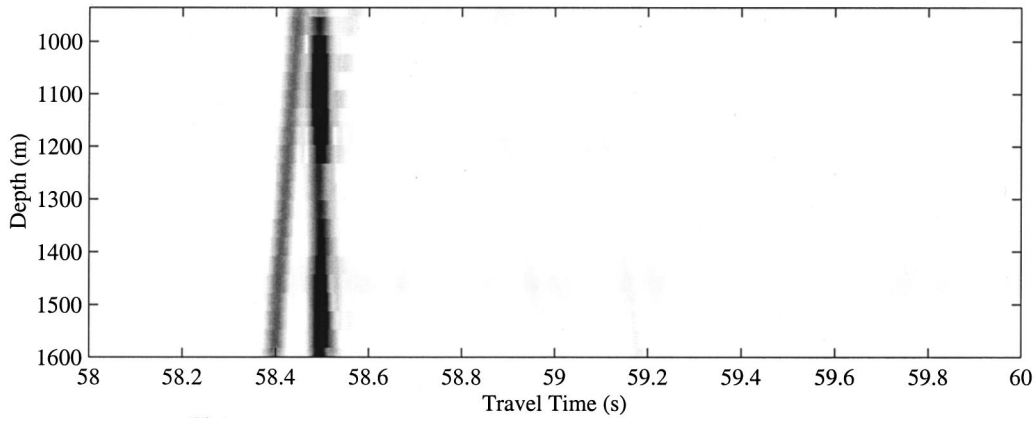


FIG. 5. A measured time front at the AET Jasper Seamount VLA from the transmission started at yearday 324 20:00 h. The earliest arrival in this pattern was used to estimate the transmitted source shape. The gray-scale indicates acoustic amplitude level.

was selected to determine the travel time. Figure 6 shows the geometrical optics template and the picked peaks for one of the AET transmissions. After the initial peak picking the peaks were further edited so that peaks which were attributed to more than one ID, usually near crossing zones or caustics, were rejected, and peaks with SNR less than 12 dB were also rejected. The noise level for each transmission was determined by taking the median value of the amplitude over the entire 27-s arrival record. A median value approach was chosen over a mean value because the median value is much less sensitive to outliers which do not fit a Gaussian noise model.<sup>26</sup> For a Gaussian distribution the median intensity differs from the mean square value by a factor of  $\ln(2)$ . The 12-dB threshold severely reduces the number of usable peaks for the early arriving wave fronts. ID+126 has the least number of points, 135 out of a possible 940 (20 hydrophones $\times$ 47 transmissions=940), while ID-138 has the most with 708 points (see Table I).

## 2. Extraction of wave front fluctuations

The method of Duda *et al.*<sup>22</sup> is used to define the distinct components of the travel time, decomposing the observed travel time for each wave front ID, depth  $z$ , and geophysical time  $t$  into

$$T(t, z, \text{ID}) = \bar{T}(z, \text{ID}) + T'(t, z, \text{ID}) + T_t(t), \quad (4)$$

where the term  $\bar{T}(z, \text{ID})$  is the mean pulse,  $T'(t, z, \text{ID})$  is the wave front fluctuation that is sought, and  $T_t(t)$  is any wave front fluctuation that is common to all of the IDs and depths. The estimation of  $T_t(t)$  is, by a double average over wave front ID and depth,

$$T_t(t) = \frac{1}{N_{\text{ID}}} \sum_{\text{ID}} \frac{1}{J(\text{ID})} \sum_j T(t, z_j, \text{ID}), \quad (5)$$

where  $N_{\text{ID}}=26$  is the number of IDs and  $J(\text{ID})$  is the number of depth points having all of the time points for a given ID (15 out of 20 on average without the 12-dB threshold). The mean value of  $T_t(t)$  is of no importance in this paper and is removed.  $\bar{T}(z, \text{ID})$  is estimated by

$$\bar{T}(z, \text{ID}) = \sum_i^L \frac{T(t_i, z, \text{ID})}{L(z, \text{ID})}, \quad (6)$$

where  $L$  is the number of elements in the sum over time at each depth and for each wave front ID, and the sum is restricted to depth points and IDs which have all of the time points. The fluctuations,  $T'(t, z, \text{ID})$ , have their common motion removed and have a zero time mean.

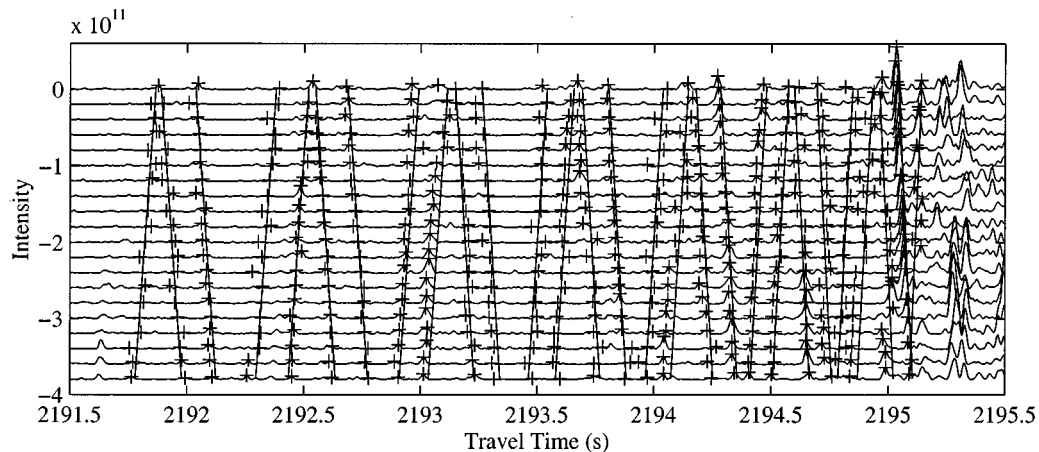


FIG. 6. A measured time front at the Hawaii VLA from the transmission started at yearday 321 00:00 h. Shown with the time front data are the geometrical optics template (sloped vertical lines) and the picked peaks (+).

TABLE I. Average signal to noise ratios (SNR), rms intensity fluctuation, and number of selected peaks for each ID after editing. A total of 940 peaks are possible for the identified wave fronts and 47 peaks are possible for the pulse termination.

(ID)	Average SNR (dB)	rms intensity fluctuation (dB)	No. of peaks
(126/–126)	(13.4/13.2)	(1.6/1.3)	(135/156)
(127/–127)	(14.0/14.3)	(1.7/1.8)	(206/224)
(128/–128)	(14.2/13.9)	(1.9/1.5)	(247/247)
(129/–129)	(14.6/14.5)	(2.0/1.7)	(222/439)
(130/–130)	(14.8/15.0)	(2.1/2.0)	(419/358)
(131/–131)	(14.1/15.1)	(1.6/2.6)	(327/526)
(132/–132)	(14.9/15.1)	(2.2/2.5)	(463/454)
(133/–133)	(14.7/14.8)	(2.5/2.2)	(444/564)
(134/–134)	(15.4/15.3)	(2.6/2.4)	(534/588)
(135/–135)	(15.7/15.7)	(2.8/2.6)	(548/665)
(136/–136)	(15.7/16.5)	(2.4/2.9)	(597/646)
(137/–137)	(16.5/16.2)	(2.9/2.8)	(652/623)
(138/–138)	(16.9/17.5)	(2.9/3.1)	(662/708)
(139/–139)	(18.3/17.1)	(3.1/3.1)	(507/506)
–256 (pulse termination)	...	...	47

Figure 7 shows the estimate of the mean pulse  $\bar{T}(z, \text{ID})$ , and the mean pulse time history  $T_t(t)$ . The trend in  $T_t(t)$  is consistent with the changes underway in the mesoscale sound-speed field,<sup>1</sup> and the high-frequency variations in  $T_t(t)$  appear to be partially accounted for by the barotropic tides [the correlation coefficient between the tide prediction in Fig. 7 and the detrended  $T_t(t)$  series is 0.23]. The series  $T_t(t)$  may have a significant contribution from the baroclinic

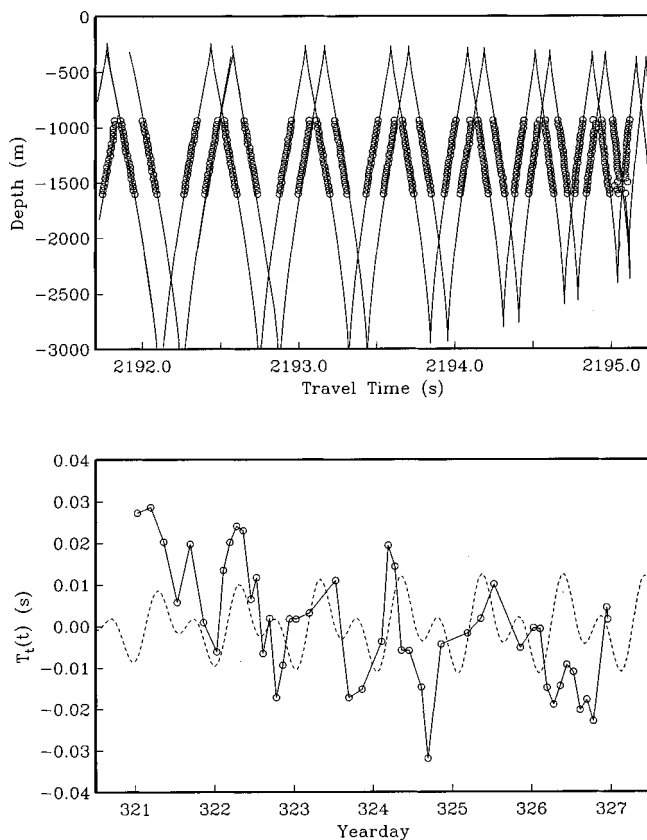


FIG. 7. Upper: Mean pulse  $\bar{T}(z, \text{ID})$  calculated from the AET data (circles) and a ray prediction (solid lines). Lower: The mean pulse time series  $T_t$  (solid) and a barotropic tide prediction (dash).

tides since they are predominantly low-mode internal waves, and the acoustic sampling properties of the identified rays are very similar (see Fig. 11). The series  $T_t(t)$  may also have a contribution from residual mooring or source motion. Figure 8 shows time series of the travel-time fluctuations  $T'$  for hydrophone 10 at a depth of 1270 m. Figure 9 shows the depth variability for some of the observed fluctuations for wave front ID–138. The fluctuations are piecewise continuous, with wave front breaks occurring where there are strong multiple arrivals.

### 3. Travel-time fluctuations

An important issue in determining internal-wave-induced acoustic fluctuations is the time scale of the variability. Coherence times for internal waves are on the order of

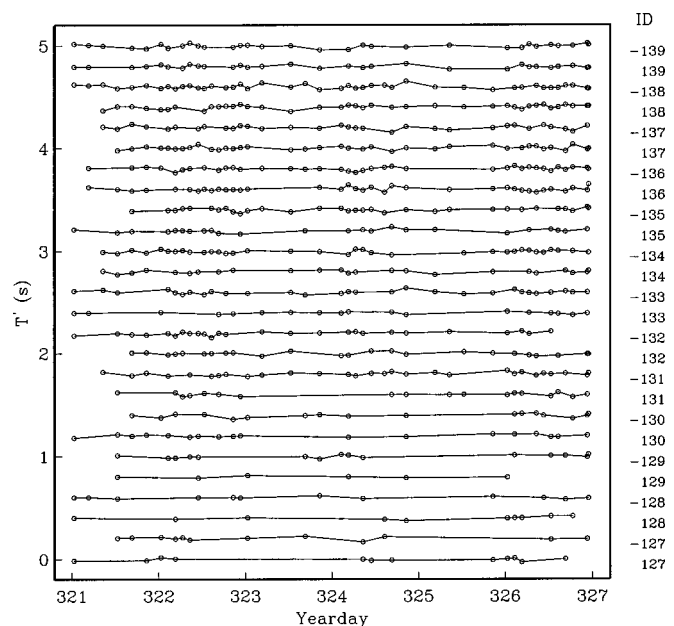


FIG. 8. Travel-time fluctuations  $T'$  for the resolved time fronts for hydrophone 10 at a depth of 1270 m. The timeseries have been offset by 0.2 s.

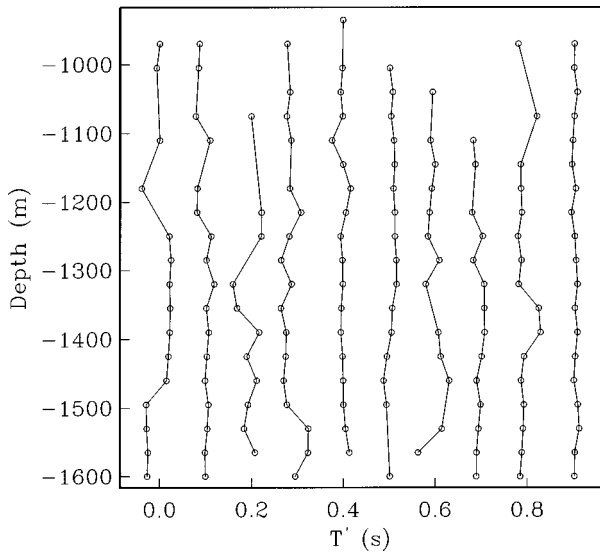


FIG. 9. Travel-time fluctuations  $T'$  for timefront ID-138 at all hydrophones for the first ten transmissions. The fronts have been offset by 0.1 s.

hours while for megameter, low-frequency acoustic propagation through internal waves the acoustic coherence times are on the order of tens of minutes.<sup>9</sup> A calculation of the time-lagged autocovariance functions for each identified wave front and for each depth has been done; an example is shown in Fig. 10 for ID-138 at hydrophone 10. It is found that for all IDs the travel-time fluctuations show no significant time coherence past the first lag (2 h), so the variance can be considered a good measure of the internal-wave effect with little contamination by baroclinic tides.

Figure 11 shows the travel-time variance,  $\tau^2$ , for the identified wave fronts plotted as a function of average upper turning point (UTP) depth, where the average UTP depth was determined from the identified ray path. The identified rays have very similar travel-time variances. The spread of ray UTP depths is only 120 m. There is not much difference in how the identified rays have sampled the ocean. Variance estimates for each ID were calculated at the hydrophones and

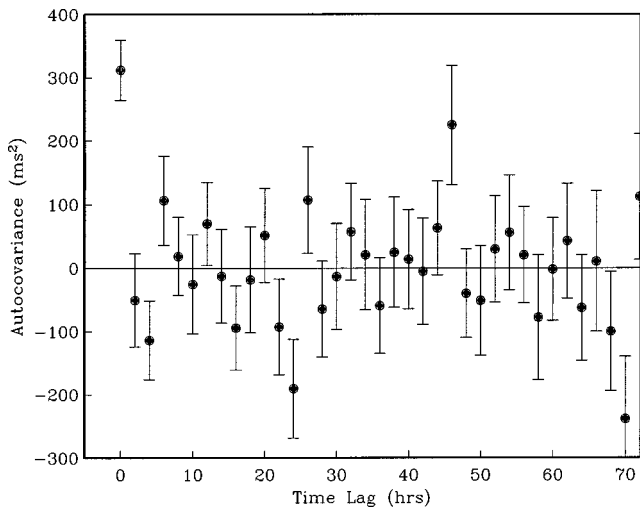


FIG. 10. Time lagged covariance function for wave front ID-138 at hydrophone number 10 (depth=1270 m).

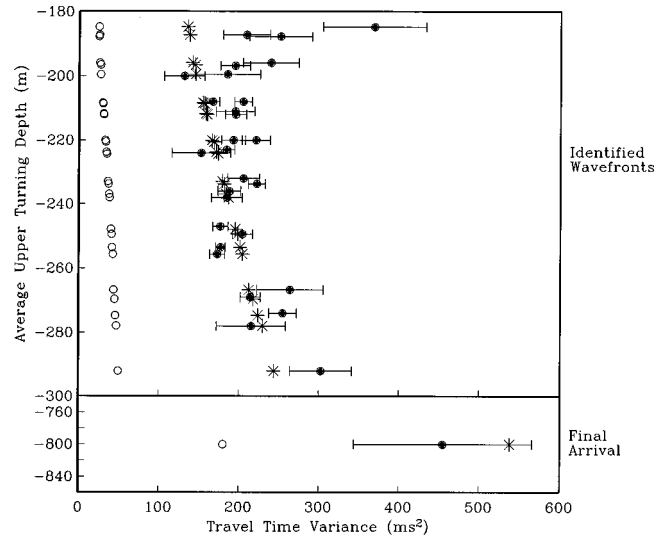


FIG. 11. Travel-time variance for identified wave fronts and the final arrival are shown with error bars (solid circles). Variance estimates using the new calculation of  $L_p$  are displayed with stars and variance estimates using the old evaluation of  $L_p$  with the parabolic-turning-point correction are displayed with open circles. The predictions are calculated using one-half the standard GM internal-wave energy.

then averaged over hydrophones to give a single variance estimate. The uncertainty estimate on the variance was determined from the spread of variance estimates at each hydrophone and assuming independent estimates of variance at each depth.

#### 4. Pulse shapes

The average pulse shape is calculated by aligning the peaks of individual arrivals and summing over hydrophone number and time, with the restriction that the SNR of the peak exceeds 12 dB,

$$\langle I(t, \text{ID}) \rangle = \frac{1}{N_z} \sum_{j=1}^{20} \frac{1}{N_t} \sum_{i=1}^{47} I_{i,j}(\tilde{t} - T(t_i, z_j, \text{ID})), \quad (7)$$

where  $N_z$  is the number of depth points,  $N_t$  is the number of time points, and  $T$  is the travel time of the arrival. The transmitted pulse shape is estimated using the first time front arrival from the Jasper Seamount VLA data (see Fig. 5). This first arrival at the Jasper VLA has an ID of -3. Figure 12 compares the average pulse shapes for the 26 identified time fronts at the Hawaii VLA with the pulse shape at the Jasper Seamount VLA.

The pulse spread,  $\tau_0$ , can be defined in terms of the mutual coherence function for frequency separations,  $\Delta\sigma$ ,

$$\langle \psi^*(\Delta\sigma)\psi(0) \rangle = \exp\left[-\frac{(\Delta\sigma)^2}{2}(\alpha^2 + \tau_0^2)\right], \quad (8)$$

where  $\alpha$  is the transmitted pulse width and  $\tau_0$  is the pulse spread. From the Fourier transform of Eq. (8), we call the  $e^{-0.5}$  point of the pulse envelope,  $\gamma$ , whose value in terms of  $\alpha$  and  $\tau_0$  is  $\gamma = \pm\sqrt{\alpha^2 + \tau_0^2}$ . For the average AET arrival (upper panel of Fig. 12)  $\gamma = 9.75$  ms while  $\alpha = 9.2$  ms. The quadrature difference of these values gives an estimate of 3.2 ms for  $\tau_0$ . Figure 13 shows the pulse spread,  $\tau_0$ , for all

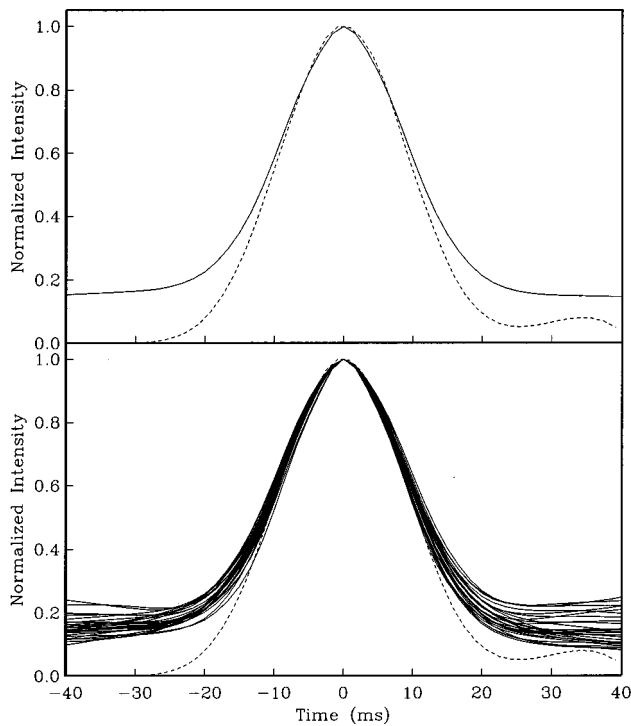


FIG. 12. Lower: Average pulse shapes for 26 identified wave fronts (solid) are shown with the pulse shape observed at the near VLA (dash). Upper: Average pulse shape over all 26 IDs (solid) is shown with the near VLA pulse shape (dash).

identified arrivals as a function of average UTP. Rays with deeper UTP have more spread than shallower UTP rays, and it is evident that the pulse spread  $\tau_0^2$  is much smaller than the travel-time variance  $\tau^2$ .

### 5. Probability density function for intensity

Peak intensities for the identified time fronts are used to calculate the cumulative probability distributions (CPFs) of

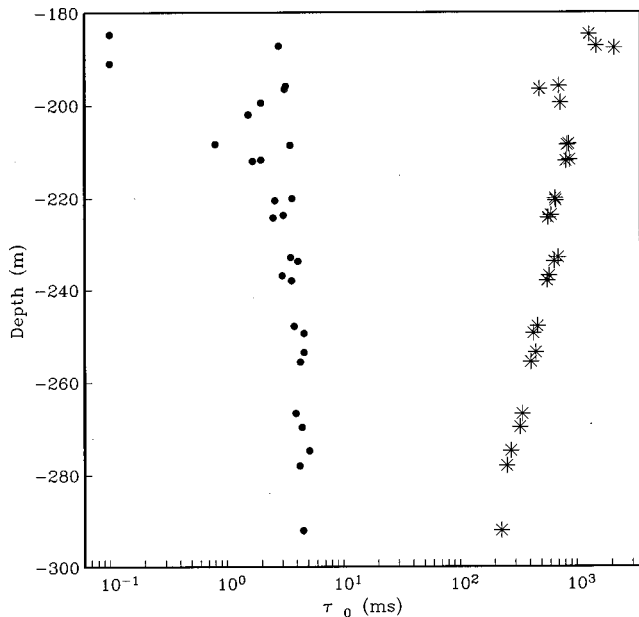


FIG. 13. Predicted (star) and observed (circle) pulse spread,  $\tau_0$  for the identified wave fronts. The predictions are calculated using half the reference GM internal-wave energy level.

intensity for each ID; Fig. 14 shows the observed CPFs for ID-138 compared to CPFs for log-normal and exponential probability density functions (PDFs). The CPF is generated for each ID using data from all hydrophones and all times subject to the restriction that the SNR is above 12 dB. Therefore the computed CPFs from the data should be considered conditional CPFs. For the earliest arrivals the 12 dB threshold severely reduces the data available for a CPF calculation, but for the later arrivals like ID-138 there is very little reduction (see Table I). Figure 14 shows that the intensity variations are closer to a log-normal distribution than to an exponential distribution. The Kolmogorov-Smirnov test<sup>27</sup> for  $D$ , the maximum value of the difference between the cumulative distributions of the data and the model distributions, was performed. For the exponential distribution  $D$  has a value of 0.1768 and the probability of rejecting the null hypothesis is essentially zero (i.e., very low probability that the measurements were drawn from the exponential distribution). For the log-normal distribution  $D$  has a much smaller value of 0.04048, and the probability of rejecting the null hypothesis is 0.19 (i.e., a suggestive but not conclusive probability that the measurements were drawn from a log-normal distribution). The variance of log-intensity for ID-138 is  $(3.1 \text{ dB})^2$  and the scintillation index  $(S_I = \langle I^2 \rangle / \langle I \rangle^2 - 1)$ , is 0.74. Other IDs show the near log-normal CPF, but the total variance decreases for the earlier arrivals, due to the 12-dB threshold.

### 6. The transmission finale

The pulse termination time of the arrival pattern at the Hawaii VLA has been used as a pseudo-adiabatic mode 1 arrival time by Worcester *et al.*<sup>1</sup> in ocean acoustic tomography inversions, and so the travel-time fluctuations of this quantity are of interest. The pulse termination time was determined by averaging the intensities of the shallowest eight hydrophones to create a time record of intensity for each transmission, which was then convolved with a step function to locate the edge of the cutoff. Figure 15 shows the time series of pulse termination times at the Hawaii VLA after correcting for the common-motion offset  $T_i$  and after removing a linear trend. The autocovariance of this time series is also shown in Fig. 15. Some low-frequency variability is still evident; recall that the autocovariance for identified rays (Fig. 10) showed no variance at time scales greater than 2 h. The difference between the zeroth lag covariance and the first lag covariance is used to estimate the internal-wave-induced travel-time variance giving  $455 \text{ ms}^2$ . This variance is greater than that for any of the identified wave front fluctuations (see Fig. 11).

The intensity fluctuations in the last 2 s of the arrival where there are no identifiable wave fronts are also of interest, since the scintillation behavior is closely linked to the scattering physics. A search for peaks with travel times greater than 2195.75 s and with SNR greater than 14 dB over 20 hydrophones and 47 transmissions yielded a total of 23 474 peaks. The intensity CPF of these peaks is shown in Fig. 16. As was the case for the the identified wave fronts, the PDF is closer to log-normal than to exponential; how-



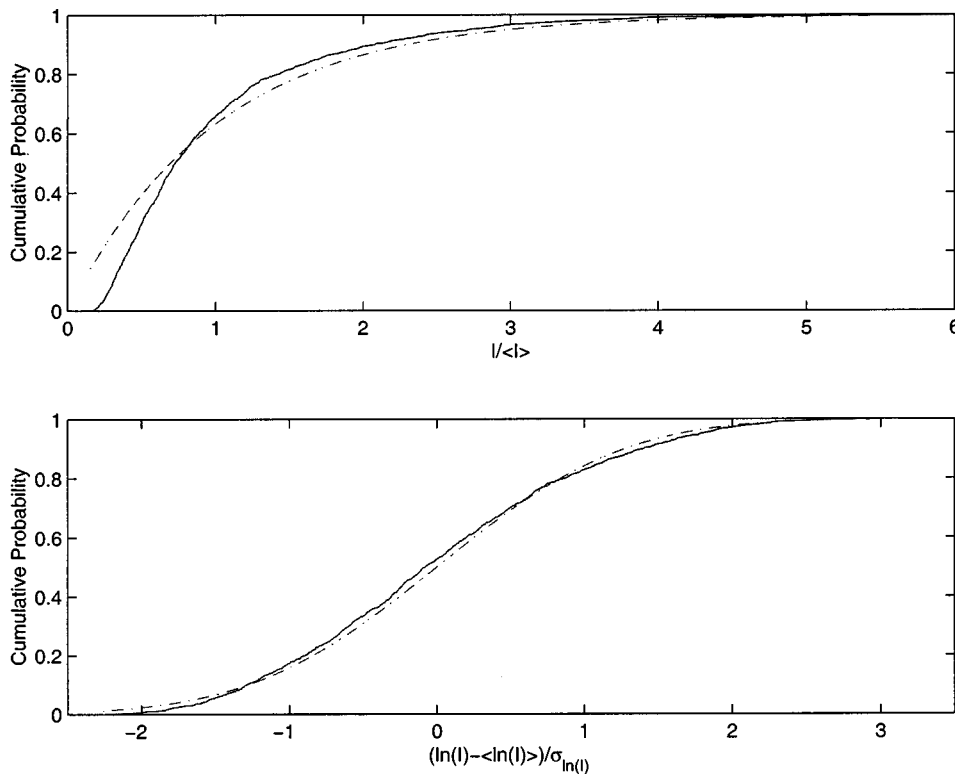


FIG. 14. Observed (solid) and model (dashdot) cumulative probability functions (CPF) for intensity (upper) and log-intensity (lower). In the upper panel we compare the observations to a CPF based on an exponential intensity PDF. In the lower panel we compare the observations to a CPF based on a normal distribution of log-intensity. Observation are for wave front ID-138.

ever, the Kolmogorov–Smirnov test<sup>27</sup> for  $D$  rejects both distributions. The variance of log intensity is  $(3.5 \text{ dB})^2$ , and the scintillation index is 0.92.

## II. CALCULATION OF ACOUSTIC FLUCTUATIONS

### A. Phase fluctuation variance

The strength parameter,  $\Phi$ , which is the rms acoustic phase variability in the geometrical optics approximation, is calculated for a given eigenray path  $\Gamma = z_r(x)$  and is given by

$$\Phi^2 = \sigma^2 \int_{\Gamma} ds \frac{\langle \mu^2(z_r) \rangle L_p(z_r)}{c^2}, \quad (9)$$

where  $\langle \mu^2(z_r) \rangle$  is the fractional sound-speed variance from internal waves, and  $L_p(z_r)$  is an effective internal-wave correlation length along a ray path.

The original method of calculating  $L_p$  for internal waves, which involved approximating the ray by a straight line that is tangent to the ray, has recently been shown to be inaccurate in many situations.<sup>28</sup> An analytic improvement has been used that compares well with numerical simulations of the SLICE89 experiment. The analytic improvement involves approximating the ray as a parabola that is locally tangent to the ray. In this section a different approach is taken which is numerical in nature and makes fewer approximations than the analytic approach of Flatté and Rovner.<sup>28</sup>

Neglecting the effect of internal-wave currents, which have been shown to be a small effect,<sup>10</sup> the phase difference between signals which have traversed fluctuating and non-fluctuating oceans can be written for geometrical optics as

$$\delta\phi = \sigma \left[ \int_{\Gamma} \frac{ds}{c(z) + \delta c(\mathbf{r})} - \int_{\Gamma_0} \frac{ds}{c(z)} \right] \approx \sigma \int_{\Gamma_0} ds \frac{\delta c(\mathbf{r})}{c^2(z)}, \quad (10)$$

where it is assumed that  $\delta c \ll c(z)$  and, as a consequence of Fermat's principle, to a good approximation the path  $\Gamma$  in the presence of fluctuations is the same as the ray path  $\Gamma_0$  for the case of no fluctuations.<sup>17</sup> The strength parameter is given by  $\Phi^2 = \langle (\delta\phi)^2 \rangle$ , so using the expression  $\delta c = (dc/dz)_p \zeta$ , where  $\zeta$  is the internal-wave displacement and  $(dc/dz)_p$  is the potential sound-speed gradient, the internal-wave correlation length along a ray,  $L_p$ , can be written as

$$L_p(s_1) = \frac{1}{\langle \mu^2(s_1) \rangle} \left[ \left( \frac{dc}{dz} \right)_p \right]_{p1} \int_{\Gamma_0} ds_2 \left[ \frac{1}{c^2} \left( \frac{dc}{dz} \right)_p \right]_{p2} \times \rho(\Delta x, \Delta z; \bar{z}), \quad (11)$$

where  $\rho(\Delta x, \Delta z; \bar{z}) = \langle \zeta(s_1) \zeta(s_2) \rangle$  is the correlation function of internal-wave displacements. This approach has been taken by Henyey<sup>18</sup> in examining the validity of the Markov approximation which is used in the ray-tangent approximation for  $L_p$ .

Numerical solution of Eq. (11) for  $L_p(s_1)$  can be carried out efficiently because the integrand is concentrated around  $s_2 = s_1$ . In practice it has been found that precision in  $L_p(s_1)$  to five decimal places can be obtained by calculating the  $s_2$  integral in the following way:

$$L_p(s_1) = \frac{1}{\langle \mu^2(s_1) \rangle} \left[ \left( \frac{dc}{dz} \right)_p \right]_{p1} \int_{x_1 - \Delta}^{x_1 + \Delta} dx_2 \sec \theta_2 \left[ \frac{1}{c^2} \left( \frac{dc}{dz} \right)_p \right]_{p2} \times \rho(\Delta x, \Delta z; \bar{z}), \quad (12)$$

where  $\Delta = 300 \text{ km}$ .

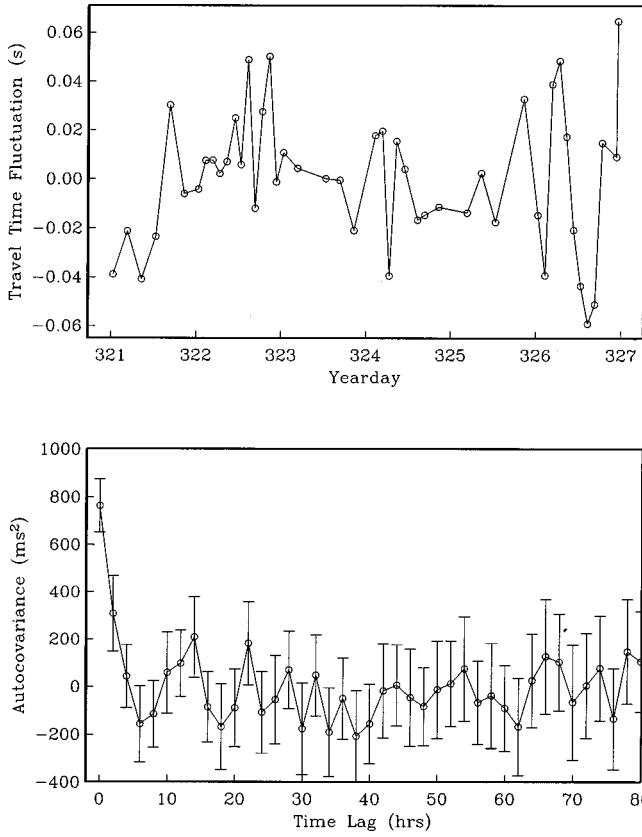


FIG. 15. Upper panel: Time series of the pulse termination time determined from the shallowest eight hydrophones. Lower panel: Autocovariance of the final arrival time series.

The correlation function  $\rho$  is calculated using the internal-wave spectrum,  $S_{\zeta}(j, k_x; \bar{z})$ ,

$$\rho(\Delta x, \Delta z; \bar{z}) = \sum_{j=1}^{j_{\max}} \cos(k_z(j; \bar{z}) \Delta z) \times \int_{-\infty}^{\infty} S_{\zeta}(j, k_x; \bar{z}) \cos(k_x \Delta x) dk_x, \quad (13)$$

where  $j$  is the internal-wave vertical mode number,  $k_z(j; \bar{z})$  is the internal-wave vertical wave number, and  $k_x$  is the internal-wave horizontal wave number along one of the horizontal axes. The Garrett–Munk internal-wave spectrum in terms of vertical mode number,  $j$ , and  $k_x$  is given by<sup>3,10,29</sup>

$$S_{\zeta}(j, k_x; \bar{z}) = \int_{-\infty}^{\infty} dk_y F_{\zeta}(j, k_x, k_y; \bar{z}) = \frac{2\zeta_0^2}{\pi^2 M_j} \frac{N_0}{N(\bar{z})} \frac{G(j, k_x)}{j^2 + j_*^2}, \quad (14)$$

$$G(j, k_x) = \frac{k_j}{k_x^2 + k_j^2} + \frac{1}{2} \frac{k_x^2}{(k_x^2 + k_j^2)^{3/2}} \ln \left( \frac{\sqrt{k_x^2 + k_j^2} + k_j}{\sqrt{k_x^2 + k_j^2} - k_j} \right), \quad (15)$$

with  $k_j = \pi f j / N_0 B$ ,  $M_j = \sum_{j=1}^{\infty} (j^2 + j_*^2)^{-1}$ , and

$$N_0 B = \int_0^{z_b} N(z) dz, \quad N(z) > f, \quad (16)$$

where  $z_b$  is the ocean depth. In the GM model  $N(z)$  is the buoyancy frequency,  $f = 2\Omega \sin(\text{latitude})$  is the local vertical component of the earth's rotation ( $\Omega$  is the angular velocity of the earth), and parameters of the GM model are  $j_* = 3$ ,  $N_0 = 3$  cph, and  $\zeta_0 = 7.3$  m. In the evaluation of Eq. (13), to an excellent approximation the second term in Eq. (15) can be ignored, the spectrum renormalized, and the  $k_x$  integral done analytically, which yields

$$\rho(\Delta x, \Delta z; \bar{z}) = \frac{\rho_0}{M_j} \sum_{j=1}^{j_{\max}} \frac{\cos(k_z(\bar{z}) \Delta z)}{j^2 + j_*^2} e^{-k_j \Delta x}, \quad (17)$$

where the use of the normalization factor  $\rho_0 = \zeta_0^2 N_0 / \sqrt{N(z_1)N(z_2)}$  is a crude attempt to account for the depth nonstationarity of the internal-wave displacement variance. For the calculation of  $\rho$  the WKB approximation for the vertical wave number is used,  $k_z(j; \bar{z}) = \pi j N(\bar{z}) / N_0 B$ , and  $j_{\max} = 100$ . The values of  $N_0 B$  at the Jasper and Hawaii sites derived from the CTD data are 8.28 and 9.74 rad m/s (see Table II).

It must be emphasized here that our use of WKB results is consistent with our use of the GM internal-wave spectrum.<sup>29</sup> The GM model is clearly not an upper ocean model even though we are applying it to the entire water column. We choose the GM model for lack of another model which might describe the upper ocean.

Figure 17 shows the correlation function at the depth where  $N = 3$  cph for the Hawaii and Jasper seamount sites as calculated using Eq. (17). A direct calculation of the correlation function using Eq. (13) and the full expression for the GM spectrum shows that Eq. (17) differs from the exact result at the 1% level. The correlation functions from Eqs. (13) and (17) have nonzero values out to quite large separations, because they have contributions from internal waves with large vertical wavelengths. Our numerical approximation is to truncate the correlation function at the first zero crossing, as shown in Fig. 17. Internal-wave correlation lengths are defined by doing integrals of the correlation functions;

$$L_x = \frac{1}{\rho_0} \int_0^{\infty} \rho(\Delta x, \Delta z = 0) d\Delta x, \quad (18)$$

$$L_z = \frac{1}{\rho_0} \int_0^{\infty} \rho(\Delta x = 0, \Delta z) d\Delta z. \quad (19)$$

Table II shows the values of  $N_0 B$ ,  $L_x$ , and  $L_z$  at values of  $N = 1, 3$ , and 5 cph for the Jasper Seamount and Hawaii VLA sites. Larger values of  $L_z$  occur where  $N$  is small, as can be seen from Eq. (17) and the WKB expression for  $k_z$ . The correlation lengths at the Hawaii VLA are larger than the values at Jasper seamount due to  $N_0 B$  having a larger value at the Hawaii VLA site. The truncation of the correlation function allows for the correlation lengths to be evaluated easily, and it also provides a convenient way of evaluating  $\{k_v^2\} = L_z^{-2}$  which appears in the formulas in the following sections (see Sec. II C).

Figure 18 shows the weighting function,  $\langle \mu^2 \rangle L_p / c^2$ , for ray ID-133 near the Jasper Seamount source calculated from Eq. (11). The weighting function is largest near the ray

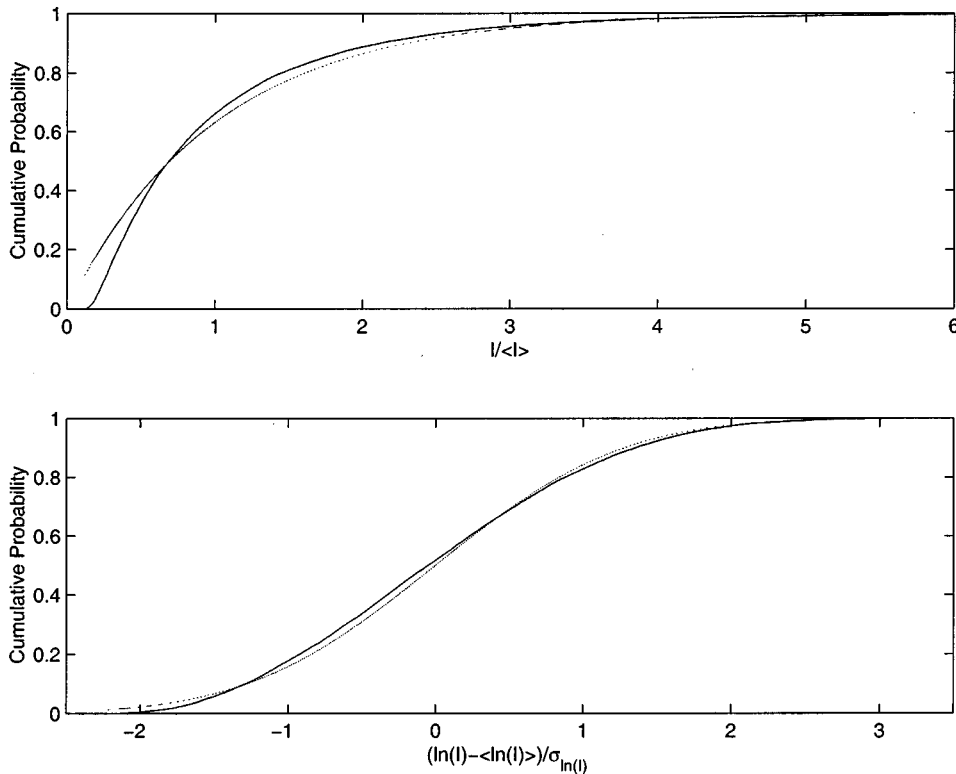


FIG. 16. Observed (solid) and model (dashdot) cumulative probability functions (CPF) for intensity (upper) and log-intensity (lower). In the upper panel we compare the observations to a CPF based on an exponential intensity PDF. In the lower panel we compare the observations to a CPF based on a normal distribution of log-intensity. Observation are for the wave front finale region.

UTP, but has a double peak slightly off the ray apex; this is because of the anisotropy of internal waves, whose vertical correlation length is an order of magnitude less than the horizontal correlation length (see Table II). As discussed in the Appendix, Eq. (11) does not concentrate the internal-wave effect as strongly at the UTP as previous approximations.<sup>4,10</sup> A new analytical approach by Flatté<sup>28</sup> which takes into account the curvature of the ray and the depth variability of  $\langle(\delta c)^2\rangle$  gives results for  $\langle\mu^2\rangle L_p/c^2$  which are qualitatively similar to the direct calculation.

## B. Ensemble-averaged pulse

The average pulse shape is an important measure of acoustic fluctuations, since sound-speed variations can distort the pulse in the saturated and partially saturated regimes. The Fourier transform of the mutual coherence function of frequency combined with the appropriate source transfer function,  $P(\Delta\sigma)$ , gives the ensemble-averaged pulse (EAP),

$$\langle I(t) \rangle = \int_{-\infty}^{\infty} d\Delta\sigma P(\Delta\sigma) \langle \psi^*(\Delta\sigma) \psi(0) \rangle e^{i\Delta\sigma t}. \quad (20)$$

Following Dashen and Flatté,<sup>5,9</sup> in the case of full saturation the mutual coherence function for small frequency separations  $\Delta\sigma$  can be written as

TABLE II. Internal wave correlation lengths.

Location	$N_0 B$ (rad m/s)	$L_x$ (km) ( $N = 3$ cph)	$L_z$ (m) ( $N = 5/3/1$ cph)
Jasper VLA	8.28	14.4	75/123/385
Hawaii VLA	9.74	25.7	88/145/454

$$\langle \psi^*(\Delta\sigma) \psi(0) \rangle \approx \exp \left[ -\frac{(\Delta\sigma)^2}{2} (\tau^2 + \tau_0^2) \right] \cdot \exp(i\Delta\sigma \tau_1) \quad (21)$$

with

$$\tau^2 = \Phi^2 / \sigma^2, \quad (22)$$

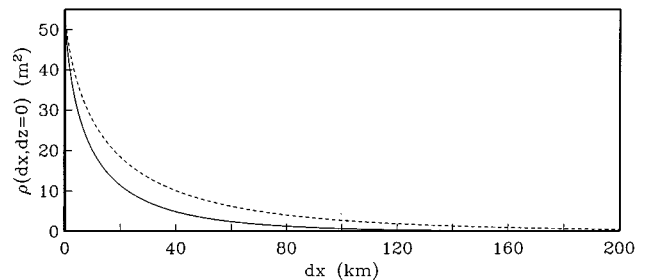
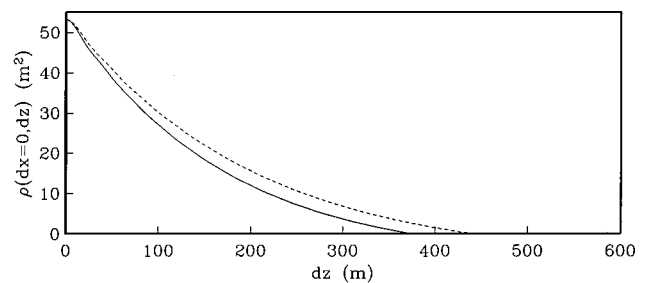


FIG. 17. Internal wave displacement correlation functions at the depth where  $N(z) = 3$  cph at Jasper Seamount (solid) and the Hawaii VLA (dash).

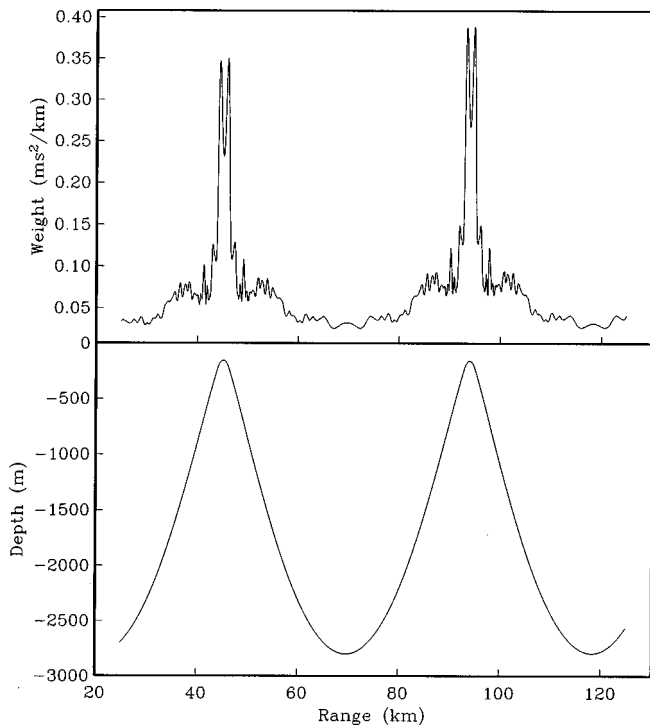


FIG. 18. Ray weighting function,  $\langle \mu^2 \rangle L_p / c^2$ , calculated using the new expression for  $L_p$ , for ray ID-133, for the first two UTPs is shown in the upper panel. The ray path is shown in the lower panel.

$$\tau_0^2 = \left( \frac{\ln(\Phi)}{2c_0} \right)^2 \int_{\Gamma} ds \langle \mu^2(z_r) \rangle L_p(z_r) \{k_v^2\} \times \int_{\Gamma} ds' \langle \mu^2(z'_r) \rangle L_p(z'_r) \{k_v^2\} [g(x, x')]^2, \quad (23)$$

$$\tau_1 = \frac{\ln(\Phi)}{2c_0} \int_{\Gamma} ds \langle \mu^2(z_r) \rangle L_p(z_r) \{k_v^2\} g(x, x). \quad (24)$$

The parameters  $\tau$ ,  $\tau_0$ , and  $\tau_1$  can be interpreted as follows. The quantities  $\tau^2$  and  $\tau_0^2$  combine in quadrature to give the width of the pulse while  $\tau_1$  represents a shift of the mean pulse arrival time (i.e., a travel-time bias). The quantity  $\tau^2$  represents the travel-time variance that one would measure in the unsaturated region. The quantity  $\tau_0^2$  represents the effect of pulse spreading due to loss of coherence between frequencies. Physically this coherence loss can be understood as a result of the interference of many uncorrelated microrays. In practice the magnitude of  $\tau_1$  is very close to the magnitude to  $\tau_0$ , that is to say, in the saturated region the bias and spread come hand-in-hand.<sup>10</sup>

### C. Calculation of $\Lambda$

The diffraction parameter,  $\Lambda$ , is the weighted average along a ray of  $(R_f^2 \{k_v^2\}) / (2\pi)$  where  $R_f(x)$  is the first Fresnel zone radius and  $\{k_v^2\}$  is the spectrum-weighted average value of the square of the vertical wave number of internal waves. Thus

$$\Lambda = \Phi^{-2} \sigma^2 \int_{\Gamma} ds \frac{\langle \mu^2(z_r) \rangle L_p(z_r) \{k_v^2\} R_f^2}{c^2 2\pi}, \quad (25)$$

where  $R_f^2(x) = \lambda |g(x, x)|$ ,  $\lambda$  is the acoustic wavelength, and the Green's function<sup>11</sup>  $g(x, x')$  gives the vertical distance from a ray to a nearby "broken" ray having unit slope discontinuity at  $x = x'$ . In the unsaturated regime the diffraction parameter,  $\Lambda$ , tells us about the variance of log-amplitude<sup>10</sup> ( $\langle \chi^2 \rangle$ ); that is for  $\Lambda \ll 1$ ,  $\langle \chi^2 \rangle = (\frac{1}{4}) \Lambda \Phi^2$ .

The parameters  $\Lambda$  and  $\Phi$  are defined for single-frequency propagation. Because the AET experiment transmitted pulses, the definitions of these parameters are not precise. In particular, the  $\Lambda$  parameter is ambiguous to a large degree because the concept of a Fresnel zone is inherently a single-frequency idea. Nonetheless, we calculate  $\Phi$  so that we can compare to the travel time variance [using Eq. (22)], and we calculate  $\Lambda$  so that together with  $\Phi$  we can predict the wave propagation regime.<sup>10</sup> This has been the standard procedure for all broadband experiments.<sup>10,17</sup>

### III. AET PREDICTIONS

Acoustic fluctuation predictions were made using range-dependent sound-speed profiles derived from 110 XBT profiles taken along the Hawaii VLA acoustic path that were merged onto the Levitus94 climatological profiles for November (see Ref. 1). Two buoyancy-frequency and potential sound-speed gradient profiles obtained from CTD casts at Jasper Seamount and the Hawaii VLA (Fig. 2) were used to characterize the internal-wave fluctuation field (see Sec. IV B). These profiles were interpolated in range along the acoustic path using a triangular interpolation scheme which connects the maximum value of the Jasper profile with the maximum value of the Hawaii profile.

In the calculations of acoustic fluctuations, eigenray paths are generated between the source and the center of the Hawaii VLA (depth = 1270 m) using the range-dependent sound-speed profiles. These are the same eigenray paths discussed by Worcester *et al.*<sup>1</sup> For predictions of the pulse termination time the eigenray with the longest travel time is found using a receiver depth of 900 m. Equations (9), (24), and (25) are integrated in range using the range-interpolated buoyancy-frequency and potential sound speed gradient profiles and the exact expression for  $L_p$  given by Eq. (11). For comparison Eq. (9) is integrated in range using the standard formulas for  $L_p$  (see the Appendix) with the parabolic-turning-point correction. The results of these calculations are described next.

#### A. Travel-time fluctuations

Predictions of  $\tau^2$  calculated from the old and new expressions for  $L_p$  for one-half the GM reference level are shown in Fig. 11. The old expression underestimates the value of  $\tau^2$  by about a factor of 3. Estimates of  $\tau^2$  based on the new expression for  $L_p$  are in excellent agreement with the observations with the biggest disagreement for rays with the shallowest UTPs. These are the rays which interact most with the highly variable upper-ocean internal-wave field, which is not described by the GM correlation function. The one-half GM energy level is also consistent with the XBT displacement data (see Fig. 3 and Sec IB).

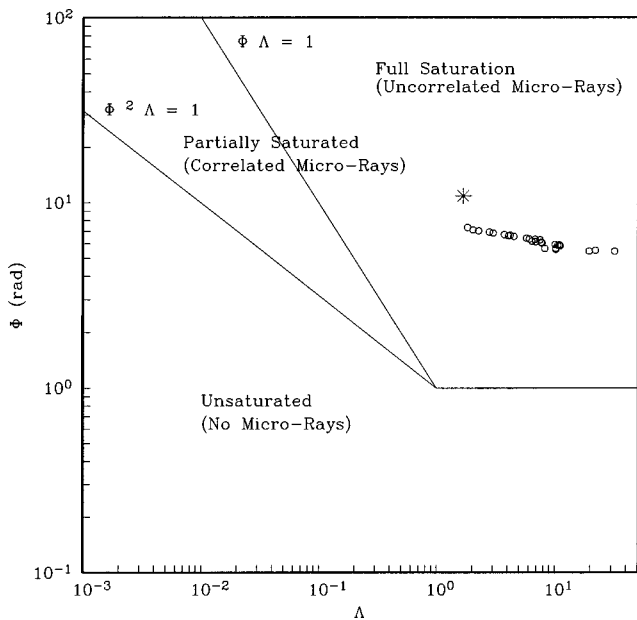


FIG. 19. The  $\Lambda$ - $\Phi$  diagram for identified wave fronts (open circles) and the pulse termination (\*). Calculations are for one-half of the GM reference energy.

### B. Pulse shape

Predictions of pulse spread,  $\tau_0^2$ , calculated using Eq. (24) (we assume that  $\tau_0 \approx \tau_1$ )<sup>9,10</sup> at one-half the GM reference energy level are shown in Fig. 13 together with the observations. The predictions for  $\tau_0$  are 40–1000 times larger than the observations! This result differs from previous results due to the use of the exact expression for  $L_p$ , which does not concentrate the weighting along the ray at the ray UTP where  $g(x,x)$  is close to zero. As will be shown in Sec. III C,  $g(x,x)$ , which is closely linked to the Fresnel zone (see Fig. 20), achieves very large values for long-range

propagation. In addition the predictions show that pulse spread should increase with decreasing ray UTP depth whereas the observations show pulse spread decreasing with decreasing ray UTP depth.

### C. Wave propagation regime: $\Lambda$ vs $\Phi$

Figure 19 shows the  $\Lambda$ - $\Phi$  diagram for the identified rays and the pulse termination. The calculated and observed values of  $\Phi$  are larger than  $2\pi$ , but the  $\Lambda$  values are predicted to be very large in this case, which forces the predictions into the fully saturated regime. On average the Fresnel zone is spread over roughly 60 vertical correlation lengths of the internal-wave field! Figure 20 shows the calculated Fresnel zone for one of the ray IDs and it is clear that the Fresnel zone is comparable in size to the scale of the wave guide. This can be understood qualitatively in terms of the constant-sound-speed Fresnel zone,

$$R_f^2(x) = \lambda x(R-x)/R, \quad (26)$$

where for  $x=R/2$ ,  $R=3000$  km, and  $\lambda=20$  m;  $R_f=3870$  m. For the case of a sound channel or wave guide the envelope of  $R_f^2$  is close to the parabolic shape of Eq. (26), but  $R_f^2$  oscillates between zero at the ray turning points and its maximum somewhere in between.<sup>11</sup> Use of the new formula for  $L_p$  dramatically increases the calculated value of  $\Lambda$  because  $L_p$  is not so strongly peaked at the UTPs where  $R_f$  is close to zero.

All of the arrivals for the  $\Lambda$ - $\Phi$  calculation are predicted to fall well within the saturated region of the diagram, where the pulse is expected to be a complex interference pattern of many uncorrelated microrays. In full saturation the travel-time variance,  $\tau$ , is expected to be much smaller than the pulse spread,  $\tau_0$ , and the intensity PDF is expected to be

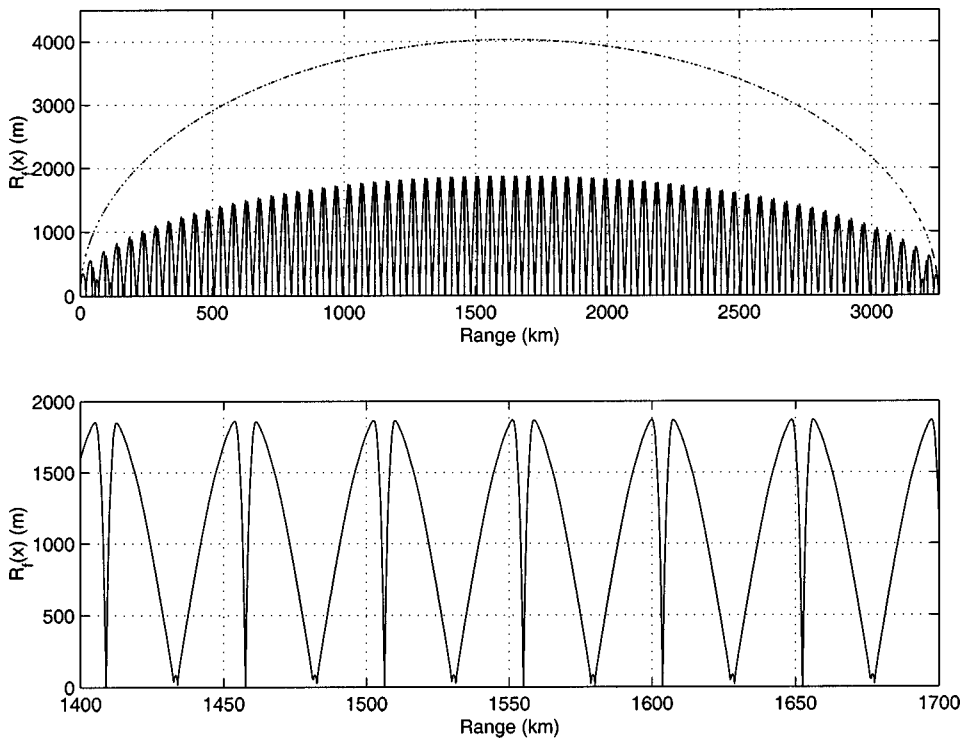


FIG. 20. Calculated Fresnel zone radius for wave front ID-133 (solid) and for the case of no sound channel (dash). The lower panel is a blowup of the center region of the upper panel.

close to exponential. The intensity variance for the exponential distribution is  $(5.6 \text{ dB})^2$  and the scintillation index is unity.

Figures 14 and 16 show that the observed PDFs are closer to log-normal than exponential. Further, as noted above, the observed pulse spread,  $\tau_0$ , is much less than the travel-time variance,  $\tau$ . On the other hand, for the late arrivals and the peaks in the pulse crescendo,  $S_l=0.74$  and  $0.92$ , respectively, and these are well above the weak fluctuation limit of  $0.3$ . The conclusion is that the propagation is not in the weak fluctuation limit. Yet collectively these results imply that the propagation is most likely in the partially saturated or unsaturated regimes rather than saturated.

## IV. DISCUSSION

### A. Ray weighting function for internal waves

Previous analytical approximations to the ray weighting function for sound propagation through internal waves are clearly inaccurate as the Appendix shows (see Fig. 18). Furthermore, the exact evaluation of  $L_p$  dramatically changes the acoustic fluctuation predictions for quantities like  $\tau_0$ ,  $\tau_1$ , and  $\Lambda$ . New and better approximations need to be made to improve the estimates of the ray weighting function without resorting to solving the full integral of Eq. (11). The new method outlined by Flatté and Rovner<sup>28</sup> shows some promise in gaining a better analytical hold on this problem.

### B. Travel-time variance

Predictions of travel-time variance from Eq. (9) with the exact expression for  $L_p$  and for one-half the GM reference energy are in good agreement with the observations for most of the identified arrivals and the pulse termination. The one-half GM level is also consistent with the XBT observations. The biggest discrepancy exists for the shallowest rays, which are preferentially sampling the non-GM internal-wave field near the ocean surface. Differences between acoustic fluctuation predictions made using the GM model and observations of acoustic energy which has sampled the upper ocean have been documented by Colosi<sup>23</sup> for the SLICE89 experiment. The rise in observed travel-time variance for turning depths above 200 m could also be related to the rise in the observed displacement variance derived from the XBTs above 200 m which is shown in Fig. 3. The predictions of  $\tau$  do not show this rise, nor do the WKB-scaled displacements which were used in the calculation of  $\tau$ .

### C. Wave propagation regime: Intensity PDF and pulse spread

We arrive at the very surprising result that the AET experiment is observed to behave as though it were near the boundary of the weak-fluctuation and medium-fluctuation regimes. This is a very different result than other experiments which were conducted at higher frequencies and shorter range.<sup>7,10,12,14–16</sup>

The observed scintillation indices in the AET experiment are between  $0.74$  and  $0.92$ ; these are well above weak-fluctuation values. Since the PDFs are closer to log-normal

rather than exponential, the transmissions appear to be not far from the weak-fluctuation regime, if analogies with optical propagation through turbulence are any guide.<sup>30</sup>

The observed pulse spread,  $\tau_0$ , is smaller than the rms travel-time variance by a factor of  $3$ . This is again consistent with unsaturated or partially saturated behavior. Predictions of  $\tau_0$  are off by two orders of magnitude and the variability of  $\tau_0$  as a function of ray UTP is incorrect. In addition the observation of small  $\tau_0$  suggests that the bias  $\tau_1$  is also small. This is an important result for ocean acoustic tomography since a change in travel time bias cannot be distinguished from a change in heat content.

In comparing PDFs the log-normal (unsaturated) and the exponential (full saturation) PDFs are the only models we have used though other models exist. Ewart<sup>31</sup> proposes the use of the generalized Gamma distribution, which has the log-normal and exponential PDFs as limiting cases. Flatté *et al.*<sup>30</sup> have suggested a log-normal convolved with an exponential distribution, which was developed and tested for optical propagation in the atmosphere. The AET data are consistent with the log-normal distribution and probably cannot distinguish between more complicated distributions.

These results show the dramatic limitations of the CW theory of acoustic fluctuations, and they illustrate the need for a fully broadband fluctuation theory.

### D. Transmission finale

The observed log-normal distribution for the crescendo peaks (Fig. 16) is again a surprise since the present understanding of this region, in ray language, is that the wave field is a complex interference of many deterministic as well as stochastic rays.<sup>24</sup> At the same time, numerical simulations of acoustic propagation through internal waves by Colosi and Flatté,<sup>21</sup> and analytical work by Dozier and Tappert,<sup>32,33</sup> indicate that acoustic normal mode intensities, in this region of strong mode coupling, should obey an exponential PDF. The exact connection between the normal mode PDF and the full-field PDF observed here is unclear.

## ACKNOWLEDGMENTS

This work was supported by the Strategic Environmental Research and Development Program through Defense Advanced Research Projects Agency (DARPA) Grant No. MDA 972-93-1-0003. This work was partially supported by the Office of Naval Research through the Ocean Acoustics Program Grant No. N00014-97-1-0068 and N00014-96-1-0254. John Colosi wishes to acknowledge support from an ONR Young Investigator Award, from the J. Lamar Worzel Assistant Scientist Fund, and the Penzance Endowed Fund in Support of Scientific Staff. This is Woods Hole Oceanographic Institution contribution no. 9585.

## APPENDIX: CALCULATION OF $L_p$

The calculation of  $L_p$ , the correlation length of internal waves along a ray tangent, is well described in the literature (see Ref. 4, 10, or 34), but the ray-tangent approximation is clearly inaccurate near the ray turning depth. In this section a

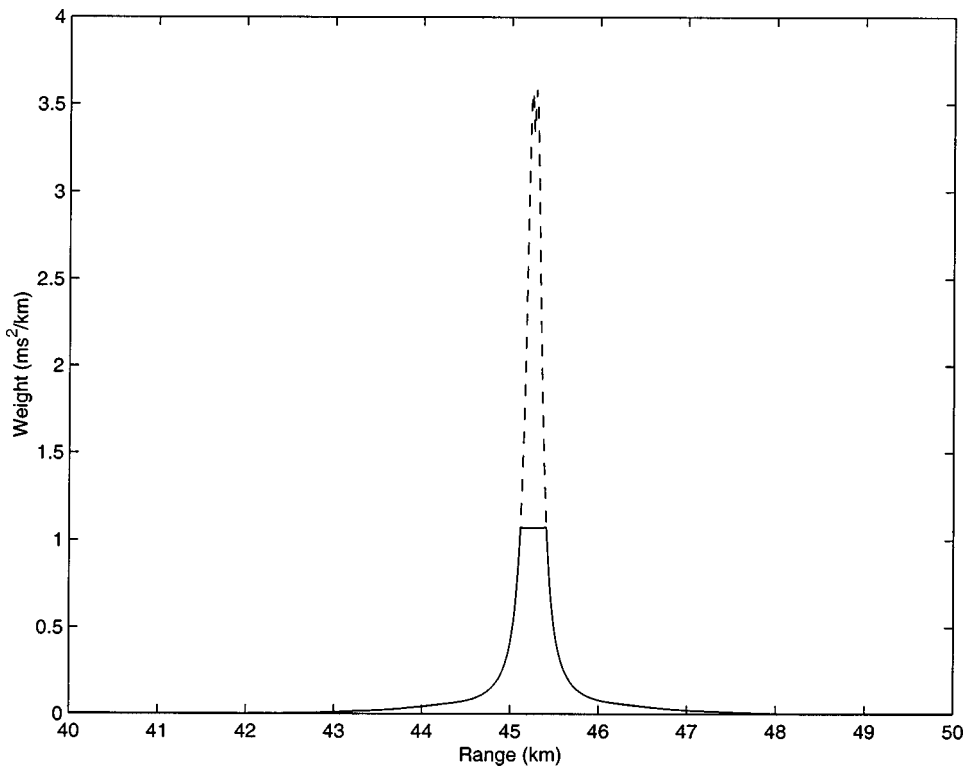


FIG. A1. Ray weighting function  $\langle \mu^2 \rangle L_p / c^2$  along ray ID-133 with (solid) and without (dash) the parabolic-turning-point correction. For this example the profiles of sound speed, buoyancy frequency, and potential sound-speed gradient were independent of range and were taken from the midpoint of the transmission path.

discussion of some analytical modifications to the evaluation of the function  $L_p$  that make the calculation more accurate near ray turning points will be given. The function  $L_p$  is conventionally evaluated by approximating the ray as a straight-line segment over distances on the order of the correlation length of internal waves. That is,

$$L_p(\theta, z) = \frac{1}{\langle \mu^2(z) \rangle} \int_{-\infty}^{\infty} du \rho(s_1, s_1 + (u, u \tan \theta)), \quad (\text{A1})$$

where  $\rho$  is the correlation function of sound speed fluctuations,  $s_1 = (x, z)$  of the ray, and  $\theta$  is the slope of the ray. Away from ray turning points the straight-line approximation can be used, but it clearly fails at the ray apex where  $\theta = 0$ . Therefore a limiting value of  $L_p$  will be derived based upon the curvature of the ray at the upper apex.

The correlation function  $\rho$  can be expressed in terms of the internal-wave displacement spectrum  $S_{\zeta}(\mathbf{k}, z)$  as

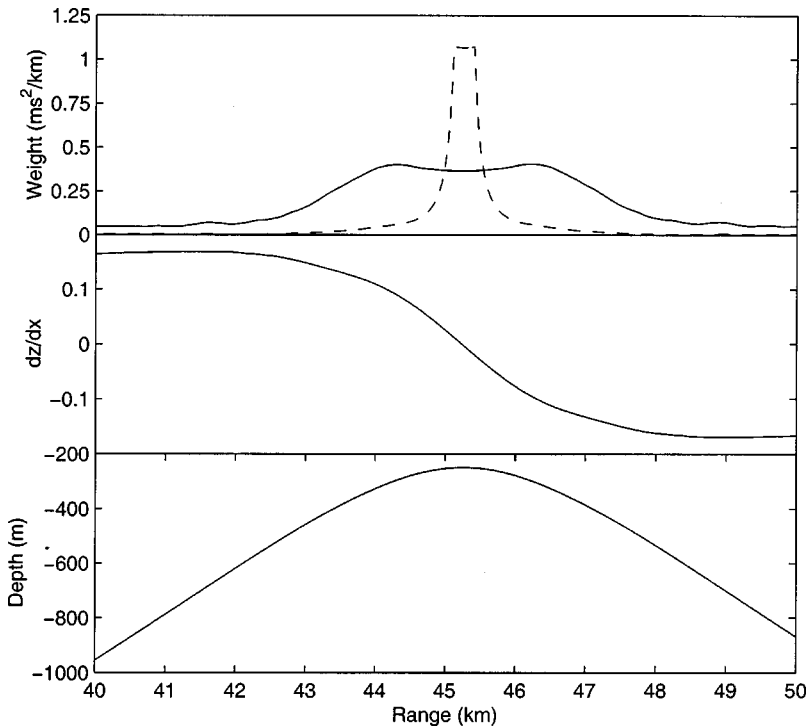


FIG. A2. Upper panel: Ray weighting functions  $\langle \mu^2 \rangle L_p / c^2$  using the parabolic-turning-point correction calculation of  $L_p$  (dash) and the new calculation of  $L_p$  (solid) along ray ID-133. Middle panel: Ray slope. Lower panel: Ray trajectory. For this example the profiles of sound speed, buoyancy frequency, and potential sound-speed gradient were independent of range and were taken from the midpoint of the transmission path.

$$\rho(s_1, s_2) = \text{Re} \left[ \int d\omega \int d^3\mathbf{k} S_\zeta(\mathbf{k}, z) e^{i(k_x \Delta x + k_y \Delta y + k_z \Delta z)} \right]. \quad (\text{A2})$$

For a parabolic turning point,  $\Delta z = au^2$  with  $a = (0.5/c) \times (dc/dz)$ , and assuming  $\Delta y = 0$ ,

$$\rho(s_1, s_1 + (u, au^2)) = \text{Re} \left[ \int d\omega \int dk_z \int k dk \int_0^{2\pi} \frac{d\phi}{2gp} \times S_\zeta e^{i(k \cos(\phi)u + k_z au^2)} \right]. \quad (\text{A3})$$

Then  $L_p$  is given by

$$L_p(a, z) = \frac{1}{\langle \mu^2(z) \rangle} \int d\omega \int dk_z \int k dk S_\zeta \text{Re} [I(a, k, k_z)], \quad (\text{A4})$$

where

$$I(a, k, k_z) = \int_{-\infty}^{\infty} du e^{ik_z au^2} \int_0^{2\pi} \frac{d\phi}{2\pi} e^{ik \cos(\phi)u}. \quad (\text{A5})$$

The angular integral can be done to yield a Bessel function,  $J_0$ , and the  $u$  integral has a closed form solution<sup>35</sup> resulting in

$$\text{Re} [I] = 4\pi \left[ \sqrt{\frac{\pi}{k_z a}} \cos(k^2/(8k_z a) - \pi/4) J_0(k^2/(8k_z a)) \right]. \quad (\text{A6})$$

Using the WKB internal-wave dispersion relation,  $\omega^2 = N^2 k^2/k_z^2 + f^2$ , the WKB vertical wave number relation,  $k_z = \pi N j / N_0 B$ , and the internal-wave spectrum in terms of vertical mode number  $j$  and internal-wave frequency  $\omega$ ,

$$S_\zeta = \langle \mu^2 \rangle N_\omega f \frac{\sqrt{\omega^2 - f^2}}{\omega^3} \frac{N_j}{j^2 + j_*^2}, \quad (\text{A7})$$

the result is

$$L_p(a, z) = 2\pi f \sqrt{\frac{N_0 B}{aN}} \sum_{j=1}^{\infty} \frac{H(j)}{\sqrt{j}} \int_f^N d\omega \frac{\sqrt{\omega^2 - f^2}}{\omega^3} \times \cos(\beta_j(\omega^2 - f^2) - \pi/4) J_0(\beta_j(\omega^2 - f^2)) \quad (\text{A8})$$

with  $\beta_j = \pi j / (8aN_0 B N)$ . For small  $j$  ( $1 \leq j \approx 30$ ), the  $J_0(\beta_j(\omega^2 - f^2))$  term and the  $\cos(\beta_j(\omega^2 - f^2) - \pi/4)$  term can be approximated with the  $\omega = f$  values because the integral is highly peaked for  $\omega = f$ . For  $j > 30$  the weighting of the spectrum is small so to a good approximation the result is

$$L_p(a, z) = \frac{\pi^2}{2} \sqrt{\frac{N_0 B}{aN(z)}} \langle j^{-1/2} \rangle. \quad (\text{A9})$$

This can be compared to the straight-line case where

$$L_p(\theta, z) = \frac{\pi^2}{8} f N_0 B \langle j^{-1} \rangle M_\omega^{-1} \quad (\text{A10})$$

and

$$M_\omega^{-1} = \int_{\omega_L}^N \frac{d\omega}{\omega^3} \left( \frac{\omega^2 - f^2}{\omega^2 - \omega_L^2} \right)^{1/2}, \quad (\text{A11})$$

where  $\omega_L^2 = f^2 + N^2 \tan^2(\theta)$ .

We adopt the rule in calculating the quantity  $L_p$  along the ray path that  $L_p = \min(L_p(\theta, z), L_p(a, z))$ . Figure A1 shows the ray weighting function  $\langle \mu^2 \rangle L_p / c^2$  with and without the parabolic-turning-point correction for one of the ray arrivals. It is clear that the vertical correlation length of internal waves limits the growth of the weighting function.

Figure A2 shows a comparison between the modified weighting function  $\langle \mu^2 \rangle L_p / c^2$  using the parabolic-turning-point correction and the calculation from Eq. (11) which uses fewer assumptions. The influence of the internal-wave field on the acoustic wave propagation is not as strongly concentrated at the UTP as has been previously assumed.<sup>4,10,9</sup>

<sup>1</sup>P. F. Worcester, B. D. Cornuelle, M. A. Dzieciuch, W. H. Munk, B. M. Howe, J. A. Mercer, R. C. Spindel, J. A. Colosi, K. Metzger, T. G. Birdsall, and A. B. Baggeroer, "A test of basin-scale acoustic thermometry using a large-aperture vertical array at 3250-km range in the eastern North Pacific Ocean," *J. Acoust. Soc. Am.* **105**, 3185–3201 (1999).

<sup>2</sup>V. I. Tatarskii, "The effects of a turbulent atmosphere on wave propagation," *Isreal Program for Scientific Translation* (1971).

<sup>3</sup>C. Garrett and W. Munk, "Space-time scales of ocean internal waves," *Geophys. Fluid Dyn.* **2**, 225–264 (1972).

<sup>4</sup>W. H. Munk and F. Zachariassen, "Sound propagation through a fluctuating stratified ocean: Theory and observation," *J. Acoust. Soc. Am.* **59**, 818–838 (1976).

<sup>5</sup>R. Dashen, S. M. Flatté, and S. A. Reynolds, "Path-integral treatment of acoustic mutual coherence functions for rays in a sound channel," *J. Acoust. Soc. Am.* **77**, 1716–1722 (1985).

<sup>6</sup>S. M. Flatté and R. B. Stoughton, "Theory of acoustic measurement of internal wave strength as a function of depth, horizontal position, and time," *J. Geophys. Res.* **91**, 7709–7720 (1986).

<sup>7</sup>S. M. Flatté, S. A. Reynolds, R. Dashen, B. Buehler, and P. Maciejewski, "AFAR measurement of intensity and intensity moments," *J. Acoust. Soc. Am.* **82**, 973–979 (1987).

<sup>8</sup>S. M. Flatté, S. A. Reynolds, and R. Dashen, "Path-integral treatment of intensity behavior for rays in a sound channel," *J. Acoust. Soc. Am.* **82**, 967–972 (1987).

<sup>9</sup>S. M. Flatté and R. B. Stoughton, "Predictions of internal-wave effects on ocean acoustic coherence, travel time variance, and intensity moments for very long-range propagation," *J. Acoust. Soc. Am.* **84**, 1414–1424 (1988).

<sup>10</sup>S. Flatté, R. Dashen, W. Munk, K. Watson, and F. Zachariassen, *Sound Transmission through a Fluctuating Ocean* (Cambridge U. P., Cambridge, 1979).

<sup>11</sup>S. M. Flatté, "Wave propagation through random media: Contributions from ocean acoustics," *Proc. IEEE* **71**, 1267–1294 (1983).

<sup>12</sup>P. F. Worcester, "Reciprocal acoustic transmission in a midocean environment: Fluctuations," *J. Acoust. Soc. Am.* **66**, 1173–1181 (1979).

<sup>13</sup>J. L. Spiesberger and P. F. Worcester, "Fluctuations of resolved acoustic multipaths at long range in the ocean," *J. Acoust. Soc. Am.* **70**, 565–576 (1981).

<sup>14</sup>P. F. Worcester, G. O. Williams, and S. M. Flatté, "Fluctuations of resolved acoustic multipaths at short range in the ocean," *J. Acoust. Soc. Am.* **70**, 825–840 (1981).

<sup>15</sup>S. A. Reynolds, S. M. Flatté, R. Dashen, B. Buehler, and P. Maciejewski, "AFAR measurements of acoustic mutual coherence functions of time and frequency," *J. Acoust. Soc. Am.* **77**, 1723–1731 (1985).

<sup>16</sup>R. B. Stoughton, S. M. Flatté, and B. M. Howe, "Acoustic measurements of internal-wave rms displacement and rms horizontal current off Bermuda in late 1983," *J. Geophys. Res.* **91**, 7721–7732 (1986).

<sup>17</sup>W. Munk, P. Worcester, and C. Wunsch, *Ocean Acoustic Tomography* (Cambridge U. P., Cambridge, 1995).

<sup>18</sup>F. S. Henyey and C. Macaskill, "Sound through the internal wave field," in *Stochastic Modeling in Physical Oceanography* (Birkhauser, Boston, MA, 1996), pp. 141–184.

<sup>19</sup>P. F. Worcester, "Instrumentation for the acoustic thermometry of ocean climate (ATOC) prototype pacific ocean network," *MTS/IEEE Oceans '95 Proceedings* (1995).

<sup>20</sup>J. A. Colosi, S. M. Flatté, and C. Bracher, "Internal-wave effects on 1000-km oceanic acoustic pulse propagation: Simulation and comparison to experiment," *J. Acoust. Soc. Am.* **96**, 452–468 (1994).



- <sup>21</sup>J. A. Colosi and S. M. Flatté, "Mode coupling by internal waves for multi-megameter acoustic propagation in the ocean," *J. Acoust. Soc. Am.* **100**, 3607–3620 (1996).
- <sup>22</sup>T. F. Duda, S. M. Flatté, J. A. Colosi, B. D. Cornuelle, J. A. Hildebrand, W. S. Hodgkiss, P. F. Worcester, B. M. Howe, J. A. Mercer, and R. C. Spindel, "Measured wave-front fluctuations in 1000-km pulse propagation in the Pacific Ocean," *J. Acoust. Soc. Am.* **92**, 939–955 (1992).
- <sup>23</sup>J. A. Colosi, "Random media effects in basin-scale acoustic transmissions," in Proceedings of the 1997 Aha Hulikoá Winter Workshop (1997), pp. 157–166.
- <sup>24</sup>J. Simmen, S. M. Flatté, and G.-Y. Wang, "Wavefront folding, chaos, and diffraction for sound propagation through ocean internal waves," *J. Acoust. Soc. Am.* **102**, 239–255 (1997).
- <sup>25</sup>T. F. Duda and J. B. Bowlin, "Ray-acoustic caustic formation and timing effects from ocean sound-speed relative curvature," *J. Acoust. Soc. Am.* **96**, 1033–1046 (1994).
- <sup>26</sup>J. F. Claerbout, *Fundamentals of Geophysical Data Processing* (McGraw-Hill, New York, 1976).
- <sup>27</sup>W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes in FORTRAN: The Art of Scientific Computing*, 2nd ed. (Cambridge U. P., Cambridge, 1992).
- <sup>28</sup>S. M. Flatté and G. Rovner, "Path-integral expressions for fluctuations in acoustic transmission in the ocean waveguide," in Proceedings of the 1997 Aha Hulikoá Winter Workshop (1997).
- <sup>29</sup>J. A. Colosi and M. G. Brown, "Efficient numerical simulation of stochastic internal wave induced sound speed perturbation fields," *J. Acoust. Soc. Am.* **103**, 2232–2235 (1998).
- <sup>30</sup>S. M. Flatté, C. Bracher, and G.-Y. Wang, "Probability-density functions of irradiance for waves in atmospheric turbulence calculated by numerical simulation," *J. Opt. Soc. Am. A* **11**, 2080–2092 (1994).
- <sup>31</sup>T. E. Ewart, "A model of the intensity probability distribution for wave propagation in random media," *J. Acoust. Soc. Am.* **86**, 1490–1498 (1989).
- <sup>32</sup>L. B. Dozier and F. D. Tappert, "Statistics of normal mode amplitudes in a random ocean. I. Theory," *J. Acoust. Soc. Am.* **63**, 353–365 (1978).
- <sup>33</sup>L. B. Dozier and F. D. Tappert, "Statistics of normal mode amplitudes in a random ocean. II. Computations," *J. Acoust. Soc. Am.* **64**, 533–547 (1978).
- <sup>34</sup>R. Esswein and S. M. Flatté, "Calculation of the phase-structure function density from oceanic internal waves," *J. Acoust. Soc. Am.* **70**, 1387–1396 (1981).
- <sup>35</sup>I. S. Gradshteyn and I. W. Ryzhik, *Table of Integrals, Series and Products* (Academic, New York, 1965).

# Geoacoustic inversion via local, global, and hybrid algorithms

Mark R. Fallat and Stan E. Dosso<sup>a)</sup>

*School of Earth and Ocean Sciences, University of Victoria, Victoria, British Columbia V8W 3P6, Canada*

(Received 27 April 1998; accepted for publication 26 February 1999)

In this paper, local, global, and hybrid inversion algorithms are developed and applied to the problem of determining geoacoustic properties by minimizing the mismatch between measured and modeled acoustic fields. Local inversion methods are sensitive to gradients in the mismatch and move effectively downhill, but generally become trapped in local minima and must be initiated from a large number of starting points. Global inversion methods use a directed random process to search the parameter space for the optimal solution. They include the ability to escape from local minima, but as no gradient information is used, the search can be relatively inefficient. Hybrid inversion methods combine local and global approaches to produce a more efficient and effective algorithm. Here, downhill simplex (local) and simulated annealing (global) methods are developed individually and combined to produce a hybrid simplex simulated annealing algorithm. The hybrid inversion is found to be faster by more than an order of magnitude for a benchmark testcase in which the form of the geoacoustic model is known. The hybrid inversion algorithm is also applied to a testcase consisting of an unknown number of layers representing a general geoacoustic profile. Since the form of the model is not known, an underparameterized approach is employed to determine a minimum-structure solution. © 1999 Acoustical Society of America. [S0001-4966(99)01606-9]

PACS numbers: 43.30.Pc, 43.60.Pt [DLB]

## INTRODUCTION

The problem of determining ocean-bottom properties from measurements of ocean acoustic fields has received considerable attention in recent years (e.g., Refs. 1–10). Geoacoustic inversion is representative of a class of nonlinear geophysical inverse problems for which a direct solution is not available. These problems can be formulated by assuming a discrete form of the model of unknown parameters  $\mathbf{m} = \{m_i, i = 1, \dots, M\}$ , and defining an objective or cost function  $E(\mathbf{m})$  which represents the mismatch between the measured data and replica data computed for a particular realization of the model. The goal of the inversion is to determine the set of model parameters that minimizes the mismatch. This can be a challenging problem due to the size of the parameter space, which increases geometrically with the number of parameters, and the presence of local minima due to the nonlinearity of the inverse problem. Local, global, and (recently) hybrid inversion methods have been applied to geophysical inverse problems. The goal of this paper is to consider each of these approaches in order to develop an efficient and effective algorithm for geoacoustic inversion.

Local inversion methods are based on iteratively improving a starting model by moving down the local gradient of the objective function. Since gradient information is utilized, these methods move efficiently downhill, but typically become trapped in the local minimum closest to the starting model. Hence, the results of local methods can be strongly dependent on the starting model. A variety of local inversion methods exist and the computational effort associated with individual iterations of these methods can differ substantially. For example, each iteration of the Gauss–Newton

method requires the inversion of a matrix formed from the Jacobian and Hessian matrices of first and second partial derivatives of the objective function. Somewhat simpler approaches, such as the conjugate gradient method, do not require computation of the Hessian or matrix inversion. Simpler still, the downhill simplex (DHS) method is based on an intuitive geometric scheme of moving a simplex of  $M + 1$  models downhill without any actual gradient calculations. Rajan *et al.*<sup>1</sup> and Zala and Ozard<sup>8</sup> have successfully applied local methods to geoacoustic inversion.

Global inversion methods are designed to widely search the parameter space by using a random process to repeatedly perturb the model, and include the ability to move uphill in the objective function in order to escape from local minima. Global methods require only computation of the objective function, not its derivatives, and are relatively insensitive to the starting model. However, since the model perturbations are computed randomly (without gradient information), global methods are inefficient at moving downhill and can be computationally expensive. The two most widely used global inversion methods are simulated annealing (SA) and genetic algorithms (GA), which are both based on analogies with natural optimization processes (crystal growth by thermodynamic annealing and genetic evolution, respectively). SA randomly perturbs the model, and accepts all steps which lower the objective function (referred to as energy) while probabilistically accepting some uphill steps. The probability of accepting uphill steps is decreased as the inversion progresses, eventually leading to convergence to a minimum in the parameter space. GA are based on simulating the evolution of a population of models through random processes that mimic genetic crossover (recombinations of existing models) and mutation (random variations) in a manner that

<sup>a)</sup>Electronic mail: sdosso@uvic.ca

favors models with a low mismatch (or conversely, a high match, referred to as fitness). Both SA and GA have been applied to geoacoustic inversion by a number of authors.<sup>2-7,9</sup> The relative efficiency of the algorithms has been debated, but SA is certainly the simpler of two.

Whether a local or global method is best suited to a particular inverse problem depends on the topology of the multi-dimensional parameter space.<sup>7</sup> If the space includes a relatively small number of local minima, a local method initiated from a number of random starting models is more efficient than global inversion. However, for parameter spaces with a large number of minima, this approach becomes impractical, and global inversion is the method of choice. Unfortunately, for a given inverse problem, it can be difficult to assess the topology of the parameter space and determine *a priori* which approach is more applicable. However, this issue would seem to have been precluded by the recent development of hybrid inversion methods for geophysical inverse problems.

Hybrid inversion methods combine local and global inversions in an attempt to retain the advantages of each while overcoming their respective weaknesses. Hybrid methods have the potential to move effectively downhill, yet avoid becoming trapped in local minima, and lead to a solution that is independent of the starting model. Hybrid inversions typically use a local method to replace or improve some or all of the random steps of a global inversion algorithm. This can be accomplished in a variety of ways. For instance, Gerstoft<sup>7</sup> used several Gauss-Newton iterations to improve each newly generated model in the population of a GA inversion, and applied this approach to the geoacoustic inverse problem. Chunduru *et al.*<sup>11</sup> used a conjugate gradient method to improve the models obtained at various stages of a SA inversion for geo-electrical resistivity models (a number of variations of the algorithm were considered). Liu *et al.*<sup>12</sup> used a weighted combination of DHS and random steps within a SA algorithm. At early stages of the inversion, the random steps were weighted more heavily; near convergence, the DHS steps were favored. The algorithm was applied to seismic waveform inversion and residual statics corrections. Each of these hybrid methods combine, in some manner, a deterministic local step down the gradient with an independent random step. Press *et al.*<sup>13</sup> suggested a somewhat different approach based on incorporating a random component directly into the local inversion at every step in a hybrid inversion based on DHS and SA.

The hybrid geoacoustic inversion developed in this paper represents a variation of the method of Press *et al.*<sup>13</sup> This approach was chosen for a number of reasons. Foremost is the inherent simplicity of the DHS and SA methods compared to other local and global inversion methods. Second is the efficiency of individual DHS iterations, which provide a local downhill step without computing partial derivatives or solving systems of equations. Individual DHS iterations may not step down the gradient as effectively as more sophisticated local methods, but as a random component to the model perturbations is required in SA, precise downhill steps are not needed and may result in unnecessary computational expense. Finally, as DHS operates on a simplex of  $M + 1$

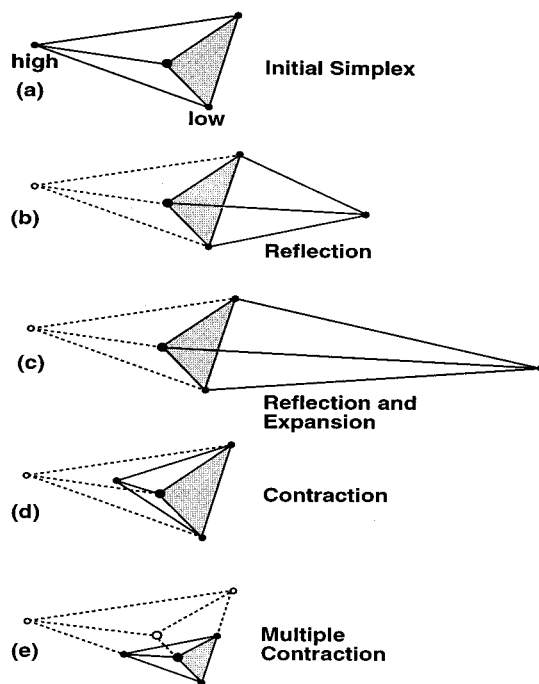


FIG. 1. Types of steps attempted by the DHS algorithm in three dimensions (after Press *et al.*, Ref. 12).

models rather than a single model, the method effectively retains a memory of good regions of the parameter space as the inversion proceeds, a feature notably lacking in SA (although present in GA).

In this paper, local and global geoacoustic inversions based on DHS and SA are developed in Secs. I and II, respectively, leading to a hybrid inversion algorithm, referred to as simplex simulated annealing, in Sec. III. For comparison, each of the inversions are applied to a benchmark testcase developed at the 1997 Matched-field Workshop.<sup>10</sup> This testcase involves nine unknown geoacoustic and geometric parameters; however, the form of the model is exactly known (i.e., a single sediment layer over a basement). In Sec. IV, the hybrid inversion is applied to a more practical and challenging benchmark testcase in which the form of the model is not known (i.e., a general sediment profile) and must be included as part of the inversion. The Appendix presents a further comparison of the three methods for a known multimodal function.

## I. LOCAL INVERSION: THE DOWNHILL SIMPLEX METHOD

The DHS method<sup>13,14</sup> is a local inversion technique based on an intuitive geometric scheme for moving downhill in parameter space. Although DHS is not necessarily the most efficient method for finding local minima, individual iterations are both efficient and simple, which are important properties for the approach to hybrid inversion developed in Sec. III. The DHS method operates on a simplex of  $M + 1$  models in an  $M$ -dimensional space [e.g., Fig. 1(a), for  $M = 3$ ]. The simplex undergoes a series of transformations in order to work its way downhill. Each model is ranked ac-

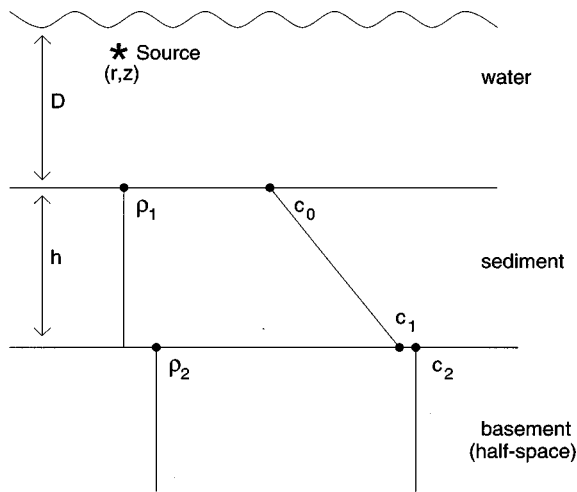


FIG. 2. Schematic diagram of the ocean environment for the first test case. The unknown parameters are described in the text.

according to its mismatch  $E$ . The algorithm initially attempts to improve the model with the highest mismatch by reflecting it through the face of the simplex [Fig. 1(b)]. If this new model has the lowest mismatch in the simplex, an extension by a factor of 2 in the same direction is attempted [Fig. 1(c)]. If the model obtained by the reflection still has the highest mismatch, the reflection is rejected and a contraction by a factor of 2 towards the lowest-mismatch model is attempted [Fig. 1(d)]. If none of these steps decrease  $E$ , then a multiple contraction by a factor of 2 in all dimensions toward the lowest-mismatch model is performed [Fig. 1(e)]. If any of these steps result in parameter values outside the given search bound, the parameters are set to the bound. The above series of steps is repeated until the difference between the highest- and lowest-mismatch values relative to their average is less than some tolerance  $\epsilon$ , i.e.,

$$\frac{E_{\text{high}} - E_{\text{low}}}{(E_{\text{high}} + E_{\text{low}})/2} < \epsilon, \quad (1)$$

or until a maximum number of iterations is reached. The DHS method moves progressively downhill, but as it has no mechanism to move uphill, it is prone to becoming trapped in local minima. Therefore, to improve the chances of finding a good approximation to the global minimum, it is necessary to initiate the DHS method from a number of starting models. The model solution with the lowest mismatch is

taken as the best estimate of the true parameter values.

The DHS inversion algorithm described above was applied to a benchmark test case developed at the 1997 Matched-field Workshop.<sup>10</sup> This test case involved determining the geoacoustic and geometric properties for the environment shown in Fig. 2. In this test case, the form of the geoacoustic model is known to consist of a single sediment layer over a semi-infinite basement. There are nine unknown parameters including the water depth,  $D$ , source range and depth,  $r$  and  $z$ , sediment thickness,  $h$ , compressional speeds at the top and bottom of the sediment layer,  $c_0$  and  $c_1$  (linear gradient assumed), compressional speed of the basement,  $c_2$ , and the densities of the sediment and basement,  $\rho_1$  and  $\rho_2$ , respectively. The true parameter values and the assumed lower and upper parameter bounds,  $m_i^-$  and  $m_i^+$ , are given in Table I. The data consist of acoustic field measurements at a frequency of 100 Hz on a vertical array of 20 hydrophones evenly spaced over the water column. The benchmark data were produced using the full-wave numerical propagation model SAFARI.<sup>15</sup> The replica fields are computed using the normal mode model ORCA,<sup>16</sup> which provides an efficient far-field solution to the (elastic) wave equation. The measure of mismatch adopted in this paper is based on the (normalized) Bartlett correlator at a single frequency or averaged incoherently over multiple frequencies:

$$E(\mathbf{m}) = 1 - \frac{1}{F} \sum_{i=1}^F \frac{|\mathbf{p}(f_i) \cdot \mathbf{p}^*(\mathbf{m}, f_i)|^2}{|\mathbf{p}(f_i)|^2 |\mathbf{p}(\mathbf{m}, f_i)|^2}, \quad (2)$$

where  $F \geq 1$  is the number of frequencies,  $\mathbf{p}$  is the acoustic field measured at the array of sensors, and  $\mathbf{p}(\mathbf{m})$  is the replica field computed for model  $\mathbf{m}$ . For this test case, the mismatch between the SAFARI- and ORCA-generated fields when  $\mathbf{m}$  was taken to be the true geoacoustic model was  $E = 1.0 \times 10^{-6}$ . This model-based mismatch places a lower limit on the useful range of inversion mismatch values (i.e., it is not meaningful to invert the data to  $E \leq 10^{-6}$ ). Figures 3 and 4 show one-dimensional (1-D) and 2-D cross sections of the parameter space for this test case. For each figure, the parameters that are not varied are held fixed at their true value. These figures clearly illustrate three features that make geoacoustic inversion a challenging problem: a large number of local minima, correlations between parameters (resulting in narrow oblique valleys in the parameter space), and a wide range of parameter sensitivities.

TABLE I. The true parameter values, lower and upper bounds ( $m_i^-$  and  $m_i^+$ ), and inversion results for the first test case. CPU indicates the computation time for the inversion. The parameter symbols are defined in the text and shown in Fig. 2.

Case	$h$ (m)	$c_0$ (m/s)	$c_1$ (m/s)	$c_2$ (m/s)	$\rho_1$ (g/cm <sup>3</sup> )	$\rho_2$ (g/cm <sup>3</sup> )	$D$ (m)	$r$ (km)	$z$ (m)	$E$	CPU (h)
True	27.08	1516	1573	1751	1.54	1.85	115.33	1.22	26.40		
$m_i^-$	10	1500	1550	1600	1.4	1.6	100	1.0	10		
$m_i^+$	50	1600	1750	1800	1.85	2.0	120	1.4	30		
DHS	27.06	1516	1574	1751	1.54	1.88	115.33	1.22	26.42	$2.6 \times 10^{-6}$	30
SA	27.50	1518	1572	1761	1.52	1.71	115.42	1.22	26.44	$1.4 \times 10^{-4}$	12
SSA	27.09	1517	1573	1751	1.54	1.81	115.35	1.22	26.42	$3.0 \times 10^{-6}$	1

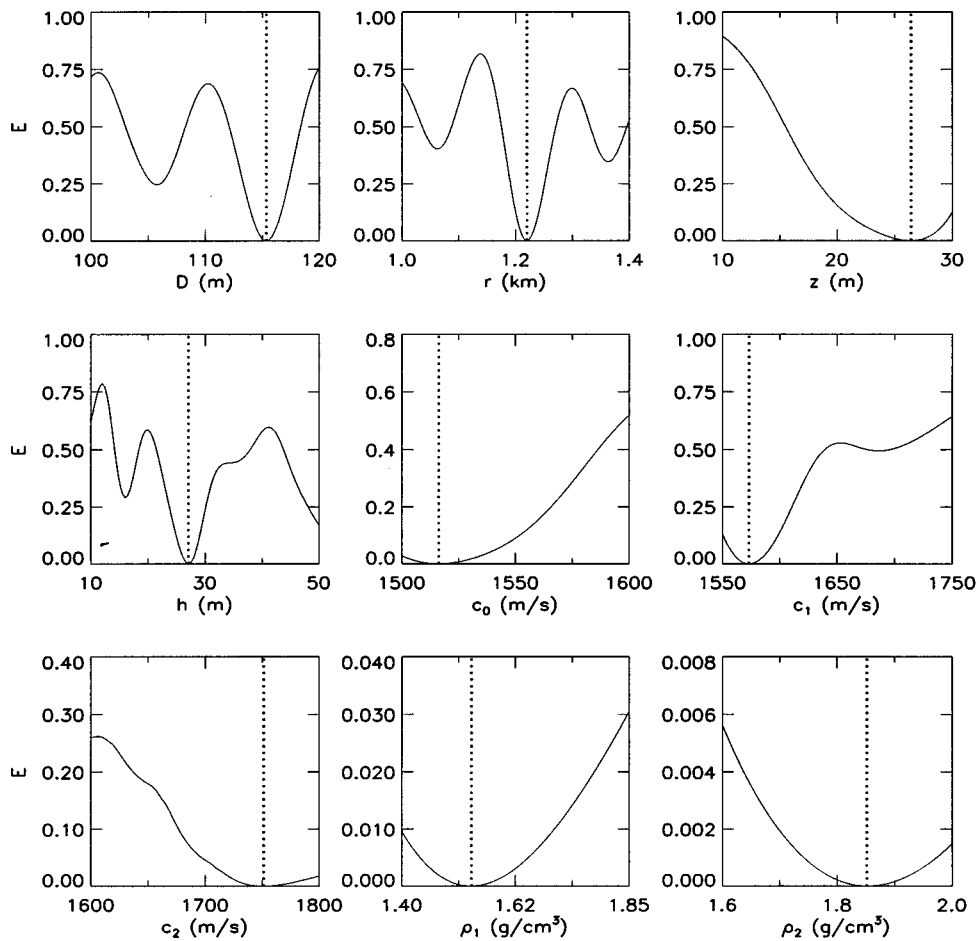


FIG. 3. 1-D cross sections of the parameter space for the first workshop test case. The dotted lines represent the true parameter values.

Figure 5 shows a compilation of the final model parameters obtained by initiating the DHS algorithm from 250 randomly chosen starting simplexes (i.e., 2500 random starting models, since there are 10 models in each simplex). The tolerance used for convergence according to Eq. (1) was  $\epsilon = 10^{-3}$  and a maximum of 1000 iterations were allowed

which lead to all but 3 of the cases running to convergence. The inversion required  $\sim 30$  h of computation time on a 200-MHz Pentium PC. The algorithm frequently obtained models with relatively low mismatch, with 40% of the solutions achieving mismatches of  $E < 10^{-2}$  and 20% obtaining  $E$

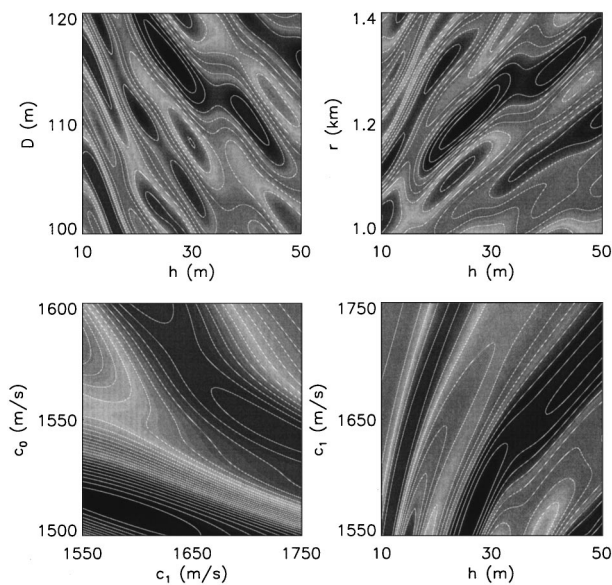


FIG. 4. 2-D cross sections of the parameter space for the first workshop test case selected to illustrate local minima and correlations between parameters. The dark areas represent low mismatches.

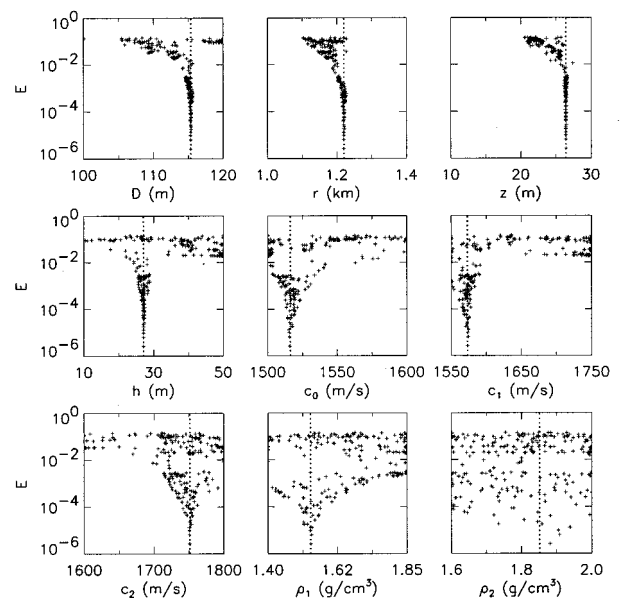


FIG. 5. Results for 250 independent DHS inversions for the first test case. The dotted lines represent the true parameter values, and the range of abscissa values indicates the parameter search interval.

$<10^{-3}$ . The plots in Fig. 5 show the mismatch as a function of model parameter values. The range of abscissa values indicate the parameter search interval and the true parameter values are given by the dotted line. The geometric parameters  $D$ ,  $h$ ,  $r$ , and  $z$  are generally well determined, with reasonable estimates obtained for  $E \leq 10^{-2}$ . The compressional speeds  $c_0$ ,  $c_1$ , and  $c_2$  are fairly well determined for  $E \leq 10^{-3}$ . However, the densities  $\rho_1$  and  $\rho_2$  are generally poorly determined: even with  $E < 10^{-4}$  good estimates for  $\rho_2$  are not always obtained. It is apparent that the acoustic fields are relatively insensitive to the densities, since the measured fields can be matched to a high precision despite significant discrepancies in these parameters. The set of model parameters which achieved the smallest mismatch ( $E = 2.6 \times 10^{-6}$ ) is given in Table I, and represents excellent estimates for all parameters. Note that this mismatch approaches the level of model-based mismatch.

## II. GLOBAL INVERSION: SIMULATED ANNEALING

Local inversion methods, such as DHS, move directly downhill, becoming trapped in local minima, and must be initiated from a number of starting models. Global inversion methods, such as SA, have the ability to avoid local minima and more effectively search the space, albeit in a somewhat inefficient manner. In this section, a fast SA algorithm is described and applied to the geoacoustic inversion benchmark testcase.<sup>9</sup>

The SA algorithm consists of a series of iterations involving random perturbations of the unknown parameters. After each iteration, a control parameter, the temperature  $T$ , is decreased slightly. Perturbations which decrease  $E$  are always accepted; perturbations which increase  $E$  are accepted conditionally, with a probability  $P$  that decreases with  $T$  according to the Boltzmann distribution

$$P(\Delta E) = \exp(-\Delta E/T). \quad (3)$$

Accepting some perturbations which increase  $E$  allows the algorithm to escape from local minima in search of a better solution. At early iterations (high  $T$ ), the algorithm searches the parameter space in an essentially random manner. As  $T$  decreases, accepting increases in  $E$  becomes increasingly improbable, and the algorithm spends more time searching regions of low  $E$ , eventually converging to a solution which should approximate the global minimum. The starting temperature, rate of reducing  $T$ , and the number and type of perturbations define the annealing schedule. Adopting an annealing schedule that is too fast (i.e., decreases  $T$  too quickly or allows too few perturbations) can lead to sub-optimal solutions. Alternatively, adopting an annealing schedule that is overly cautious wastes computation time. SA is a heuristic search technique, and there are a variety of approaches to defining the annealing schedule. An appropriate annealing schedule is generally problem specific and requires some experimentation and familiarity with the inverse problem.

In this paper, a variation of fast simulated annealing,<sup>17</sup> as implemented by Liu *et al.*<sup>12</sup> and Lane,<sup>18</sup> is employed. Fast SA is based on using a temperature-dependent Cauchy distribution to generate the parameter perturbations, and reducing the temperature in a geometric manner. The Cauchy dis-

tribution has the desirable properties of a Gaussian-like peak and Lorentzian tails which provide concentrated local sampling of the parameter space while allowing occasional large perturbations. The perturbations are implemented in the following manner. Each model parameter  $m_i$  is perturbed according to

$$m_i = m_i' + \xi \Delta_i, \quad (4)$$

where  $m_i'$  is the value of the model parameter prior to the perturbation,  $\Delta_i$  is a random variable uniformly distributed on the interval  $[-\Delta m_i, \Delta m_i]$ , where  $\Delta m_i = m_i^+ - m_i^-$ , and the quantity  $\xi$  is a Cauchy-distributed random variable computed as

$$\xi = [T_j/T_0]^{1/2} \tan[\pi(\eta - 1/2)]. \quad (5)$$

In Eq. (5),  $\eta$  is a uniform random variable on  $[0, 1]$ , and  $T_j$  is the temperature at the  $j$ th step. The size of the perturbations decreases with the square root of the temperature (this is somewhat more conservative than the approach in Ref. 17 in which  $\xi$  depended directly on  $T$ ). Thus at high temperatures, large perturbations search the entire parameter space, while at low temperatures a more local search is performed. However, because of the long tails of the Cauchy distribution, large perturbations are possible at any temperature. Each parameter is perturbed individually a predetermined number of times according to Eq. (4) at each temperature step. After each perturbation, the change in the objective function is examined to determine if the new parameter is accepted or rejected. After the set of perturbations is complete, the temperature is reduced according to

$$T_j = \beta^j T_0, \quad (6)$$

where  $\beta$  is a constant less than one. An appropriate value for the starting temperature  $T_0$  can be determined by requiring that at least 90% of all perturbations are accepted initially. Appropriate values for  $\beta$  and the number of perturbations per temperature step depend on the difficulty of the inversion and require some experimentation. Once the desired number of iterations is completed, a process known as ‘‘quenching’’ is carried out to move to the bottom of the closest minimum. In quenching, the size of the model perturbations are held fixed and the temperature is set to zero for a preset number of iterations (this removes the possibility of taking uphill steps).

The SA inversion algorithm described above was applied to the same geoacoustic testcase considered in Sec. I (Fig. 2). The annealing schedule was initiated at a temperature of  $T_0 = 0.3$  and involved 2500 iterations (temperature steps) with 5 perturbations of each parameter at each iteration, and a temperature reduction factor of  $\beta = 0.995$ . This inversion required  $\sim 12$  h of computation time. The results of the inversion are displayed in Fig. 6 and summarized in Table I. Figure 6(a) shows that the mismatch  $E$  decreases steadily (although not monotonically) with temperature to a small final value of  $E = 1.4 \times 10^{-4}$ . Figure 6(b)–(j) shows the convergence of the geoacoustic parameters. Good estimates of the true values are found for all parameters, with the exception of the basement density  $\rho_2$ . However, it is interesting to note from Table I that for this example, the local

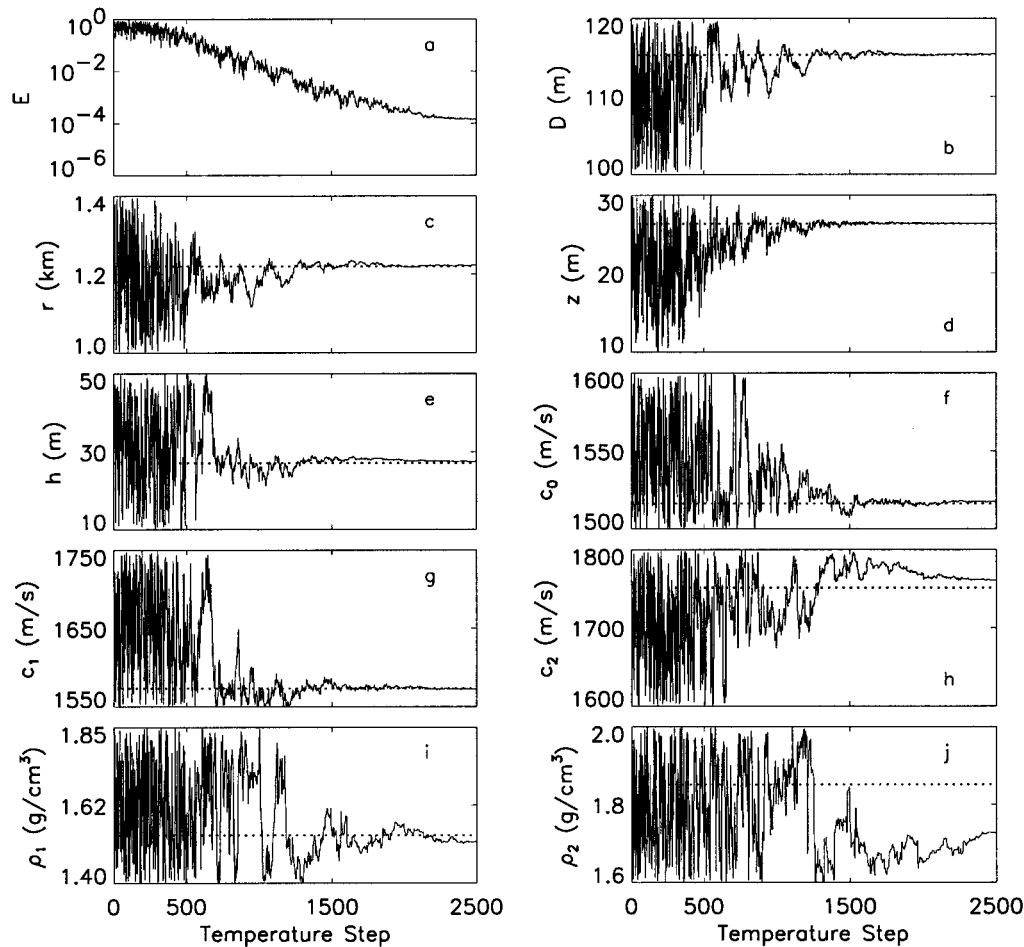


FIG. 6. Convergence of the mismatch  $E$  and the model parameters for the SA inversion of the first test case. The dotted line represents the true parameter value, and the range of ordinate values indicates the parameter search interval.

DHS method produced a substantially better result than SA in approximately twice the computation time.

### III. HYBRID INVERSION: SIMPLEX SIMULATED ANNEALING

SA has been successfully applied to geoacoustic inversions by a number of authors (e.g., Refs. 2–4, 6, 9). However, the method has several shortcomings. For instance, it is relatively inefficient near convergence and in cases where correlated parameters produce narrow valleys in the parameter space that are not aligned with the parameter axes. To address the latter problem, Collins and Fishman<sup>6</sup> suggested using the covariance matrix of the objective function gradient (sampled over the parameter space) to rotate the parameter space, searching for the optimal solution in the new space, then rotating back to the original parameters (this approach applies to parameter spaces containing a small number of local minima with prominent features oriented in one direction<sup>7</sup>). Another shortcoming of SA is its lack of any form of memory: A particularly good model may be discarded at an early stage and never revisited. To address this, Lindsay and Chapman<sup>4</sup> devised an adaptive SA algorithm that accumulated information about the topography of the parameter space as the search progressed, and adaptively guided the search to preferred regions of the parameter

space. In this section, a hybrid inversion based on SA and DHS is developed which overcomes each of these shortcomings in a natural manner.

The hybrid inversion, referred to as simplex simulated annealing (SSA), incorporates the local DHS method described in Sec. I into the global SA search described in Sec. II. In particular, unlike standard SA, the SSA inversion operates on a simplex of models rather than on a single model, and instead of employing purely random model perturbations, DHS steps with a random component are applied to perturb the model. The procedure developed here to introduce the random component into the DHS steps is somewhat subtle, but has proven to be highly effective. The DHS steps are not computed directly from the current simplex of models, but rather from a secondary simplex which is formed by applying random perturbations to all the model parameters and mismatches associated with the current simplex. The perturbations to the model parameters of the current simplex are computed using the temperature-dependent Cauchy distribution, given by Eqs. (4) and (5). The perturbation to the mismatch associated with each model in the simplex is computed according to

$$E = E' + \xi \bar{E}, \quad (7)$$

where  $\xi$  is given by Eq. (5),  $E'$  is the current mismatch

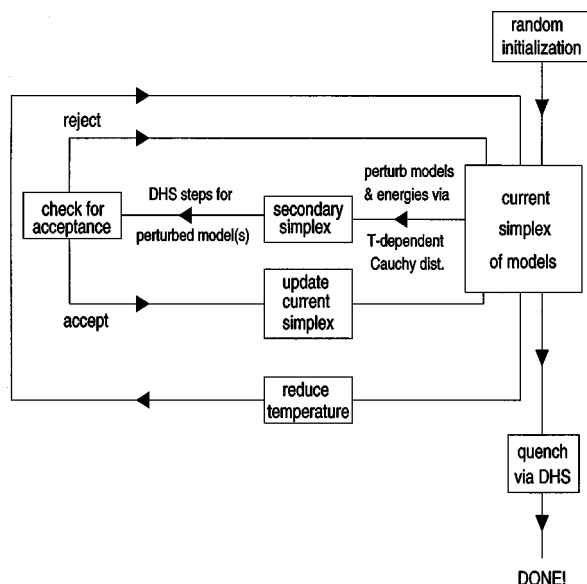


FIG. 7. Simplified block diagram illustrating the SSA algorithm.

corresponding to that model, and  $\bar{E}$  represents the mean mismatch of the current simplex. The random component introduced by computing the DHS steps from a secondary (perturbed) simplex means that the steps are not always downhill, but may be uphill as well. Each proposed DHS step is evaluated for acceptance using the standard probabilistic criterion of SA, providing a mechanism for accepting uphill steps and escaping from local minima. The accepted steps are used to update the current simplex. After each DHS step (accepted or rejected) a new secondary simplex is computed. As in SA, the form of the parameter search of the SSA algorithm is controlled by the temperature. At high temperatures, the random component of the DHS steps dominates and the probability of accepting uphill steps is high, resulting in an essentially random search. At low temperatures, the random component and probability of accepting uphill steps is small, and the search approximates the local DHS method, which moves efficiently downhill. At intermediate temperatures, the algorithm makes a smooth transition from a random search to the DHS method. In practice, once the temperature has been reduced to a sufficiently low value, it is often effective to quench the inversion by setting the temperature to zero and using the DHS method to move directly to the bottom of the closest minimum. A block diagram illustrating the basic SSA algorithm is given in Fig. 7.

One final note on the SSA algorithm has to do with efficiency. The multiple-contraction step of the DHS method [e.g., Fig. 1(e)] is the most time consuming since it requires the evaluation of the mismatch for all but one model of the simplex. At early (random-search) stages of the inversion, the additional computational expense of performing multiple contractions is generally wasted. Therefore, the efficiency of the SSA algorithm can be improved by introducing a simple procedure which allows only some fraction of the multiple contractions to be performed. At high temperatures, this fraction can be small (or zero); the fraction is increased as the temperature decreases, to a final value of unity near convergence.

The hybrid SSA approach addresses the shortcomings associated with standard implementations of SA. Since DHS steps adjust all parameters at once according to the local gradient, the algorithm moves effectively down narrow oblique valleys in parameter space. Also, near convergence (low  $T$ ), the algorithm more and more closely approximates the DHS algorithm and moves efficiently downhill. An additional feature of SSA is that the current best model is always retained in the simplex, effectively providing the method with a form memory. As a result of these factors, we have found that the SSA algorithm can achieve substantially better results with a much faster annealing schedule than SA.

The SSA algorithm was applied to invert data from the testcase described in Fig. 2. The annealing schedule was initiated at a temperature of  $T_0=0.3$  and involved 800 temperature steps with 3 DHS steps at each iteration and a temperature reduction factor of  $\beta=0.985$ . The inversion required  $\sim 1$  h of computation time. The results of the inversion are shown in Fig. 8 and summarized in Table I. Figure 8(a) shows the mismatch  $E$  of all models in the simplex throughout the inversion. The final mismatch for the inversion is exceedingly low ( $E=3.0\times 10^{-6}$ ) and approaches the level of model-based mismatch. Figure 8(b)–(j) shows the convergence of the geoacoustic parameters for all models in the simplex. The SSA inversion produced excellent estimates of the true values for all of the unknown parameters, including the basement density. To assess the reliability of the SSA algorithm, the inversion was repeated from ten different random initializations. In each case the results were similar to those described here. In fact, one case obtained a mismatch of  $E<10^{-6}$ , which is below the level of model-based mismatch. The results in Table I indicate that the SSA inversion is substantially more efficient than either DHS and SA. In the Appendix, a further comparison of SSA to DHS and SA is considered using a known multimodal function.

#### IV. GEOACOUSTIC PROFILE INVERSION

The testcase considered in developing the DHS, SA, and SSA inversions in Secs. I–III assumed that the form of the geoacoustic model was exactly known (i.e., a single sediment layer over a semi-infinite basement). In many practical cases, however, the form of the model is not well known, and must be included as part of the inversion. This section considers applying the SSA algorithm to a second testcase from the Matched-field Workshop,<sup>10</sup> which consists of inverting data generated for an unknown (large) number of sediment layers, representing a general geoacoustic profile. In cases where the form of the model is not known, it is well advised to seek as simple a solution as possible (i.e., a minimum-structure model). There are two general approaches to accomplish this. One approach is to control the amount of structure by the parameterization of the model. A typical application of this approach begins by inverting for a purposely underparameterized model, and successively repeats the inversion as the number of model parameters is increased until a point is reached where further increases in parameters lead to negligible decreases in mismatch. Beyond this point, including more parameters allows more model structure; however, this structure is clearly not required by



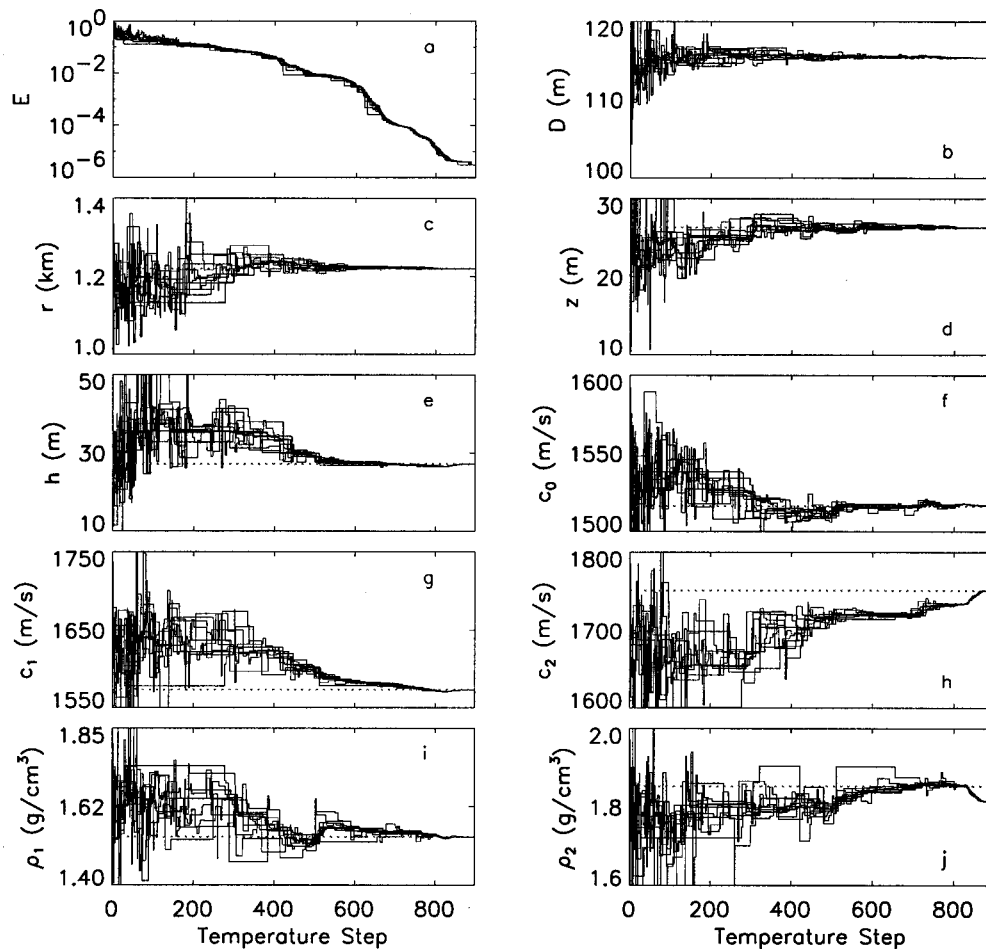


FIG. 8. Convergence of the mismatch  $E$  and the model parameters for the SSA inversion of the first test case (all models in the simplex are shown). The dotted line represents the true parameter value, and the range of ordinate values indicates the parameter search interval.

the data. A second approach is to overparameterize the model, but include an explicit penalty for model structure along with the data mismatch in the objective function being minimized (i.e., regularize the inversion by minimizing model structure). A well-developed theory for this approach exists for local inverse methods;<sup>19</sup> however, applications to global inversion to date are somewhat ad hoc.<sup>20</sup>

Here, an underparameterized approach is applied to geoacoustic-profile inversion. The test case consists of inverting acoustic data at two frequencies (40 and 120 Hz) for compressional-speed and density profiles; geometric parameters are considered known and the properties of the water column and sensor array are identical to the previous test case. Figure 9(a) shows mismatches obtained via SSA inversion when the geoacoustic model consisted of one to nine layers. The parameters included in the inversion were the compressional speed and density of each layer, and the thickness of each layer except the basement. Hence, for a model with  $N$  layers, a total of  $3N-1$  parameters are included in the inversion. The mismatch in Fig. 9(a) decreases significantly as the number of layers increases from one to five; beyond five layers the mismatch increases slightly. This slight increase is likely due to the fact that as the number of

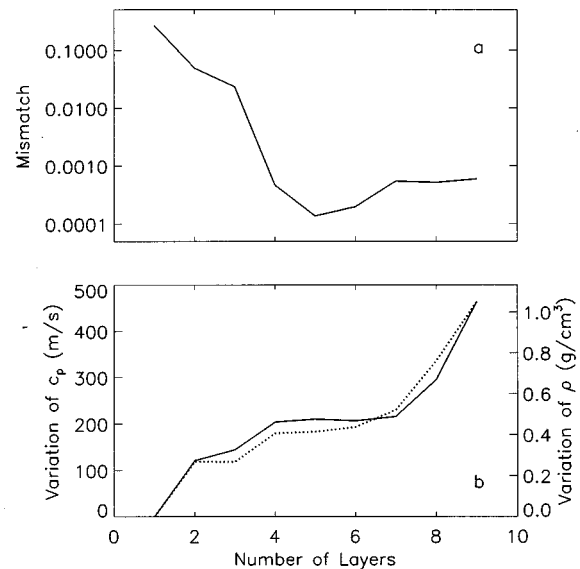


FIG. 9. The mismatch, for the second test case, as a function of the number of layers included in the inversion is shown in (a). The  $L_1$  norm of the variation of the compressional speed  $c_p$  and density  $\rho$  is indicated by the solid and dotted lines, respectively, in (b).

model layers increases beyond the number that are well resolved by the data, the number of local minima in the parameter space increases greatly, and the SSA inversion algorithm has increasing difficulty finding a good approximation to the global minimum. Adding more layers allows the amount of model structure to increase. The structure of a geoacoustic model profile can be quantified by the  $L_1$  norm of the variation:

$$V = \sum_{i=1}^{N-1} |q_{i+1} - q_i|, \quad (8)$$

where  $q$  can represent either the compressional speed  $c_p$  or density  $\rho$ . Figure 9(b) shows the variation of the compressional speed and density as a function of the number of layers included in the inversion. The variation increases as the number of layers is increased from one to four. For four to seven layers the variation remains approximately constant; for more than seven layers the variation again increases with the number of layers. This increase in variation is illustrated in Fig. 10 which shows examples of compressional-speed and density profiles obtained by the SSA inversion for models with two, five, and nine layers. In each case the dotted line indicates the inversion result, and the solid line indicates the true profile. The two-layer model [Fig. 10(a), (b)] represents a reasonable, if somewhat simplified, approximation to the true profiles. A good approximation to the true profiles is obtained from the inversion that included five layers in the model [Fig. 10(c), (d)]. The nine-layer case produced a reasonable estimate of the compressional-speed profile [Fig. 10(e)], but produced a relatively poor estimate of the density profile [Fig. 10(f)]. Also, the nine-layer case clearly contains extra structure (compared to the five-layer case), in the form of thin low-speed and/or high- and low-density layers.

According to Figs. 9 and 10, a five-layer model would seem to represent the best trade-off between a low mismatch and a small amount of structure, and hence is the preferred

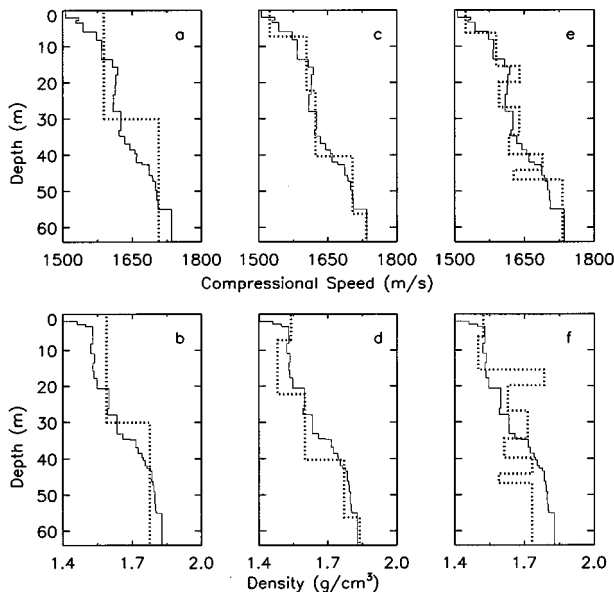


FIG. 10. Compressional-speed and density profiles, for the second testcase, obtained by SSA inversion. The inversion results for two, five and nine layers are shown by the dotted lines in (a, b), (c, d), and (e, f), respectively. The solid lines indicates the true compressional-speed and density profiles.

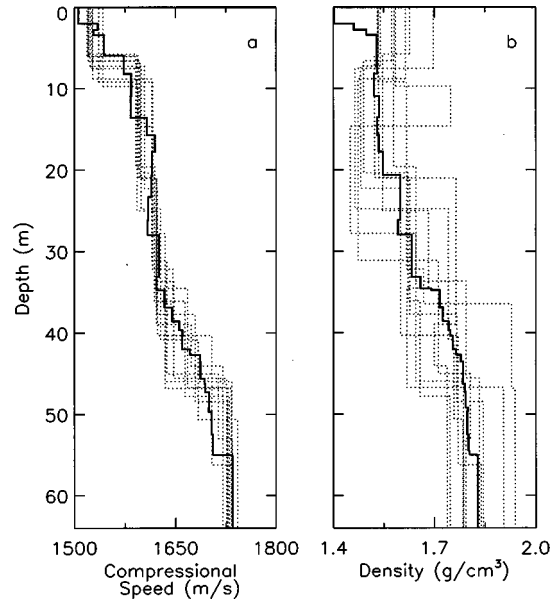


FIG. 11. Results of ten SSA inversions, for the second testcase, showing (a) compressional-speed profiles, and (b) density profiles. The dotted lines are the inversions results, and solid line indicates true profile.

interpretation. To investigate the range of possible five-layer solutions, the SSA inversion was carried out ten times initiated from different starting points. Figure 11 shows the results of these inversions. The inversions produced consistently good estimates for the compressional-speed profile [Fig. 11(a)], with a tight envelope closely following the true profile. Since the acoustic-field data are relatively insensitive to the sediment density, the density profile is not as well constrained [Fig. 11(b)]; however, the underlying trend of density increasing with depth is clearly evident.

## V. SUMMARY

In this paper, local, global, and hybrid inversion algorithms were developed and applied to the problem of determining geoacoustic properties by minimizing the mismatch between measured and modeled acoustic fields. Local inversion methods are sensitive to local gradients in the mismatch and move effectively downhill, but typically become trapped in local minima. Hence, a large number of starting points may be required for nonlinear problems. Global inversion methods use directed random processes to widely search the parameter space and include the ability to move uphill to escape from local minima. However, as gradient information is not employed, the search can be relatively inefficient. Hybrid inversion methods combine local and global approaches in an attempt to retain the advantages of each and produce a more efficient and effective inversion algorithm. Here, local and global inversion algorithms based on the downhill simplex method and simulated annealing, respectively, were considered individually and combined to produce a hybrid simplex simulated annealing algorithm for geoacoustic inversion. DHS is based on an intuitive geometric scheme for moving downhill in parameter space without computing partial derivatives or solving systems of equations. SA uses an analogy to the physical process of annealing to search the

parameter space. The hybrid SSA algorithm consists of local DHS steps incorporated into the SA global search. DHS and SA were chosen for the hybrid inversion algorithm because of their inherent simplicity and efficiency, and because the resulting algorithm retains a memory of good regions of the parameter space.

All three algorithms were evaluated using a benchmark testcase developed for the 1997 Matched-field Workshop. The testcase involved inverting noise-free acoustic fields for nine unknown geoacoustic and geometric parameters in a case where the form of the model was exactly known (i.e., a single sediment layer over a semi-infinite basement). The multi-start DHS algorithm produced an excellent result, but was computationally expensive, requiring approximately 30 h of computation time. SA produced a somewhat less precise result in about one-third the time. The hybrid SSA inversion proved to be the most efficient by far, producing an excellent result in 1 h of computation time.

SSA was also applied to a more practical testcase which consisted of inverting for an unknown (large) number of sediment layers representing a general geoacoustic profile. In this case the form of the model was not known, and had to be included as part of the inversion. To this end, an underparameterized approach was employed to determine the model which achieved the best trade-off between low mismatch and a simple structure. The inversion was successively repeated as increasing numbers of layers were included in the geoacoustic model. The relationship between the mismatch and the amount of model structure (measured by the  $L_1$  norm of the model variation) as a function of the number of layers was then examined to determine the preferred model representation. For this testcase, SSA produced a good, if somewhat simplified, approximation to the true geoacoustic model. Finally, the SSA algorithm developed in this paper has also been applied to other geoacoustic testcases from the workshop (not shown here), to a known multimodal function (see the Appendix), and to an unrelated inverse problem,<sup>21</sup> with the same result as obtained here: i.e., the hybrid inversion was much more efficient than either local or global methods.

## ACKNOWLEDGMENTS

This work was supported by the National Science and Engineering Research Council (NSERC) and the Department of National Defense (DND) through an NSERC research grant and the NSERC/DND Ocean Acoustics Chair Program at the University of Victoria.

## APPENDIX: MINIMIZATION OF A MULTIMODAL FUNCTION

This Appendix compares the local, global, and hybrids methods developed in this paper for a known multimodal function. The function considered is a version of that used by Musil<sup>22</sup> extended to higher dimensions:

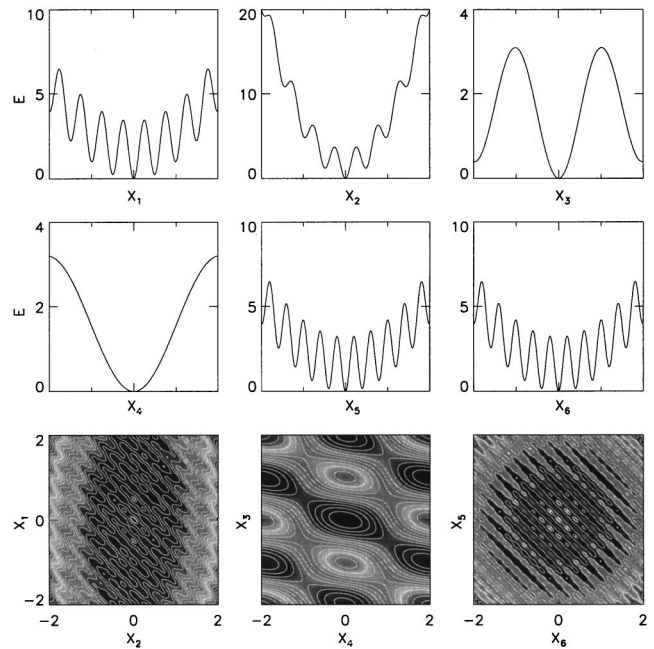


FIG. A1. 1-D and 2-D cross sections of function (A1). In each case, the parameters that are not varied are held at their value at the global minimum ( $x_i=0$ ).

$$\begin{aligned}
 E(x_1, x_2, x_3, x_4, x_5, x_6) &= 4.8 + x_1^2 + 5x_2^2 + 0.1x_3^2 + 0.05x_4^2 + x_5^2 + x_6^2 \\
 &\quad - 0.3 \cos 4\pi(x_1 - x_2) - 1.4 \cos 4\pi(x_1 + x_2) \\
 &\quad - 0.5 \cos 10\pi(0.05x_4 - 0.1x_3) \\
 &\quad - 1.0 \cos 10\pi(0.05x_4 + 0.1x_3) - 0.25 \cos 5\pi(x_5 - x_6) \\
 &\quad - 1.35 \cos 5\pi(x_5 + x_6). \tag{A1}
 \end{aligned}$$

Function (A1) has three important features: a large number of local minima, correlations between some parameters, and

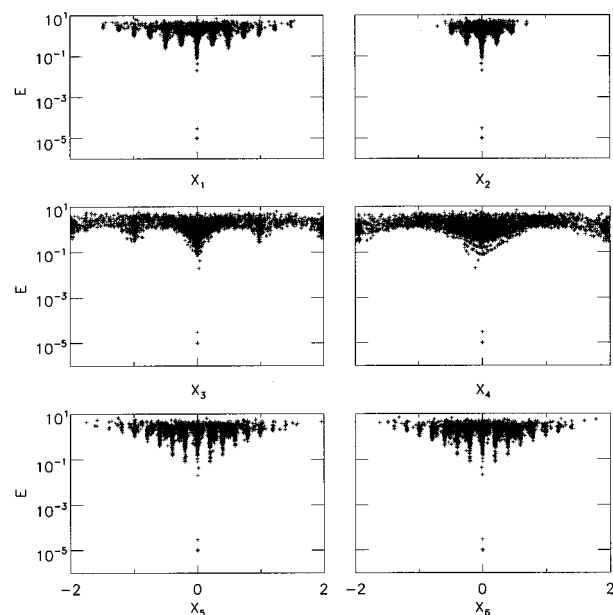


FIG. A2. The results of 5000 independent DHS runs. The point at  $E = 10^{-5}$  represents all results that achieved  $E \leq 10^{-5}$  (48 results).

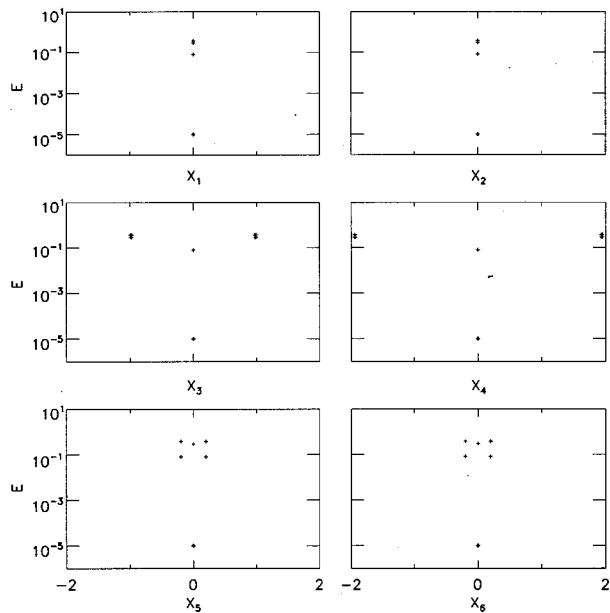


FIG. A3. The results of 100 independent SA runs. The point at  $E=10^{-5}$  represents all results that achieved  $E \leq 10^{-5}$  (44 results).

a range of parameter sensitivities. Figure A1 shows the function in one and two dimensions. The global minima of the function is  $E(x_i)=0$  which occurs at the origin (i.e.,  $x_i=0$ ,  $i=1,\dots,6$ ). Two criteria are adopted to define a successful inversion. The first criterion is that the model parameters closely approximate their values at the global minima (i.e.,  $x_i \leq 10^{-3}$ ,  $i=1,\dots,6$ ). The second criterion is that a low function value is obtained (i.e.,  $E \leq 10^{-5}$ ).

As DHS is a local method that frequently becomes trapped in local minima, it must be started from a large number of points for a nonlinear problem. Figure A2 shows the results of the 5000 independent runs of DHS with 8% of the parameter values within  $x_i \leq 10^{-3}$  and 1% of the solutions

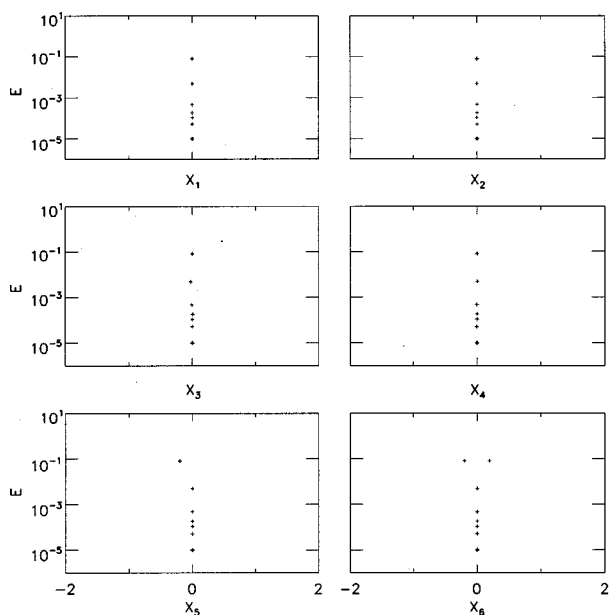


FIG. A4. The results of 100 independent SSA runs. The point at  $E=10^{-5}$  represents all results that achieved  $E \leq 10^{-5}$  (93 results).

achieving  $E \leq 10^{-5}$ . SA and SSA avoid local minima; however, to assess their effectiveness for this example, they were each initiated from 100 random starting points. The annealing schedule for SA consisted of 5000 temperature steps, with 10 perturbations per temperature step, and a temperature reduction factor of  $\beta=0.995$ . Figure A3 shows the results of the 100 SA inversions with 79% of the parameter values within  $x_i \leq 10^{-3}$  and 44% of the solutions achieving  $E \leq 10^{-5}$ . A similar annealing schedule was employed for SSA except that the inversion was quenched after 4000 temperature steps (even with similar annealing schedules, SSA required far fewer function evaluations since it perturbs all parameters at once). Figure A4 shows the results of the 100 different SSA inversions with 98% of the parameter values within  $x_i \leq 10^{-3}$  and 93% of the solutions achieving  $E \leq 10^{-5}$ .

The above results indicate that all three methods are capable of good results for a multimodal function. The DHS method frequently became trapped in local minima but did obtain some good results. SA is more effective (although not foolproof) at obtaining the global minimum, but was not particularly effective at moving downhill within this minimum. In this example, SSA achieved solutions deep in the global minimum virtually every time. Of the three methods, SSA also required an order of magnitude fewer function evaluations.

<sup>1</sup>S. D. Rajan, J. F. Lynch, and G. V. Frisk, "Perturbative inversion methods for obtaining bottom geoacoustic parameters in shallow water," *J. Acoust. Soc. Am.* **82**, 998–1017 (1987).

<sup>2</sup>M. D. Collins, W. A. Kuperman, and H. Schmidt, "Nonlinear inversion for ocean-bottom properties," *J. Acoust. Soc. Am.* **92**, 2770–2783 (1992).

<sup>3</sup>S. E. Dosso, M. L. Jeremy, J. M. Ozard, and N. R. Chapman, "Estimation of ocean-bottom properties by matched-field inversion of acoustic field data," *IEEE J. Ocean Eng.* **18**, 232–239 (1993).

<sup>4</sup>C. E. Lindsay and N. R. Chapman, "Matched field inversion for geoacoustic model parameters using adaptive simulated annealing," *IEEE J. Ocean Eng.* **18**, 224–231 (1993).

<sup>5</sup>P. Gerstoft, "Inversion of seismoacoustic data using genetic algorithms and a posteriori probability distributions," *J. Acoust. Soc. Am.* **95**, 770–782 (1994).

<sup>6</sup>M. D. Collins and L. Fishman, "Efficient navigation of parameter landscapes," *J. Acoust. Soc. Am.* **98**, 1637–1644 (1995).

<sup>7</sup>P. Gerstoft, "Inversion of acoustic data using a combination of genetic algorithms and the Gauss–Newton approach," *J. Acoust. Soc. Am.* **97**, 2181–2190 (1995).

<sup>8</sup>C. A. Zala and J. M. Ozard, "Estimation of geoacoustic parameters from narrow band data using a search-optimization technique," *J. Comput. Acoust.* **6**, 223–244 (1998).

<sup>9</sup>M. R. Fallat and S. E. Dosso, "Geoacoustic inversion for the Workshop97 benchmark test cases using simulated annealing," *J. Comput. Acoust.* **6**, 29–44 (1998).

<sup>10</sup>N. R. Chapman, A. Tolstoy, and G. H. Brooke, "Workshop97: Benchmarking for geoacoustic inversion in shallow water," *J. Comput. Acoust.* **6**, 1–28 (1998).

<sup>11</sup>R. K. Chunduru, M. K. Sen, and P. L. Stoffa, "Hybrid optimization methods for geo-physical inversion," *Geophysics* **62**, 1196–1207 (1997).

<sup>12</sup>P. Liu, S. Hartzell, and W. Stephenson, "Non-linear multiparameter inversion using a hybrid global search algorithm: Applications in reflection seismology," *Geophys. J. Int.* **122**, 991–1000 (1995).

<sup>13</sup>W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes in FORTRAN*, 2nd ed. (Cambridge University Press, Cambridge, 1992), Chap. 10.

<sup>14</sup>J. A. Nelder and R. Mead, "A simplex method for function minimization," *Comput. J.* **7**, 308–313 (1965).

<sup>15</sup>H. Schmidt, "SAFARI: Seismo-acoustic fast field algorithm for range-

- independent environments,” Rep. SR-113 SACLANT ASW Research Center, La Spezia, Italy (1988).
- <sup>16</sup>E. K. Westwood, C. T. Tindle, and N. R. Chapman, “A normal model for acousto-elastic ocean environments,” *J. Acoust. Soc. Am.* **100**, 3631–3645 (1996).
- <sup>17</sup>H. Szu and R. Hartley, “Fast simulated annealing,” *Phys. Lett. A* **122**, 157–162 (1987).
- <sup>18</sup>F. Lane, “Estimation of the kinematic rupture parameters from historical seismograms: An application of simulated annealing to a nonlinear optimization problem,” Ph.D. thesis, Colorado School of Mines, Golden, Colorado (1992).
- <sup>19</sup>S. C. Constable, R. L. Parker, and C. G. Constable, “Occam’s inversion: A practical algorithm for generating smooth models from electromagnetic sounding data,” *Geophysics* **52**, 289–300 (1987).
- <sup>20</sup>S. E. Dosso, J. M. Ozard, and J. A. Fawcett, “Inversion of acoustic field data for bathymetry and bottom sound speed via simulated annealing,” in *Acoustic Signal Processing for Ocean Exploration*, edited by J. M. F. Moura and I. M. G. Lourtie (Kluwer Academic, The Netherlands, 1993), pp. 51–56.
- <sup>21</sup>S. E. Dosso and B. J. Sotirin, “Optimal array element localization,” *J. Acoust. Soc. Am.* (submitted).
- <sup>22</sup>M. Musil, “Range-dependent matched-field inversion using the downhill simplex algorithm,” M. Sc. thesis, University of Victoria, Victoria, B. C., Canada (1998).

# Near-caustic behavior in a 270-km acoustical experiment

Jeffrey A. Simmen and Stanley M. Flatté

*Physics Department and Institute of Tectonics, University of California at Santa Cruz, Santa Cruz, California 95064*

Harry A. DeFerrari, Hien Nguyen, and Neil J. Williams

*Rosenstiel School of Marine and Atmospheric Science, University of Miami, Miami, Florida 33149*

(Received 29 June 1998; accepted for publication 12 February 1999)

Data from a pulse-propagation experiment are analyzed to quantify fluctuations in the transmitted acoustic signal. In the experiment a bottom-moored, broadband source having a 460-Hz center frequency transmitted 273 km away to a bottom-moored, vertical hydrophone array spanning 180 m of the water column. Pulses were transmitted every 12 min for a total duration of nearly 100 h. Because the minimal hydrophone separation at the receiving array was less than 2 m, vertical behavior on small scales could be observed. Furthermore, the vertical position of the receiving array enabled an examination of the near-caustic portion of wavefronts, and in some cases, the caustics themselves. Unfortunately, because instrument motion measurements were unsuccessful, *relative* rather than *absolute* signal fluctuations were the principal observable in this analysis. © 1999 Acoustical Society of America. [S0001-4966(99)01706-3]

PACS numbers: 43.30.Re, 43.30.Zk [DLB]

## INTRODUCTION

Although ocean acoustical experiments<sup>1-3</sup> have satisfactorily tested proposed *inverse* methods<sup>4,5</sup> to determine mesoscale soundspeed or current fields, the *forward* problem of understanding the effect of the dynamical ocean on acoustic propagation still warrants attention. Of particular interest, the detailed behavior that an acoustic wavefront (or its associated timefront) exhibits as it travels through the ocean medium, especially over long ranges, is still not completely understood.

For two-dimensional sound propagation in an ocean whose soundspeed is dependent only on depth, the initial circular wavefront resulting from an impulse at a point source refracts and folds into a sawtooth shape as it propagates along the ocean sound channel.<sup>5-8</sup> The dependence of this folded geometry on propagation distance and soundspeed profile is well understood. In the real ocean, however, in addition to this average sound channel there are other range, depth, and time dependent soundspeed structures, arising from ocean fronts, mesoscale eddies, tides, internal waves, and fine structure processes. These further fold and deform, usually on smaller spatial scales, the otherwise smooth segments that comprise the already folded wavefront. Because of the temporal dependence of the ocean processes, wavefronts propagated at different times, but out to the same range, will exhibit deformations that depend not only on position along the wavefront but also on time. The ultimate goal is to understand the nature and magnitude of these wavefront deformations or fluctuations, as a function of acoustic frequency and propagation range. Theoretical efforts, often relying on numerical simulation, have addressed various aspects of this problem, such as the effect of internal waves on wavefront fluctuations,<sup>9-12</sup> and the relation between mode coupling<sup>13</sup> or chaotic rays<sup>14-16</sup> and these fluctuations.

Early pulse propagation experiments<sup>6,17,18-21</sup> showed

that arrivals of separate wavefront folds were distinguishable if the width of the acoustic pulse was sufficiently small. Although in these experiments wavefront arrival time was usually identified over several days of pulse transmissions, with few exceptions was it ever studied as a function of depth, owing to limited hydrophone coverage in the vertical. With a vertical continuum of hydrophones one could, in principal, measure the global variability of the wavefront (or timefront) geometry, distinguishing fluctuations that are extremely local from those that are more global along the wavefront. There have been only a few experimental efforts<sup>8,22-24</sup> that, by monitoring fluctuations as a function of depth, have addressed the global behavior.

Recently, Duda *et al.*<sup>8</sup> reported on the fluctuations of a wavefront that propagated a distance of 1000 km. In that experiment the receiver consisted of 50 hydrophones spanning 3 km so that fluctuations in the vertical with scales greater than 60 m were discernible. The fastest pulse transmission rate was 6 pulses every hour, so that the temporal resolution was a fraction of an hour. The dominant fluctuation of wavefront arrival time that was observed, having a variance of about 40 ms<sup>2</sup>, was termed *broadband* because it was uncorrelated over the sampling scales of 10 min and 60 m in the vertical. Although it was speculated that this broadband fluctuation might reflect folding or wiggling of the wavefront on small scales, the character of this fluctuation was never determined because of the lack of information between receiver hydrophones.

The experiment described here provided for an additional examination of the nature of arrival time fluctuations, in this case for a wavefront that propagated out to a shorter range of 273 km. Because the minimal hydrophone separation at the receiving array was only 1.64 m, vertical behavior on smaller scales could be observed. Unfortunately, because instrument motion measurements were unsuccessful, and the contributions to arrival time fluctuations of these motions

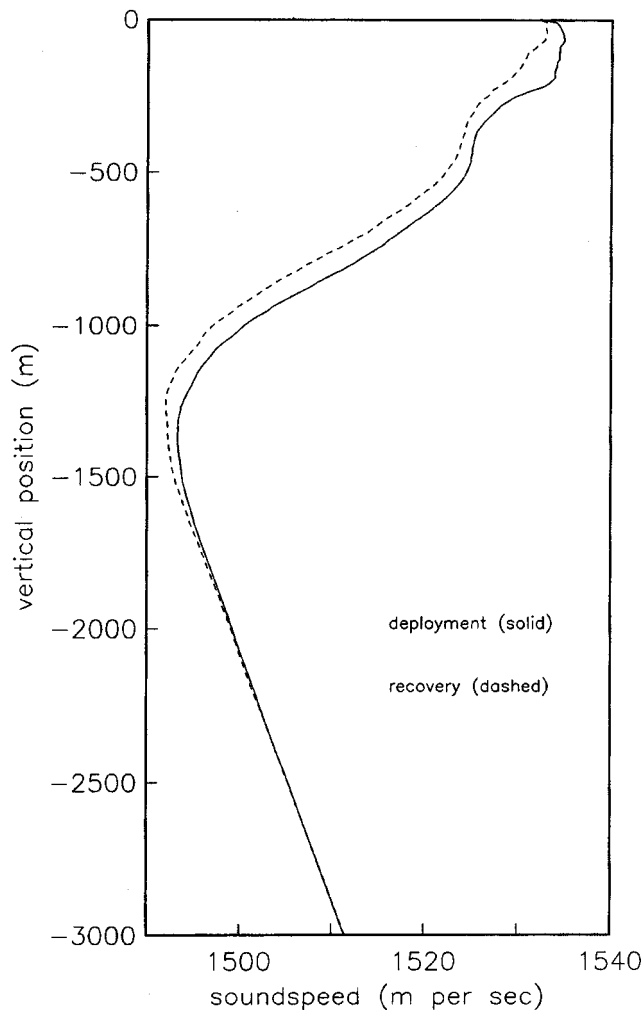


FIG. 1. Soundspeed profiles. The solid curve is the reference soundspeed profile  $C_0(z)$  constructed by averaging profiles taken during the instrument deployment cruise. The dashed curve is the average of profiles taken during the recovery cruise.

were not negligible, fluctuations in *relative* arrival time, absent of the major effects of instrument motion, became the principal observable in this analysis.

### I. DESCRIPTION OF EXPERIMENT

A vertical array of 64 hydrophones was deployed at  $27^\circ 35.42' N$ ,  $75^\circ 15.38' W$ , a distance of 272.3 km to the northwest of a broadband source deployed at  $25^\circ 24.17' N$ ,  $74^\circ 1.46' W$ . The source and receiver were both bottom moored in water having an approximate depth of 4.8 km. The source's depth (estimated at deployment near 1050 m) and the depth of the receiver's top hydrophone (estimated *a posteriori* to wander near 800–900 m) were not known accurately both at deployment, due to a malfunctioning precision depth recorder, and afterward, because instrument motion measurements were unsuccessful.

Eight conductivity-temperature-depth (CTD) casts were made along a line from the source to the receiver during the instrument deployment cruise. Similarly, on recovery of the moorings (approximately three weeks after deployment) there were an additional four CTD casts, showing a slightly

colder upper ocean. A range independent reference sound-speed profile,  $C_0(z)$ , was constructed by averaging initial CTD casts (Fig. 1). The minimal sound speed occurs at 1300 m, noticeably deeper than the depth of the deployed source and receiver hydrophones.

The receiving array, whose geometry is presented in Fig. 2, consisted of four groups of 16 hydrophones that together spanned a length of 180 m. A large, very buoyant, syntactic foam buoy at the top held the array approximately vertical; however, up-down and horizontal motion due to local currents could not be neglected. A motion measurement system was in place at the receiver, but not enough data were recorded to accurately track array movement.

Starting on 6 April 1990, pulses were transmitted every 12.0 min for a total of 494 successful transmissions. The resulting duration of the acoustic experiment was nearly 100 h. The approximate source level was 180 dB *re*:  $1 \mu Pa$  at 1 m, and to help simulate a single narrow pulse each transmission actually consisted of repetitions of an amplitude modulated sequence, 127 digits long, 4 cycles per digit, at a carrier

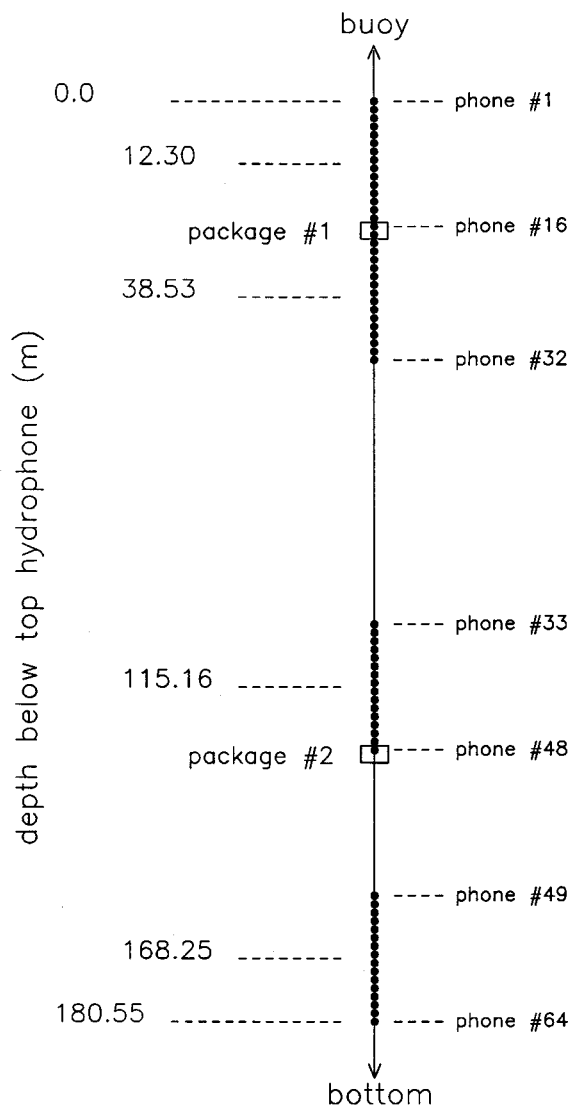


FIG. 2. Hydrophone array geometry. The array consisted of four sections of 16 hydrophones. For a straight array the centers of adjacent sections were separated by 26.23, 76.63, and 53.09 m.

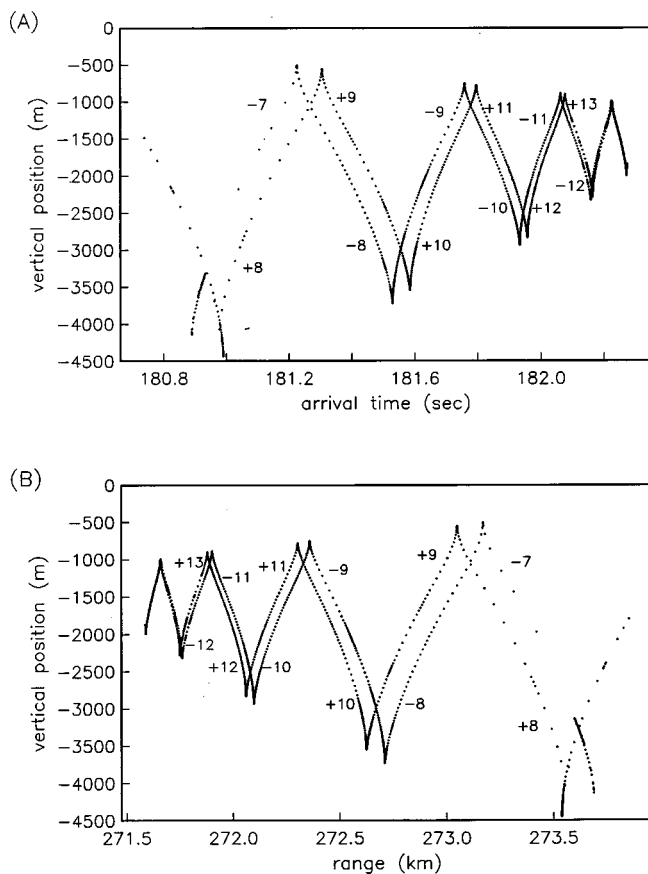


FIG. 3. (A) The geometrical optics *timefront* calculated using the reference soundspeed profile  $C_0(z)$ , a source depth of 1050 m, and a range of 272.3 km. Points on the curve correspond to uniformly incremented ray launch angles at the source. The integers shown are the segment identifiers. (B) The corresponding *wavefront* after a travel time of 181.8 s.

frequency,  $f_0$ , of 459.5588 Hz. At each array hydrophone the received signal yielded, after processing, 508 complex demodulates per pulse transmission. Each set of 508 complex demodulates spanned a period of 1.1054 s. Rubidium clocks on both moorings provided excellent timing.

## II. PREDICTED AVERAGE BEHAVIOR

For the averaged soundspeed profile,  $C_0(z)$ , the predicted arrival pattern of a transmitted pulse, based on geometrical acoustics and the estimated source depth at deployment, is presented in Fig. 3. The folded *timefront* shows the expected arrival time of a pulse as a function of depth after traveling a fixed horizontal distance of 272.3 km, and it confirms that from a single pulse transmission several signals separated in time can arrive at the same depth. The *timefront*'s construction consisted of incrementing (uniformly) the ray launch angle at the source, and plotting for each such angle the travel time and depth of the corresponding ray as it traversed the range of the receiver. Because for this particular soundspeed profile all rays that reflect from the surface are also bottom bounced, points resulting from steep positive and negative launch angles that correspond to any reflection are omitted in Fig. 3. The corresponding *wavefront* for a pulse 181.8 s after transmission is also shown in Fig. 3.

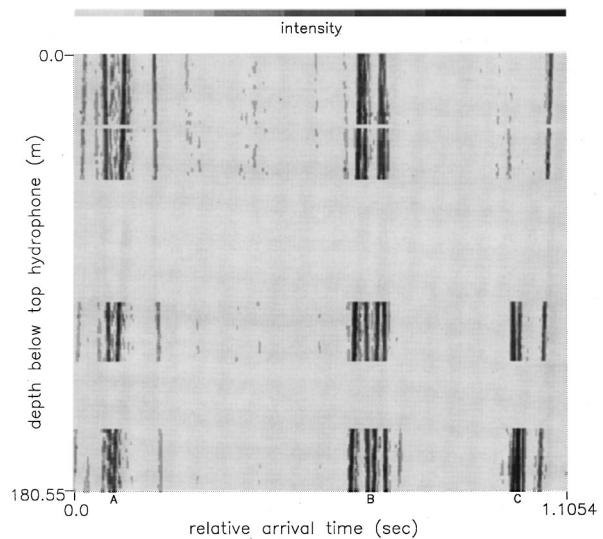


FIG. 4. The recorded intensity as a function of hydrophone position and (relative) arrival time for the 290th pulse transmission. The arrivals of separate wavefront segments are evident. Points A and B mark positions near the crossing of segments +8 and -8 and the crossing of segments +10 and -10, respectively. Point C locates the two upper caustics joining segment -11 with -12 and segment +12 with +13. The two blank horizontal strips correspond to depths where there was no hydrophone coverage.

Away from caustics, the distance between neighboring points on the wavefront is approximately inversely proportional to the local intensity.

The cusped points at the top and bottom of the *timefront* are caustics, subdividing the continuous *timefront* into smooth *segments*. Since (for simple range dependent scenarios) all the rays that make up a particular segment depart the source in the same direction and have the same total number of turning points, any segment is uniquely identifiable by an integer describing these common ray characteristics. For example, +10 and -10 are identifiers for those segments whose rays have 10 turning points and a launch angle that is initially upward and downward, respectively.

In light of the short length and (estimated) shallow depth of the hydrophone array, the *timefront* geometry in Fig. 3 suggests that if the averaged soundspeed profile predicts typical acoustic behavior, then signals received by the array should correspond to the uppermost (near-caustic) portion of *timefront* segments, and in some cases, even the caustics themselves. In fact, the two upper caustics joining segment -11 with -12 and segment +12 with +13 should fall within the receiver coverage, while segments corresponding to arrivals after +13 should be too deep to be detected (although a nongeometric arrival on the shadow zone side of a caustic may be detectable). Verifying that the vertical coverage was indeed in the near-caustic region, Fig. 4 reveals an upper swath of a *timefront* as detected by the hydrophone array for a typical pulse transmission. And in Fig. 5 plots of signal amplitude for late arrivals (corresponding to arrivals C in Fig. 4) are consistent with the amplitude pattern expected for the neighboring caustics joining segment -11 with -12 and segment +12 with +13. The 15-ms arrival time lag predicted in Fig. 6 for these caustics shows fair agreement (considering the uncertain source depth) with the 10-ms



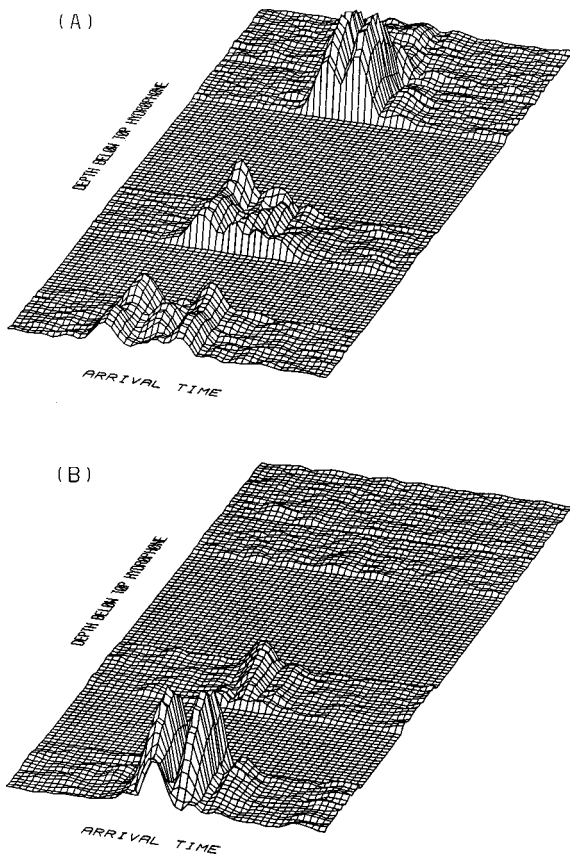


FIG. 5. (A) A close-in view for the 191st pulse transmission showing signal amplitude as a function of arrival time and vertical position along the array. The horizontal and vertical axes span 107 ms and 180 m, respectively. The two (adjacent) highest peaks correspond to the two caustics joining segment -11 with -12 and segment +12 with +13. (B) An analogous view for the 292nd pulse transmission. For pulse 292, after the array drifted upward, the coverage was just above the peaks, whereas for pulse 191 the receiver coverage was just below the caustic peaks.

separation between the twin peaks in Fig. 5. In Fig. 5(A) basic features of the amplitude pattern below the caustics can be explained as interference between four closely separated geometric arrivals (Fig. 6), each having a pulse duration of about 9 ms.

### III. DATA ANALYSIS

One of the principal goals of this analysis was to extract the temporal and vertical behavior of timefront fluctuations resulting from ocean internal waves. Since the magnitude of these fluctuations was expected to be on the order of the arrival time measurement error at a single hydrophone, an averaging or *beamforming* over several hydrophones was performed to improve signal-to-noise and consequently arrival time accuracy. Unfortunately, averaging over hydrophones eliminates small-scale vertical behavior. In order to examine small vertical scales, some arrival times were also identified for each *individual* hydrophone of a selected set.

#### A. Precision of arrival time measurement

An estimate of the arrival time measurement error at a single hydrophone was a necessary prerequisite to the analysis. The best estimate of this error was based on a simple

peak-picking simulation. For each hydrophone and transmission, a theoretical Gaussian pulse of the form  $Ae^{-(t-t_c)^2/\sigma^2}e^{-2\pi if_0(t-t_c)}$  was superimposed on the recorded noise that immediately followed the arrival of a selected wavefront segment. For each transmission the parameter  $A$  was chosen to match the measured amplitude of the selected wavefront segment. The peak position  $t_c$ , placed in a region of noise, varied linearly with hydrophone position so as to mimic a planar wavefront whose slope locally matched that of the selected segment. The parameter  $\sigma$  was fixed to make the Gaussian pulsewidth resemble the actual pulsewidth. Then for each hydrophone and transmission, the known position of the Gaussian peak was compared to that of the amplitude's peak determined after the Gaussian signal was discretized in time, added to the (already discretized) recorded noise, and interpolated (using the same method employed in the data analysis). The precision of arrival time, after simulations for several selected wavefront segments and different regions of noise, was consistently estimated at 0.4-ms rms. This compared favorably with both a slightly higher estimate based instead on the measured high frequency variance of arrival time between adjacent hydrophones, and a theoretical estimate using the average measured signal-to-noise ratio of 23 dB which predicted a lower bound of 0.3-ms rms.

#### B. Beamforming

For each of the four sections or *sets* of 16 hydrophones that comprised the array (Fig. 2), the complex demodulated data were locally beamformed to produce time series of both travel time and arrival angle for each identifiable wavefront segment. Each array section was assumed linear over its length of 24.6 m, and all of its (15 or 16) functioning hydrophones were used in the beamforming. Beamform processing is coherent averaging developed for an arriving wavefront

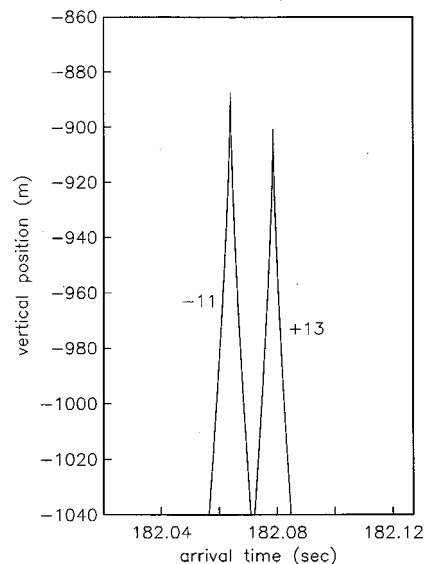


FIG. 6. The geometric optics prediction of arrival time for the upper sections of segments -11, -12, +12, and +13 based on a source depth of 1050 m. For purposes of comparison, the horizontal and vertical axes have the same length as in Fig. 5(A).

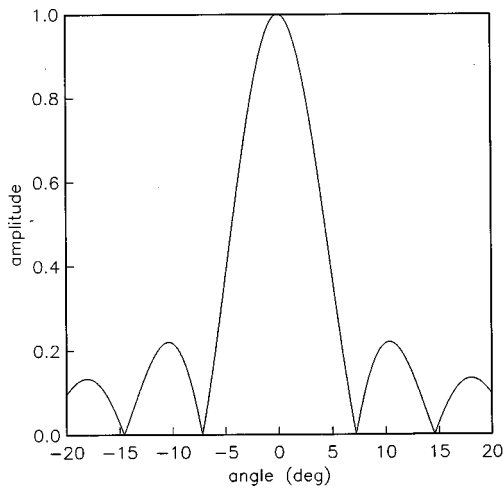


FIG. 7. The theoretical beam (amplitude) pattern for a 16 hydrophone section and an incoming plane wave of 459.5588 Hz and 0° arrival angle. Secondary lobe peaks occur at  $\pm 10.4^\circ$ .

that locally, over the set of relevant hydrophones, is planar. Then, a wavefront propagating at an angle  $\theta$  to the array will produce a time delay at hydrophone  $j$  given by  $\Delta t_j = \Delta z_j \sin \theta / C_0$ , where  $\Delta z_j$  is the distance of the  $j$ th hydrophone from some reference point (usually the center of the hydrophone set) and  $C_0$  is the local soundspeed.

For many  $\theta$  the complex demodulated time series,  $s_j(t)$ , were coherently averaged over  $N$  hydrophones by

$$S(t, \theta) = \frac{1}{N} \sum_{j=1}^N s_j(t + \Delta t_j) e^{-2\pi i f_0 \Delta t_j},$$

where  $f_0$  was the carrier frequency and  $N$  was either 15 or 16. Spline interpolation provided values for  $s_j(t)$  at all necessary times. With the exception of false peaks that resulted from side lobes in the beamforming, local maxima of  $|S(t, \theta)|$  in the  $t - \theta$  plane corresponded to the arrival of best-fit planar wavefronts at that time and angle. Angle information was useful for identifying the wavefront segment that equated to an arrival.<sup>25</sup> (Negative and positive  $\theta$  correspond to wavefronts that propagated with a downward and upward component along the array, respectively.) Figure 7 displays the beam (amplitude) pattern expected for 16 colinear hydrophones spaced at 1.64 m and an incoming plane wave of frequency  $f_0$ ; the central lobe has its first zeros near  $\pm 7^\circ$ , and the first side lobe peaks occur about  $10^\circ$  on either side of center. In Fig. 8 a sample  $|S(t, \theta)|$ , after beamforming with hydrophones 33–48, is shown for  $1^\circ$  increments of  $\theta$ . A wide central lobe and first side lobes are evident at the arrivals of several wavefront segments.

With consideration for false peaks arising from side lobes, arrival time and angle were identified, when possible, for each pulse transmission, each hydrophone set, and wavefront segments  $-7, \pm 8, \pm 9, \pm 10$ , and  $+11$ . Arrival information for a particular hydrophone set corresponded to the depth given by the center of the set. The resultant time series were very complete for segments  $-9, \pm 10$ , and  $+11$ , containing on average 480 points out of the 494 pulse transmissions; the other 4 segments had intermittent data,  $+9$  being especially deficient. Because of potential interference, adja-

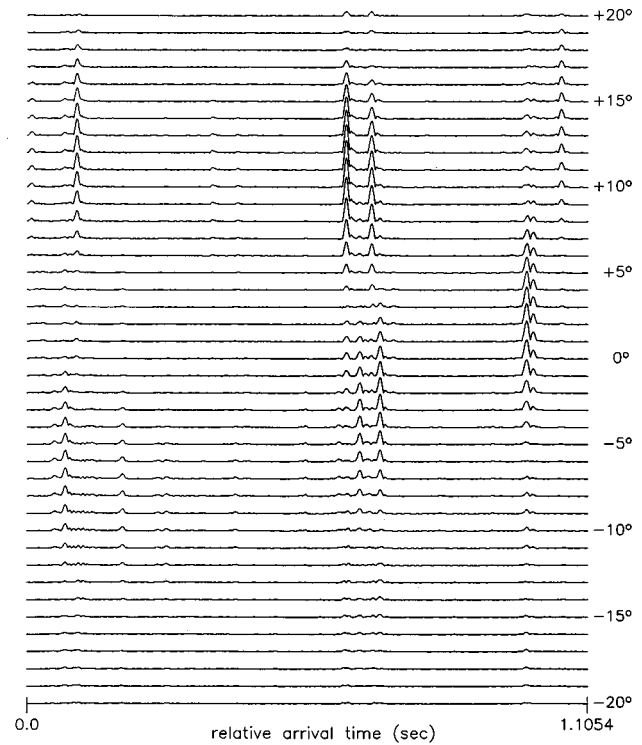


FIG. 8. The beamformed signal  $|S(t, \theta)|$  for the 290th pulse transmission (Fig. 4) and hydrophones 33–48. The dependence on  $\theta$  is only shown for  $1^\circ$  increments. Most outstanding are the arrivals of segments  $-8$  and  $+8$  (at left), segments  $-9, -10, +10$ , and  $+11$  (near the middle), and the two upper caustics formed by segments  $-11, -12, +12$ , and  $+13$  (at the center right). The  $14^\circ$  span of the central lobe is evident in each of the arrivals.

cent peaks closer than 4 ms and suspected of side lobe interference were rejected. The precision of beamformed arrival time was estimated at 0.15-ms rms, based on an estimated precision of 0.4-ms rms for arrival time at a single hydrophone and a measured 9-dB gain after beamforming. The arrival angle error, based on the high frequency variance of the difference in arrival angle between the two closest hydrophone sets (separated by 26.2 m), was estimated to be less than  $0.3^\circ$  rms.

In Fig. 9 the arrival time and angle are plotted as functions of real time for one particular wavefront segment and hydrophone set. This example is typical of all wavefront segments in that as the experiment proceeded there was a gradual trend in the measured arrival time and angle which dominated the smaller fluctuations. Unfortunately, this trend's underlying cause, as well as the primary contributor to time dependence of the measured quantities, was instrument motion. For example, the majority of the  $\sim 0.2$ -s difference in travel time between the earliest and latest arriving pulses presented in Fig. 9 is explained by a change of roughly 300 m in the relative horizontal distance between the source and receiver.

#### IV. INSTRUMENT MOTION

In this experiment, as in a similar acoustical experiment<sup>26</sup> in the same ocean region, instrument vertical drift, horizontal drift between the source and receiving array, and array tilt were on the order of tens of meters, a few

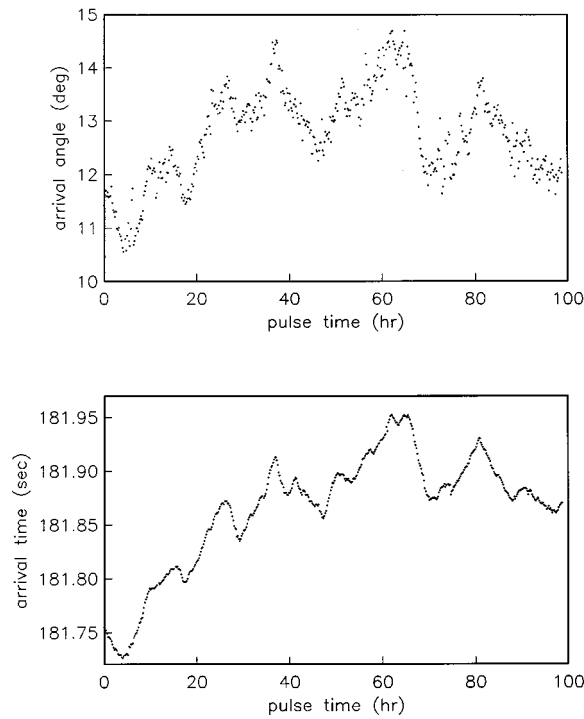


FIG. 9. The arrival angle and arrival time as functions of pulse time for segment -9 and the bottom set of 16 hydrophones. The large-scale fluctuations are attributable to instrument motion.

hundred meters, and several degrees, respectively. Instrument motions of this size have a profound influence on the measured fluctuations of arrival time and angle. Since instrument tracking was unsuccessful in this experiment, indirect *differencing* methods were used to separate the fluctuations due to mooring motion from those due to sound transmission through a changing ocean. Clearly, the fidelity of such indirect methods relies on understanding the nature and magnitude of the effects of instrument motion on the acoustical measurements.

Instrument motion had four primary components, each having a different effect on fluctuations of arrival time and angle: (A) the tilting motion of each hydrophone set, where each set of 16 hydrophones was believed to be aligned linearly; (B) the vertical movement of the source; (C) the vertical movement of each hydrophone set's center; and (D) the relative horizontal drift between the source and the center of each hydrophone set. (Here tilt means the included angle between the vertical and the projection of the hydrophone set onto the plane defined by the source-receiver and vertical directions.) The motions of the four hydrophone sets could differ since contiguous sets were either separated by a substantial distance or by a large, heavy package (Fig. 2) potentially altering the motion. Because all hydrophones were connected, significant correlation of tilting, up-down, and horizontal motions of hydrophone sets was expected.

The effect of array tilt on arrival time and angle was simple to predict; the effects of vertical and horizontal instrument motion were estimated to lowest order by ignoring ocean fluctuations and assuming that the averaged soundspeed profile governed the acoustics.

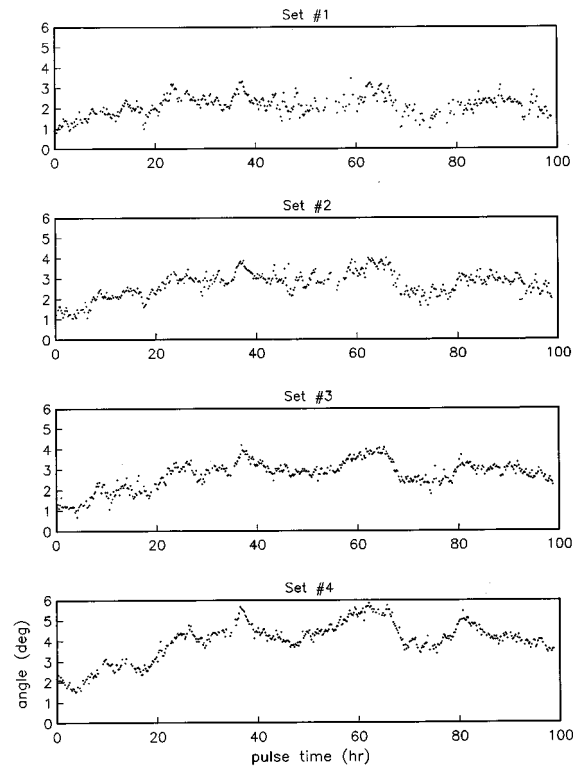


FIG. 10. Array tilt as a function of real time for all four hydrophone sets. These time series are actually the mean of the *averaged* arrival angles for wavefront segments -9 and -10 and the *averaged* arrival angles for segments +10 and +11, the (rms) tilt uncertainty being about  $.2^\circ$ . The deepest and shallowest hydrophone sets had the largest and smallest tilt swings, respectively. The middle two sets, which were not separated by system packages, had similar tilts.

### A. Array tilt

For any hydrophone set, a change of tilt results in a change of equal magnitude in the measured wavefront arrival angle. A positive tilting (the top hydrophone moves away from the source, relative to the bottom hydrophone) increases the measured arrival angle, and a negative tilting decreases it. The arrival time and the *difference* in arrival angle between any two wavefront segments are indifferent to tilt changes, as long as the center of the tilting set is stationary.

Figure 10 shows the tilt as a function of geo-time for each hydrophone set. The tilt was not measured directly but inferred from symmetry: In the neighborhood of an upper caustic, where the pair of joining wavefront segments has bilateral symmetry about a nearly vertical axis, the *average* of the measured arrival angles for the joining segments closely approximates the local array tilt. Differences between the figure's four time series suggest that the hydrophone array was not linear but rather *piecewise* linear, the top 16 hydrophones, the next 2 sets of 16 hydrophones (and the connector between them), and the bottom 16 hydrophones forming 3 separate linear pieces. The means of the time series are, from top to bottom,  $2.12$ ,  $2.71$ ,  $2.73$ , and  $3.99^\circ$ , indicating that the two bends were small: the angle between the first and second linear sections was only  $0.6^\circ$  on the average, while the angle between the second and third sec-

tions averaged  $1.3^\circ$ . These two kinks were probably caused by the 2 heavy system packages, one located at the junction of the first and second set of 16 hydrophones, and the other placed at the bottom of the third hydrophone set. As might be expected from a swaying array, the leading temporal behavior of array tilt (especially at the bottom hydrophone set) reflected that of wavefront arrival time (Fig. 9).

### B. Source vertical motion

The effect of source vertical motion is best described by Fig. 11, where sections of the predicted timefront are displayed for source depths of 1050 and 1100 m. This 50-m displacement results in a several millisecond change in the time lag between the arrivals of adjacent caustics. More specifically, it produces an arrival time shift of  $\sim 5$  ms for timefront segments  $-9$ ,  $\pm 10$ , and  $+11$ , and  $\sim 6$  ms for timefronts  $-7$ ,  $\pm 8$ , and  $+9$ ; the shifts for segments with positive identifiers have opposite sign to those with negative identifiers, implying that arrival time changes are (practically) identical for neighboring segments that share an upper caustic. Further examination of Fig. 11 shows that the changes in the vertical position of caustics and the wavefront arrival angles are negligible, and the change in arrival time along any wavefront segment is virtually independent of depth, a characteristic later exploited in the analysis.

### C. Receiver vertical motion

Because arrival time varies with depth along any timefront segment, measured arrival times are affected by receiver vertical motion. Likewise, since every wavefront has significant curvature (especially near caustics), measured arrival angles are also affected by receiver vertical motion. A depth change from 900 to 1000 m, for example, is predicted to bring about a  $\sim 8$  ms and  $\sim 2^\circ$  change in arrival time and arrival angle, respectively, for segment  $+10$ ; however, the

*difference* in the changes corresponding to two nearly parallel timefront segments (e.g.,  $-9$  and  $+10$ ) are relatively small. Vertical motion of the hydrophone array also gives rise to an apparent vertical motion of the caustics. For this experiment the approximate position along the array of the caustic joining segments  $-11$  and  $-12$  was identified for more than 50 pulse transmissions and revealed maximum displacements of greater than 100 m relative to the array (Fig. 5), mostly due to array motion.

### D. Relative horizontal motion

Relative horizontal motion between the source and any hydrophone set produces, by altering the range of propagation, changes in the arrival time measured at that set. These changes, however, differ only slightly among (and along) all wavefront segments, so that at any hydrophone set the *difference* in arrival time between any pair of segments, or any pair of caustics, changes only slightly. In the scenario of this experiment an increase of 300 m in source-receiver range, for example, equates to an increase of only a fraction of a millisecond in the difference of arrival time between the upper caustic joining segment  $+10$  with  $+11$  and that joining  $-9$  with  $-10$ . Vertical displacement of caustics and changes in wavefront arrival angle are also negligibly affected by source-receiver drift.

### V. RELATIVE FLUCTUATIONS FOR HYDROPHONE SETS

In this section the information on individual hydrophones is not kept; only the averages of arrival angle and arrival time over subarrays of 16 hydrophones each are used. The individual hydrophone information is discussed in Sec. VI. The deviations of individual hydrophone arrival times from the straight line fits implied by beamforming are typically on the order of 0.5-ms rms, which is consistent with the small-scale vertical behavior of fluctuations observed in Sec. VI. Beamforming, which can be thought of as an amplitude-weighted coherent average, gave different slopes for wavefront segments than did fitting (by least squares) straight lines to arriving peaks at individual hydrophones. The peak line fits gave wavefront arrival angles that exhibited fluctuations with an rms value several (five or so) times that of fluctuations for the angles given by beamforming; there was a difference between the two determined angles of roughly  $2^\circ$  rms.

For each set of 16 hydrophones, *differences* in the arrival time or angle between separate wavefront segments were constructed, eliminating some of the leading effects of instrument motion. Differences in the arrival time removed the major contribution of horizontal drift between the source and receiver, although lesser contributions from source and receiver vertical motion persisted. Differences in the arrival angle eliminated the consequences of array tilt, although the secondary effects of receiver vertical motion remained. And in both cases, differences avoided any depth dependence attributable to array shape. Differencing, however, shifted the emphasis from an analysis of *absolute* fluctuations to one of *relative* fluctuations<sup>19</sup> between front segments.

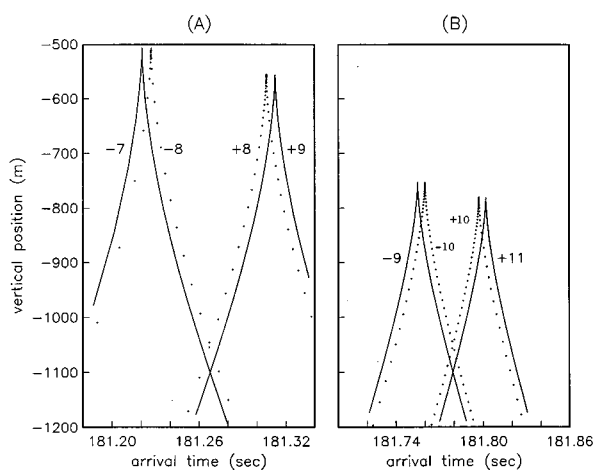


FIG. 11. (A) The predicted effect of source vertical motion in the upper part of timefront segments  $-7$ ,  $-8$ ,  $+8$ , and  $+9$ . The solid and dotted curves are the timefront predicted for a source depth of 1100 and 1050 m, respectively. With a source depth increase of 50 m segments  $-7$  and  $-8$  arrive about 6 ms earlier, while segments  $+8$  and  $+9$  arrive about 6 ms later. (B) A similar plot of the upper timefront near segments  $-9$ ,  $-10$ ,  $+10$ , and  $+11$ . Here the 50-m source displacement leads to approximately a 5-ms arrival time change.

Ocean fronts, mesoscale eddies, surface and internal tides, and internal waves will all contribute (on different time scales) to absolute fluctuations in arrival time and angle. For some of these ocean processes the induced relative fluctuations will be substantially smaller than the corresponding absolute ones, since the difference in arrival time or angle of two wavefront segments removes any contribution common to both segments. For example, because its current is fairly uniform in the vertical, the surface tide will effect arrival time changes that are almost identical over the entire wavefront, resulting in comparatively small fluctuations in the arrival time difference between separate wavefront segments.

If  $T(i,j, ID)$  and  $A(i,j, ID)$  are the measured arrival time and angle, respectively, corresponding to pulse time  $t_i$  (where  $i=1$  to 494), hydrophone set  $j$  (where  $j=1$  to 4), and wavefront segment  $ID$  (where  $ID=-7, \pm 8, \pm 9, \pm 10$ , and  $+11$ ), then  $\Delta T$  and  $\Delta A$  shall denote the respective differences between wavefront segments:

$$\Delta T(i,j; ID_1, ID_2) = T(i,j, ID_1) - T(i,j, ID_2),$$

and similarly for  $\Delta A$ . Primarily because the effect on  $\Delta T$  or  $\Delta A$  of source vertical motion is qualitatively different than that of receiver vertical motion, differences of wavefront segments having a *common caustic* and differences of *parallel neighboring* wavefront segments were analyzed separately.

### A. Arrival time differences of segments having a common caustic

For a pair of timefront segments that share an upper caustic, such as segments  $-9$  and  $-10$ , source vertical motion will not contribute significantly to fluctuations in  $\Delta T$  (Fig. 11). Unfortunately,  $\Delta T$  from segments sharing a common caustic may have contributions from vertical motion of the array as well as from ocean processes. (As shown later, the differences between neighboring timefronts that emanate from different caustics can eliminate the effects of array motion.) Figure 12 displays the temporal variations in  $\Delta T$  at the bottom hydrophone set for three consecutive pairs of contiguous timefront segments; in light of the sawtooth pattern that timefronts exhibit, the significant correlation of  $\Delta T(i,4; -7, -8)$  and  $\Delta T(i,4; -9, -10)$  and their significant anti-correlation with  $\Delta T(i,4; -8, -9)$  suggest that an underlying vertical translation, relative to the bottom hydrophone set, existed for the whole timefront. Analogous  $\Delta T$ 's at the other hydrophone sets similarly support a vertical motion of the timefront, relative to the array.

In the neighborhood of each upper caustic, the local geometry of the timefront controls the effect on  $\Delta T$  of relative vertical motion between the timefront and array. Because this local geometry is different for different caustics, vertical motion was extracted individually for each upper caustic; the guiding decomposition was

$$\Delta T(i,j) = f(z_j - Z(i)) + \epsilon(i,j),$$

where the dependence of  $\Delta T$ ,  $f$ ,  $Z$ , and  $\epsilon$  on the segments ( $ID_1$  and  $ID_2$ ) is implicit. The term involving the function  $f$  represents that part of  $\Delta T$  attributable to relative vertical translation of a locally rigid timefront; the argument of  $f$ ,  $z_j$

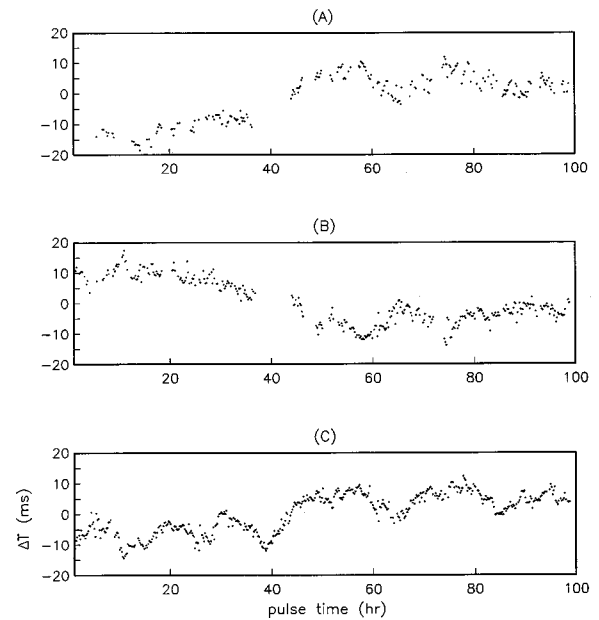


FIG. 12. (A)  $\Delta T(i,4; -7, -8)$ . (B)  $\Delta T(i,4; -8, -9)$ . (C)  $\Delta T(i,4; -9, -10)$ . (The horizontal axis is pulse time rather than pulse number  $i$ .)

$-Z(i)$ , is the vertical position of the  $j$ th hydrophone set relative to the caustic (in the reference frame moving with the caustic), while  $z_j$  and  $Z(i)$  are the hydrophone set and caustic vertical positions in the reference frame of the array. (Each  $z_j$  was approximated as constant, since array tilts less than  $6^\circ$  implied changes less than a meter.) For a short array and small vertical displacements the smooth function  $f$ , only locally resolvable, is adequately described by a truncated Taylor series:

$$f(\zeta) = \alpha + \beta(\zeta - \zeta_0) + \gamma(\zeta - \zeta_0)^2 + \delta(\zeta - \zeta_0)^3,$$

where  $\zeta_0$  was taken to be the top hydrophone's *mean* position relative to the caustic (located at  $\zeta=0$ ), and the nonlinear terms were included to account for timefront curvature. Since  $\zeta$  is the vertical coordinate in the reference frame of the caustic,  $f(\zeta)$  is the profile of arrival time difference for the locally rigid timefront. The function  $Z(i)$  and  $\alpha$ ,  $\beta$ ,  $\gamma$ , and  $\delta$ , which do not depend on  $i$  or  $j$ , were found so as to force  $\sum_i \epsilon(i,j) = 0$  for all  $j$  and  $\sum_j \epsilon(i,j) = 0$  for all  $i$ ; the resulting root-mean-square of the residuals  $\epsilon(i,j)$  was small, being about 1 ms. (These constraints were imposed on the residuals so that they would satisfy the same conditions as the residuals in Sec. V B.)

Specified to have zero mean, the inferred vertical motions  $Z(i)$  of two neighboring upper caustics are shown in Fig. 13. [Since only those pulse times with  $\Delta T$  defined at all four hydrophone sets were applicable, insufficient data ruled out extracting  $Z(i)$  for other caustics.] The figure shows very similar motions for this pair of caustics, and the power spectrum for the two time series has peaks near the inertial and semi-diurnal periods, indicating that inertial waves and tides directly (through medium changes) or indirectly (through array motion) played an important role in the relative vertical translation of caustics. The difference of these two time series is also shown in Fig. 13; its variance of almost  $100 \text{ m}^2$  gives 10 m, after accounting for an estimated 2-m rms error

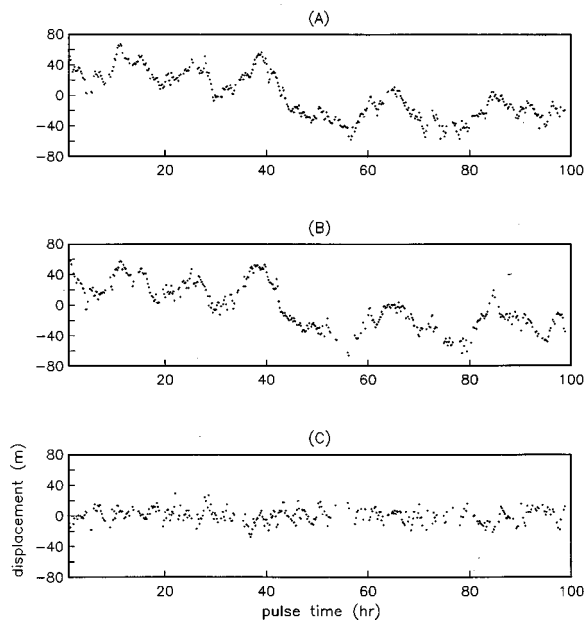


FIG. 13. (A) The vertical caustic displacement (relative to the array),  $Z$ , as a function of pulse time for the caustic joining segments  $-9$  with  $-10$  and (B) for the caustic joining segments  $+10$  with  $+11$ . These figures indicate a 30-m rms fluctuation in the caustics' vertical positions along the array, probably mostly due to array motion. (C) Free of array motion, the difference of the two time series in (A) and (B).

in each  $Z(i)$ , as a rough estimate of the rms relative vertical displacement between the two caustics. Such a relative motion cannot come from source or array motion and can be interpreted as an effect of ocean processes.

The constants  $\alpha$ ,  $\beta$ ,  $\gamma$ , and  $\delta$ , which describe  $f(\zeta)$  near  $\zeta_0$ , are presented in Table I. *Measured* values, as well as values *predicted* from the averaged soundspeed profile, are given for the two caustics referred to in Fig. 13; since  $\zeta_0$  is not independently known from the experiment, predicted values are based on a  $\zeta_0$  chosen to make the measured  $\alpha$  and the predicted  $\alpha$  agree. The measurements of  $\gamma$  show that the rigid timefront's curvature was not negligible, accounting alone for up to 3-ms difference in  $\Delta T$  from the top to bottom hydrophone set. The major source of discrepancy between measured and predicted values is the crudeness of the averaged soundspeed profile in characterizing the ocean medium. An examination of the CTDs from both the deployment and recovery cruises showed a variable, underlying range dependent ocean, having warmer water generally closer to the receiver.

The residuals  $\epsilon(i, j)$  represent that part of  $\Delta T(i, j)$  not described by a *smooth rigid* timefront that translates verti-

TABLE I. Taylor series coefficients.

	$-9$ and $-10$		$+10$ and $+11$	
	measured	predicted <sup>a</sup>	measured	predicted <sup>b</sup>
$\alpha$ (ms)	$-14.0 \pm 0.02$	$-14.0$	$-7.6 \pm 0.02$	$-7.6$
$\beta$ (ms m <sup>-1</sup> )	$0.154 \pm 0.001$	$0.145$	$0.134 \pm 0.001$	$0.118$
$\gamma$ (ms m <sup>-2</sup> )	$(-7 \pm 2) \cdot 10^{-5}$	$-24 \cdot 10^{-5}$	$(-13 \pm 2) \cdot 10^{-5}$	$-32 \cdot 10^{-5}$
$\delta$ (ms m <sup>-3</sup> )	$(3 \pm 1) \cdot 10^{-7}$	$-3 \cdot 10^{-7}$	$(0.5 \pm 1) \cdot 10^{-7}$	$-6 \cdot 10^{-7}$

<sup>a</sup>Predicted for a  $\zeta_0$  at 902-m depth.

<sup>b</sup>Predicted for a  $\zeta_0$  at 878-m depth.

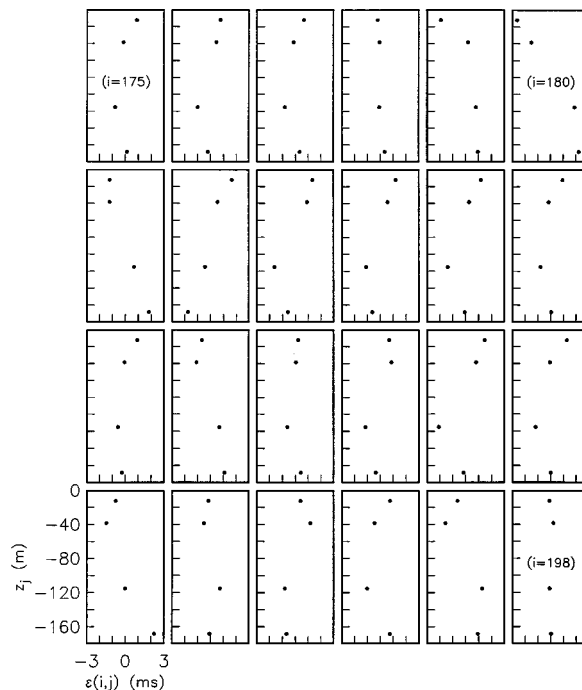


FIG. 14. The vertical profile of the residuals  $\epsilon(i, j; -9, -10)$  for pulse numbers 175–198. For each pulse the average of the residuals over depth is zero.

cally relative to the array. In addition to arrival time measurement error, contributors to  $\epsilon(i, j)$  can include steady and time dependent contortions (wiggles or folds) along the timefront; steady contortions, although fixed in place along the timefront, can still induce time dependence in  $\epsilon$  by means of vertical array movement. The residuals are small in contrast to  $\Delta T$ : The variance of each residual time series was approximately  $1 \text{ ms}^2$ , and the averages of the variances over  $j$  for  $\epsilon(i, j; -9, -10)$  and  $\epsilon(i, j; +10, +11)$  were  $1.03 \text{ ms}^2$  and  $1.08 \text{ ms}^2$ , respectively. Because the variance due to  $\Delta T$  measurement noise is estimated at  $0.04 \text{ ms}^2$  (corresponding to a  $\sim 0.15$ -ms rms error in  $T$ ), the variance of the signal in these residuals which must come from ocean processes is near  $1 \text{ ms}^2$ .

Figure 14 displays the residuals  $\epsilon(i, j; -9, -10)$  as a function of depth for 24 consecutive pulses, and it suggests a timefront wiggling that is fairly correlated between neighboring hydrophone sets. To specifically address the spatial and temporal correlations the following covariance estimator was constructed:

$$C(\Delta z, \Delta t) = C_{j_1, j_2}(k) = \frac{1}{M} \sum_{i=1}^M \epsilon(i, j_1) \epsilon(i+k, j_2),$$

where  $M$  is the number of points for which the term being summed is defined. This covariance estimator is a function of vertical lag (vertical distance  $\Delta z$  between hydrophone sets  $j_1$  and  $j_2$ ) and time lag (index  $k$  representing a time lag of  $\Delta t = 12 \cdot k$  min).  $C_{j_1, j_2}(0)$  estimates the covariance between hydrophone sets  $j_1$  and  $j_2$ , and  $C_{j, j}(k)$  estimates the auto-covariance of the time series for hydrophone set  $j$ .  $C_{j_1, j_2}(0)$  is plotted against the hydrophone set separation  $|z_{j_1} - z_{j_2}|$  in Fig. 15; for zero separation the average of  $C_{j, j}(0)$  over all four sets is used. The covariance results shown in the figure,

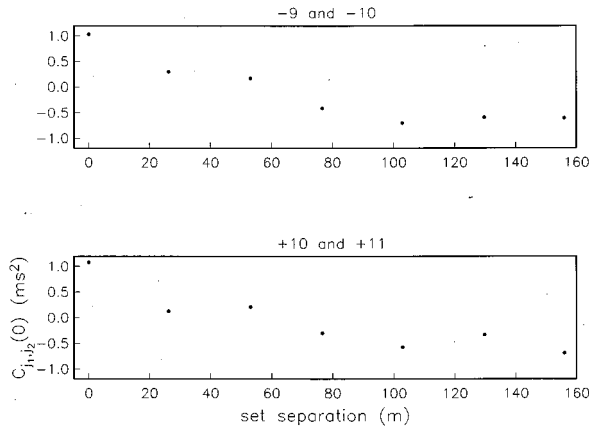


FIG. 15. Plots of  $C_{j_1, j_2}(0)$ , the residual covariance between hydrophone sets  $j_1$  and  $j_2$ , versus the hydrophone set separation  $|z_{j_1} - z_{j_2}|$ . Except for the 0 lag points (estimated twice as accurately) the precision of all other points was estimated at about  $0.04 \text{ ms}^2$ .

for the two different near-caustic regions, have similar shapes that show contributions for vertical scales up to roughly 100 m. The average of  $C_{j, j}(k)$  over  $j$ ,  $C_{ave}(k)$ , is plotted against lag time in Fig. 16, and for both regions it reveals a very quick decorrelation in time, on the order of 20–30 min. The abrupt drops from zero lag to first lag in Figs. 15 and 16 suggest a contribution from a fluctuation of roughly  $0.5 \text{ ms}^2$  that is uncorrelated over the time and (hydrophone set) depth sampling intervals. All this detailed behavior of  $C(\Delta z, \Delta t)$  is a measure of the effects of ocean processes.

### B. Arrival time differences of parallel neighboring segments

For a pair of nearly parallel timefront segments, such as  $-9$  and  $+10$ , variations in  $\Delta T$  due to instrument motion are practically independent of hydrophone set. For example, although a 50-m vertical displacement of the source produces a 10-ms change (Fig. 11) in  $\Delta T$  between segments  $-9$  and  $+10$ , and the corresponding change due to the same displacement of the receiver is 0.5 ms (inferred from the differ-

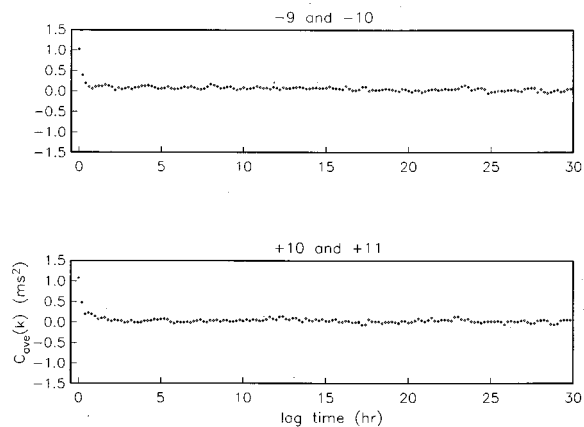


FIG. 16. Plots of  $C_{ave}(k)$ , the average residual autocovariance, versus lag time. The precision of the points was estimated to be  $0.02\text{--}0.03 \text{ ms}^2$ .

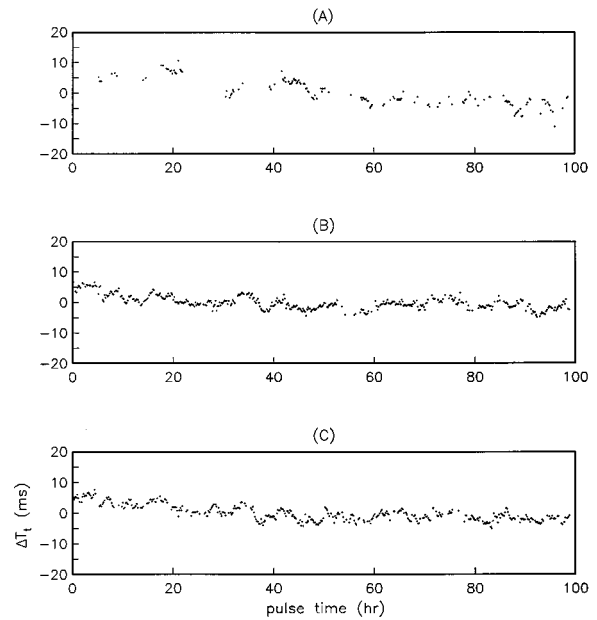


FIG. 17. (A)  $\Delta T_i(i; -7, +8)$ . (B)  $\Delta T_i(i; -9, +10)$ . (C)  $\Delta T_i(i; -10, +11)$ . (The horizontal axis is pulse time rather than pulse number  $i$ .)

ences of measurements in Table I), these changes themselves differ by less than 0.1 ms from the top to the bottom hydrophone set. Therefore, any significant variation in  $\Delta T$  that is not the same for different hydrophone sets is not ascribable to instrument motion and must be due to ocean processes. Furthermore, since the effect on  $\Delta T$  of receiver motion is small, those fluctuations in  $\Delta T$  that are common to all hydrophone sets will be dominated by source vertical motion, with inseparable contributions from ocean processes at some level.

In order to isolate fluctuations common to all hydrophone sets,  $\Delta T$  was broken up<sup>8</sup> according to

$$\Delta T(i, j) = \Delta T_i(i) + \Delta T_z(j) + \epsilon(i, j),$$

where  $\Delta T_z$  was the average over  $i$  of  $\Delta T$ , and  $\Delta T_i$  was the average over  $j$  of  $\Delta T - \Delta T_z$ . By this construction,  $\sum_i \Delta T_i(i) = 0$ . (The set of  $i$  was again restricted to only those pulse times with  $\Delta T$  defined at all four hydrophone sets.) The  $\epsilon$  term is due to ocean processes only; it automatically satisfies  $\sum_j \epsilon(i, j) = 0$  for all  $i$  and  $\sum_i \epsilon(i, j) = 0$  for all  $j$ . Source vertical motion will always show up in  $\Delta T_i$ , since in the near-caustic regions of a timefront this motion results in a relative horizontal translation between neighboring, rigid, parallel timefront segments. Vertical motion of the array will contribute a small amount to  $\Delta T_i$ . But vertical motions of the source and the array will have insignificant effects on  $\epsilon$ .

Figure 17 shows  $\Delta T_i(i)$  for three pairs of neighboring, nearly parallel timefront segments. The strong resemblance between  $\Delta T_i(i; -9, +10)$  and  $\Delta T_i(i; -10, +11)$  suggests that their dominant fluctuation, possibly due to source motion, was relative *horizontal* motion between the two upper caustics linking segments  $-9$  and  $-10$  and segments  $+10$  and  $+11$ . If the three time series were identical, it would be tempting to interpret the changes with time as due to source motion; but two are similar (although not identical) and the

third is very different. Since the amount of source motion is unknown, its effect cannot be directly removed, and hence the difference between pairs, which primarily reflects ocean processes, is taken. The difference  $\Delta T_i(i; -9, +10) - \Delta T_i(i; -10, +11)$  has a variance of  $2.5 \text{ ms}^2$ , which can be divided as follows: a nearly  $0.1\text{-ms}^2$  contribution is due to measurement error; based on the shape of the near-caustic timefronts given in Table I,  $\sim 2 \text{ ms}^2$  of the remaining  $2.4 \text{ ms}^2$  arises from the  $\sim 10\text{-m}$  rms relative vertical motion between the caustics that was previously estimated; and the other  $\sim 0.4 \text{ ms}^2$  is consistent with the effect of relative vertical motion of the array (Fig. 13). Other than the similar trend of initial decrease followed by a leveling off,  $\Delta T_i(i; -7, +8)$  is quite different from the other two time series in Fig. 17. The differences  $\Delta T_i(i; -7, +8) - \Delta T_i(i; -9, +10)$  and  $\Delta T_i(i; -7, +8) - \Delta T_i(i; -10, +11)$  both have variances slightly above  $12 \text{ ms}^2$ . Only ocean processes can account for this big difference, since the contribution of source vertical motion is estimated to be less than  $1 \text{ ms}^2$ .

An examination of  $\Delta T_z(j)$  revealed an average time lag of about 42 ms, differing from the predicted value of 37 ms, between the arrival of the caustic joining  $-9$  and  $-10$  and the caustic joining  $+10$  and  $+11$ . (Again, the difference is probably due to the averaged soundspeed profile's inaccurate description of the ocean medium.)  $\Delta T_z(j)$  also verified that each of the three segment pairs was only slightly nonparallel, and it reaffirmed a  $\sim 0.4\text{-ms}^2$  variance for fluctuations of  $\Delta T_i(i; -9, +10) - \Delta T_i(i; -10, +11)$  resulting from the  $\sim 30\text{-m}$  rms relative vertical displacement of the array (Fig. 13).

Accounting for timefront contortions and measurement error, the residuals  $\epsilon(i, j)$  describe fluctuations in  $\Delta T$  that are not common to all hydrophone sets. In all cases the residuals are small, and the residual time series show randomness and no trends. The averages of the variances over  $j$  for  $\epsilon(i, j; -10, +11)$ ,  $\epsilon(i, j; -9, +10)$ , and  $\epsilon(i, j; -7, +8)$  are, in that order,  $1.09 \text{ ms}^2$ ,  $0.72 \text{ ms}^2$ , and  $0.35 \text{ ms}^2$ . These variances represent signal due to ocean processes, since the variance from  $\Delta T$  measurement noise is  $\sim 0.04 \text{ ms}^2$  and the effect on the residuals from source and array motion is even smaller.

The residual covariance between hydrophone sets  $j_1$  and  $j_2$ ,  $C_{j_1, j_2}(0)$ , shows for all three segment pairs a decorrelation length of almost 100 m. Plots of lag time versus  $C_{ave}(k)$  show, as they did in the case of segments sharing an upper caustic, a 20–30 min decorrelation in time. The abrupt drops from zero lag to first lag in  $C_{ave}(k)$  and  $C_{j_1, j_2}(0)$  indicate a contribution from a fluctuation of typically  $0.5 \text{ ms}^2$  that is uncorrelated over the time and (hydrophone set) depth sampling intervals.

Figure 18 shows a comparison of eigenray paths for parallel neighboring segments  $-10$  and  $+11$ , as well as for segments  $+10$  and  $+11$  that share a common caustic. The eigenrays in each pair sample slightly different regions of the ocean as seen by the differences in the vertical and horizontal positions of their upper and lower turning points, and  $\Delta T$  echoes the differences in soundspeed along these two distinct ray paths. The eigenray paths for parallel neighboring segments deviate more away from the receiver, while the eigenray paths for segments sharing a common caustic deviate

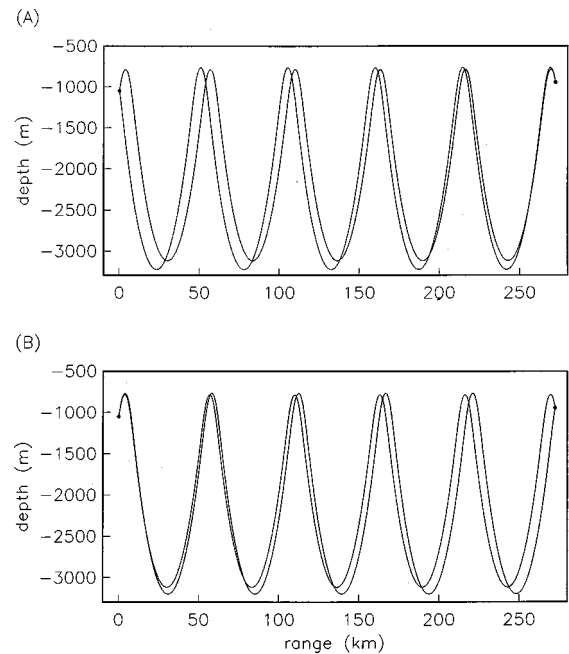


FIG. 18. (A) Comparison of eigenray paths for (parallel neighboring) segments  $-10$  and  $+11$ . Source and receiver depths are 1050 and 950 m, respectively. (B) For the same source and receiver depths, comparison of eigenray paths for segments  $+10$  and  $+11$  (having a common caustic). Because source and receiver depths for these eigenrays are comparable, eigenrays in the top figure almost look like those in the bottom figure with the source and receiver interchanged.

more away from the source; however, the geometry of the eigenray pair for segments  $-10$  and  $+11$  looks like that for segments  $+10$  and  $+11$  with the source and receiver interchanged. Recall that the residuals for segment pairs  $(-10, +11)$  and  $(+10, +11)$  are quantitatively similar (e.g., both have a variance of about  $1 \text{ ms}^2$ ).

### C. Arrival angle differences of parallel neighboring segments

Beamforming also yielded information about wavefront arrival angle, where arrival angle meant the inclination describing a planar wavefront that best fit (over the set of hydrophones) the actual wavefront. Arrival angle differences for parallel neighboring segments were examined since, unlike the angle differences for segments sharing a common caustic, the effect of the receiving array vertical motion was not expected to dominate the fluctuations. Time series displays of  $\Delta A(i, j; -10, +11)$  for each  $j$  show that fluctuations in  $\Delta A$  are small and absent of dominant instrument motion effects (unlike  $\Delta T$ ). The fluctuations in  $\Delta A$  are mostly due to noise, with a smaller contribution from relative vertical motion of the receiving array.

The averages of variances over  $j$  for the differences  $\Delta A(i, j; -10, +11)$ ,  $\Delta A(i, j; -9, +10)$ , and  $\Delta A(i, j; -7, +8)$  are  $0.26 \text{ deg}^2$ ,  $0.24 \text{ deg}^2$ , and  $0.15 \text{ deg}^2$ , respectively. As with  $\Delta T$  in Sec. VB,  $\Delta A$  can be decomposed into its components of  $\Delta A_t, \Delta A_z$ , and the residuals  $\epsilon$ .  $\Delta A_t$  accounts for about  $0.1 \text{ deg}^2$  of the variance in  $\Delta A$  for the pairs  $(-10, +11)$  and  $(-9, +10)$ , and it accounts for about  $0.05 \text{ deg}^2$  of the variance in  $\Delta A$  for the pair  $(-7, +8)$ . These values are consistent with effects due to relative vertical motion of the



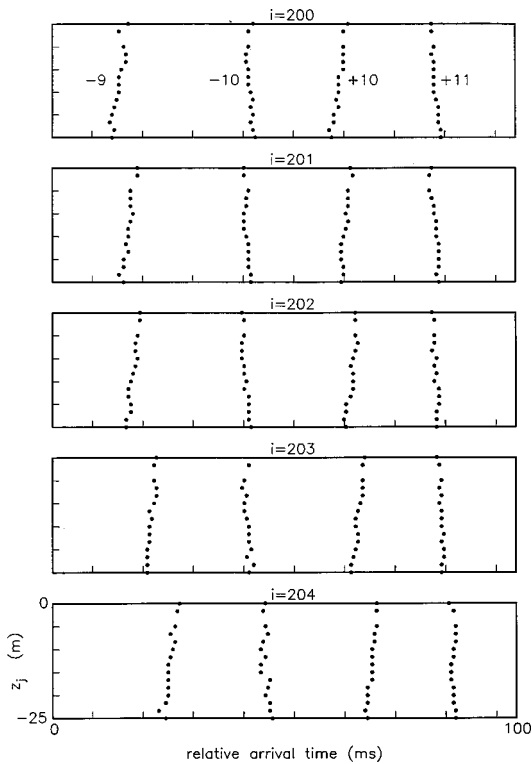


FIG. 19. For pulse numbers 200–204 and hydrophones 17–32, arrival time as a function of vertical position for timefront segments  $-9$ ,  $-10$ ,  $+10$ , and  $+11$ . Each arrival time corresponds to the arrival of a well-defined pulse peak. (Hydrophone 19 malfunctioned during the experiment.)

receiver. (The behavior of the time series for  $\Delta A_t$ , reminiscent of that for  $\Delta T_t$ , also indicates a relative vertical motion between the receiver and the wavefront segments.) Note that the contribution from receiver motion is greater for the later arriving wavefront segments where the (near-caustic) curvature is greater. The rest of the variance in  $\Delta A$  comes from the residuals.

The residuals are small, and the averages of the variances over  $j$  for  $\epsilon(i, j; -10, +11)$ ,  $\epsilon(i, j; -9, +10)$ , and  $\epsilon(i, j; -7, +8)$  are, in that order,  $0.16 \text{ deg}^2$ ,  $0.15 \text{ deg}^2$ , and  $0.11 \text{ deg}^2$ . These variances represent signal mostly due to noise and not ocean processes, since the variance from  $\Delta A$  measurement noise is estimated to be in the neighborhood of  $0.12 \text{ deg}^2$ . The residual covariance for these three pairs of parallel segments indicate a decorrelation length less than 26 m, the smallest sampling length, and this is consistent with a noise contribution. Plots of lag time versus  $C_{ave}(k)$  show a significant temporal decorrelation from zero lag to first lag.

## VI. RELATIVE FLUCTUATIONS FOR INDIVIDUAL HYDROPHONES

In order to examine the *small-scale* vertical behavior of timefront fluctuations, or even to examine the feasibility of measuring this behavior (considering the overwhelming measurement error), the preceding analysis of arrival time differences of parallel neighboring segments was performed for the *individual* hydrophones of a single set. Motivating this study of small scales was Fig. 19, which displays for 5 consecutive pulses the (relative) arrival time measured at hydrophones 17–32 and segments  $-9$ ,  $\pm 10$ , and  $+11$ . Each

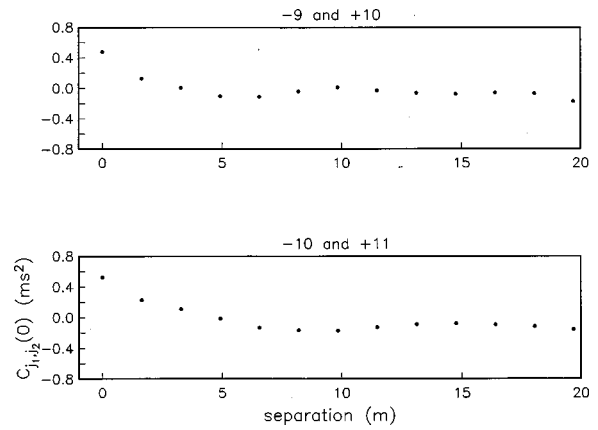


FIG. 20. Plots of  $C_{j_1, j_2}(0)$ , the residual covariance between hydrophone sets  $j_1$  and  $j_2$ , versus the hydrophone separation  $|z_{j_1} - z_{j_2}|$ . Except for the 0 lag points (estimated twice as accurately) the precision of all other points was estimated at about  $0.04 \text{ ms}^2$ .

timefront segment exhibits notable time dependent, vertical contortions; however, without further analysis it is unclear as to whether small-scale contortions like these could be fully accounted for by array wiggling and measurement error, or might instead be due to ocean internal waves.

To eliminate contributions from array wiggles, differences in the arrival time between segments  $-9$  and  $+10$  and segments  $-10$  and  $+11$  were constructed for each hydrophone and pulse time. The arrival time difference  $\Delta T$  was decomposed exactly as before by

$$\Delta T(i, j) = \Delta T_t(i) + \Delta T_z(j) + \epsilon(i, j),$$

though here  $j$  denoted hydrophone rather than set, and  $T$  meant raw arrival time rather than coherently averaged one. The residual term  $\epsilon(i, j)$  contains the  $\Delta T$  variations dependent on both time and depth, including effects of measurement error, as well as real timefront contortions.

In Fig. 20 the residual covariance  $C_{j_1, j_2}(0)$ , as previously defined, is plotted against the hydrophone separation for the two segment pairs. Variances of about  $0.5 \text{ ms}^2$  quantified the small size of the fluctuations. For comparison, Fig. 21 shows the covariance due to noise and interpolation, estimated from the simulations involving a Gaussian pulse su-

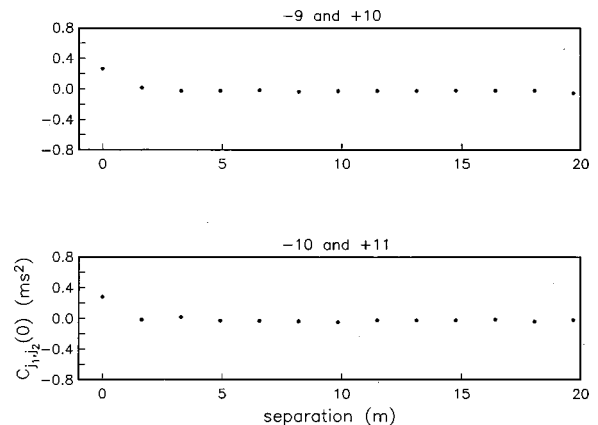


FIG. 21. Plots of the covariance due to noise and interpolation as a function of the hydrophone separation.

perimposed on actual noise. Although the differences between Figs. 20 and 21 are not outstanding, they do suggest a small signal above the measurement noise that decorrelates over about 5–10 m separation. Displays of the average (over  $j$ ) of residual autocovariance,  $C_{ave}(k)$ , show a rapid time decorrelation of these small fluctuations that is less than the 12-min sampling time.

## VII. CONCLUSION

The acoustic data from a 273-km pulse-propagation experiment have been analyzed to investigate the fluctuations in the acoustic wavefront arrival time and arrival angle. The primary purpose of the experiment was to measure the effects of ocean processes on these observables. Such measurements require subtraction of the effects of instrument motion. Although instrument-position measurement was planned for the experiment, it was not successful, requiring that the acoustic data itself be used to identify and subtract instrument motion, and thus reveal ocean-process effects. Less, but still significant, information about ocean processes is obtained by this method. The basic method for revealing ocean processes was to evaluate *relative* rather than *absolute* signal fluctuations. Some instrument motions could be inferred indirectly from the acoustics: For example, assuming a local symmetry about the vertical for timefronts sharing a common caustic, the receiving array tilt was determined as a function of time.

Comparisons between the predicted and measured geometry of timefronts in the vicinity of upper caustics have shown a fair agreement between the two. Discrepancy between them is primarily due to a simplification of the assumed soundspeed profile between source and receiver. An analysis of the arrival time differences of timefront segments having a common caustic, after beamforming over hydrophone sets of the array, has revealed: (a) a 30-m rms fluctuation in caustics' vertical positions along the array, mostly due to array motion; and (b) effects due to ocean processes: a 10-m rms relative vertical displacement between neighboring caustics, and fluctuations in travel time  $\Delta T$  of roughly 1-ms rms that have a 20–30 min decorrelation in time.

The analysis of arrival time differences for parallel neighboring segments, after beamforming over hydrophone sets of the array, has exposed ocean-process fluctuations in  $\Delta T$  and  $\Delta A$  of a few tenths of a millisecond rms and  $0.4^\circ$  rms, respectively, that also have 10–30 min decorrelation in time.

The above conclusions were based on beamforming (an amplitude-weighted coherent averaging over depth), which suppressed fluctuations associated with vertical scales much smaller than 26 m. To examine small-scale vertical behavior of timefront fluctuations, an analysis of arrival time differences for individual hydrophones was performed. Ocean-process fluctuations in  $\Delta T$  were found to be significantly smaller than 1-ms rms, with about a 5–10 m decorrelation in depth and a rapid decorrelation in time of less than 12 min.

Fluctuations that are uncorrelated over the time and depth sampling intervals are termed *broadband*.<sup>8</sup> If broadband fluctuations for *differences* between wavefront segments are assumed to have independent contributions from

each wavefront segment, then the variance of the fluctuation for each wavefront segment is half the variance corresponding to the difference. These experimental data then imply that broadband contributions for the examined wavefront segments are bounded above by 0.5-ms rms for arrival time and bounded above by  $0.3^\circ$  rms for arrival angle at this propagation range of 273 km. (For arrival angle this upper bound is primarily set by noise contributions.) All in all, *broadband* effects are very small at this range of propagation.

Further work will be necessary to evaluate the effects of various ocean processes and compare them with these data, particularly the observations of caustic movement as a whole.

## ACKNOWLEDGMENTS

This research was supported in part by the U.S. Office of Naval Research Ocean Acoustics Program. We are grateful for a grant from the W.M. Keck Foundation. This is Contribution Number 353 of the Institute of Tectonics.

- <sup>1</sup>H. A. DeFerrari and H. Nguyen, "Acoustic reciprocal transmission experiments, Florida Straits," *J. Acoust. Soc. Am.* **79**, 299–315 (1986).
- <sup>2</sup>B. M. Howe, P. F. Worcester, and R. C. Spindel, "Ocean acoustic tomography: Mesoscale velocity," *J. Geophys. Res.* **92**, 3785–3805 (1987).
- <sup>3</sup>J. L. Spiesberger and K. Metzger, "Basin-scale tomography: A new tool for studying weather and climate," *J. Geophys. Res.* **96**, 4869–4889 (1991).
- <sup>4</sup>P. F. Worcester, "Reciprocal acoustic transmission in a mid-ocean environment," *J. Acoust. Soc. Am.* **62**, 895–905 (1977).
- <sup>5</sup>W. H. Munk and C. Wunsch, "Ocean acoustic tomography: A scheme for large-scale monitoring," *Deep-Sea Res.* **26A**, 123–161 (1979).
- <sup>6</sup>M. G. Brown, W. H. Munk, J. L. Spiesberger, and P. F. Worcester, "Long-range acoustic transmission in the Northwest Atlantic," *J. Geophys. Res.* **85**, 2699–2703 (1980).
- <sup>7</sup>S. M. Flatté, J. A. Colosi, T. F. Duda, and G. Rovner, "Impulse-response analysis of ocean acoustic propagation," in *Ocean Variability and Acoustic Propagation*, edited by J. Potter and A. Warn-Varnas (Kluwer Academic, Dordrecht, 1991), pp. 161–172.
- <sup>8</sup>T. F. Duda, S. M. Flatté, J. A. Colosi, B. D. Cornuelle, J. A. Hildebrand, W. S. Hodgkiss, Jr., P. F. Worcester, B. M. Howe, J. A. Mercer, and R. C. Spindel, "Measured wavefront fluctuations in 1000-km pulse propagation in the Pacific Ocean," *J. Acoust. Soc. Am.* **92**, 939–955 (1992).
- <sup>9</sup>S. M. Flatté, R. Dashen, W. H. Munk, K. M. Watson, and F. Zachariassen, *Sound Transmission Through a Fluctuating Ocean* (Cambridge University Press, London, 1979).
- <sup>10</sup>S. M. Flatté and R. B. Stoughton, "Theory of acoustic measurement of internal wave strength as a function of depth, horizontal position, and time," *J. Geophys. Res.* **91**, 7709–7720 (1986).
- <sup>11</sup>S. M. Flatté and R. B. Stoughton, "Predictions of internal-wave effects on ocean acoustic coherence, travel-time variance, and intensity moments for very long-range propagation," *J. Acoust. Soc. Am.* **84**, 1414–1424 (1988).
- <sup>12</sup>J. A. Colosi, S. M. Flatté, and C. Bracher, "Internal-wave effects on 1000-km oceanic acoustic pulse propagation: Simulation and comparison with experiment," *J. Acoust. Soc. Am.* **96**, 452–468 (1994).
- <sup>13</sup>J. A. Colosi and S. M. Flatté, "Mode coupling by internal waves for multimegahertz acoustic propagation in the ocean," *J. Acoust. Soc. Am.* **100**, 3607–3620 (1996).
- <sup>14</sup>K. B. Smith, M. G. Brown, and F. D. Tappert, "Ray chaos in underwater acoustics," *J. Acoust. Soc. Am.* **91**, 1939–1949 (1992).
- <sup>15</sup>K. B. Smith, M. G. Brown, and F. D. Tappert, "Acoustic ray chaos induced by mesoscale ocean structure," *J. Acoust. Soc. Am.* **91**, 1950–1959 (1992).
- <sup>16</sup>J. Simmen, S. M. Flatté, and G.-Y. Wang, "Wavefront folding, chaos, and diffraction for sound propagation through ocean internal waves," *J. Acoust. Soc. Am.* **102**, 239–255 (1997).
- <sup>17</sup>W. Jobst and L. Dominijanni, "Measurements of the temporal, spatial,

- and frequency stability of an underwater acoustic channel," J. Acoust. Soc. Am. **65**, 62–69 (1979).
- <sup>18</sup>J. L. Spiesberger, R. C. Spindel, and K. Metzger, "Stability and identification of ocean acoustic multipaths," J. Acoust. Soc. Am. **67**, 2010–2017 (1980).
- <sup>19</sup>J. L. Spiesberger and P. F. Worcester, "Fluctuations of resolved acoustic multipaths at long range in the ocean," J. Acoust. Soc. Am. **70**, 565–576 (1981).
- <sup>20</sup>H. A. DeFerrari, R. I. Davis, H. Nguyen, R. F. Tusting, and N. J. Williams, "Measurements of transmission fluctuations at three ranges for refracted paths through the deep ocean," J. Acoust. Soc. Am. **74**, 1463–1488 (1983).
- <sup>21</sup>R. B. Stoughton, S. M. Flatté, and B. Howe, "Acoustic measurements of internal-wave rms displacement and rms horizontal current off Bermuda in late 1983," J. Geophys. Res. **91**, 7721–7732 (1986).
- <sup>22</sup>R. C. Sparrock, Master's thesis, University of California at San Diego (1990).
- <sup>23</sup>R. E. Williams and C. H. Wei, "Spatial and temporal fluctuations of acoustic signals propagated over long ocean paths," J. Acoust. Soc. Am. **59**, 1299–1309 (1976).
- <sup>24</sup>J. Colosi *et al.*, "Comparisons of measured and predicted acoustic fluctuations for a 3250-km propagation experiment in the eastern North Pacific Ocean," J. Acoust. Soc. Am. **105**, 3202–3218 (1999).
- <sup>25</sup>P. F. Worcester, "An example of ocean acoustic multipath identification at long range using both travel time and vertical arrival angle," J. Acoust. Soc. Am. **70**, 1743–1747 (1981).
- <sup>26</sup>P. F. Worcester, R. C. Spindel, and B. M. Howe, "Reciprocal acoustic transmissions: Instrumentation for mesoscale monitoring of ocean currents," IEEE J. Ocean Eng. **10**, 123–136 (1985).

# Mirages in shallow water matched field processing

G. L. D'Spain, J. J. Murray, and W. S. Hodgkiss

*Marine Physical Laboratory of the Scripps Institution of Oceanography, 291 Rosecrans Street,  
San Diego, California 92106*

N. O. Booth and P. W. Schey

*Code D881, Space and Naval Warfare Systems Center, San Diego, California 92152-5000*

(Received 10 November 1997; revised 19 October 1998; accepted 11 December 1998)

Broadband matched field processing (MFP)-derived estimates of 3D source location using data from hydrophone line arrays deployed in various geometries, i.e., vertical, horizontal, and tilted  $45^\circ$  from vertical, are presented in this paper. These data were collected in two shallow water (100–200 m) experiments off the coast of San Diego. Results show that estimates of source range and depth remain surprisingly coherent in the presence of large mismatch in bathymetry, but are offset from the true position by as much as 100%. The offsets are independent of array geometry. In contrast, for estimates of source azimuth, bathymetry mismatch typically causes a degradation in MFP correlation rather than an appreciable offset. However, errors in the assumed tilt from vertical of an array can lead to large offsets in the estimated source azimuth, particularly as the nominal angle of the array from vertical becomes smaller. Predictions from a simple analytical model based on adiabatic normal modes in ideal waveguides provide good fits to the broadband MFP results. © 1999 Acoustical Society of America. [S0001-4966(99)03804-7]

PACS numbers: 43.30.Wi, 43.30.Bp [SAC-B]

## INTRODUCTION

Matched field processing (MFP) is the determination of unknown parameters by the quantitative comparison of seismic acoustic field predictions, either calculated or empirical, with measurements. The unknown parameters to be determined are varied over a reasonable range of values, and estimates of the parameters are obtained where a quantitative measure of the match between the resulting predictions and measurements is greatest. Introduced into underwater acoustics in the 1970's,<sup>1</sup> it came into prominence during the 1980's when full wave field numerical propagation codes and the computer resources required to run these codes became available. It falls in the general class of inverse problems which have been addressed by geophysicists for the past several decades. The MFP approach in underwater acoustics has been used to invert independently for source location (almost exclusively for range and depth only), environmental parameters, and receiving array element positions (e.g., Refs. 2 and 3), and simultaneously for a combination of these unknowns (e.g., Ref. 4). Several excellent survey articles (e.g., Refs. 5 and 6) and books<sup>7</sup> have been written on the subject and no attempt is made here to list all the relevant literature.

Mismatch, i.e., errors in the underlying assumptions upon which the predictions (referred to as replica vectors) used in MFP are based, can have two effects. The first is a degradation in the quantitative match (e.g., correlation) between the predictions and measurements even when the predictions are based on the true values of the parameters being estimated.<sup>8</sup> The second effect is a shift in the peak of the match away from its correct location, resulting in biased estimates of the parameters. When the unknown parameters being estimated are associated with the location of an acous-

tic source, this second effect is akin to the optical phenomenon of mirages, where reflections of distant objects caused by the atmospheric temperature structure make them appear to be closer than they actually are. The focus of this work is on the second effect. It occurs when the unknown parameters being estimated are coupled to those quantities being incorrectly assumed as known in the replica vector calculations. Understanding the effects of mismatch is critical to the proper interpretation of estimates obtained from MFP inversions.

The parameters being estimated here are the range, depth, and azimuth of an underwater acoustic source. Broadband MFP processing is applied to ocean acoustic data collected by hydrophone line arrays oriented in the vertical direction (VLA), in the horizontal (HLA), and tilted at a nominal angle of  $45^\circ$  from vertical (TLA). The mismatch is associated with errors in ocean bottom bathymetry and in the assumed tilt from vertical of a line array.

Large offsets in the broadband MFP-derived source position from its true location were discovered accidentally during SWellEx-3, a shallow water experiment off the coast of San Diego in 1994. A range-independent geoacoustic model with a 200-m water column was used in the calculation of the replica vectors for a VLA in a broadband MFP inversion for the source range and depth. Although the properties of this range-independent model corresponded approximately to those along the initial part of the track, these same replica vectors were also used for the latter part of the track, even though the actual water depth decreased steadily by over 100 m. The expectation was that the MFP peak would break up, i.e., be significantly degraded, and so become randomly located in range and depth due to the severe water depth mismatch. The point at which this breakup oc-

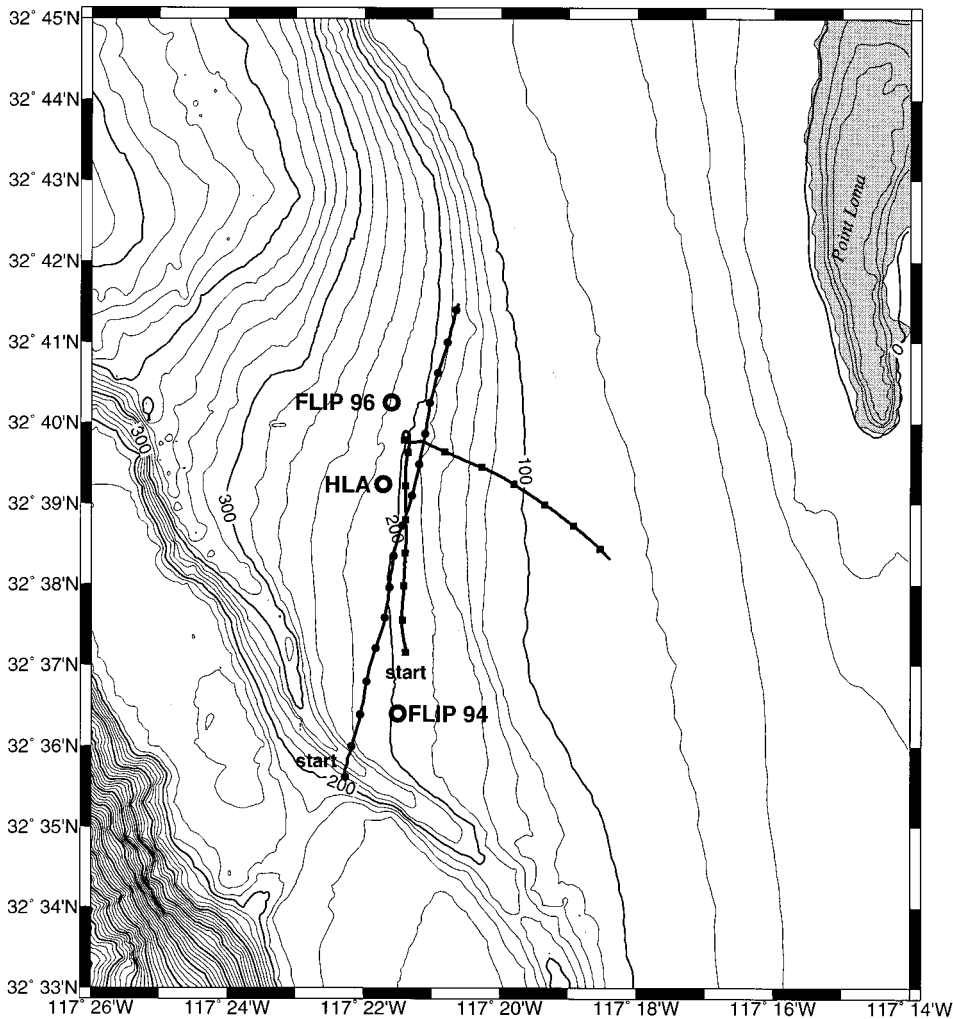


FIG. 1. Plan view of the source tow tracks for the Arc 94 and the XSlope 96 events along with bottom bathymetry contours. The contour intervals occur at 20-m increments. The location of R/P FLIP in SWellEx-3, from which the vertical array (VLA) was deployed, is indicated by the heavy circle labeled "FLIP 94." FLIP's location in SWellEx-96 was to the north, at the heavy circle labeled "FLIP 96." This location also indicates the positions of the vertical and tilted line arrays (VLA and TLA) deployed from FLIP. A heavy circle 1.9 km to the south of FLIP 96 marks the deployment position of the horizontal line array (HLA) in SWellEx-96. The tick marks along Arc 94 (solid squares) and those along XSlope 96 (solid circles) occur every 5 min.

curred would determine the azimuthal resolving capabilities of a vertical array as provided by the azimuthal dependence of the bottom bathymetry in this environment.<sup>9</sup> Rather than breaking up, the MFP peak behaved in a consistent way; it shifted progressively farther away and deeper from the actual source location as the true water depth became shallower. The broadband MFP peak remained coherent until near the end of the track, at which point the source appeared to be more than twice as deep as its actual 50-m depth, and nearly 10 km farther away than its true range of 6 km.

The purpose of this paper is to provide physically meaningful explanations for these results as well as those from the line arrays of various geometries deployed in 1996 near the same site. The first section provides a description of these two shallow water experiments and the two source tow events whose data are discussed in the paper. The method used to process the data is outlined in Sec. II and the broadband MFP results for the two events are presented in Sec. III. In Sec. IV, a simple adiabatic normal mode model is developed and its predictions are compared to the actual MFP results. The model predictions provide surprisingly good fits to the MFP results. Suggestions for useful applications and extensions of this simple model are given in Sec. V. Finally, some conclusions and implications from the work are listed in Sec. VI.

## I. DESCRIPTION OF THE EXPERIMENTS AND EVENTS

### A. The experiments

The data discussed in this paper were collected during two shallow water experiments 10–15 km west of San Diego. SWellEx-3 was conducted in the summer months of July and August of 1994, and SWellEx-96 in May 1996. These experiments are part of the set of "Shallow Water Evaluation Cell Experiments" (SWellEx) jointly conducted by the Marine Physical Laboratory (MPL) and the Naval Command, Control, and Ocean Surveillance Center, RDT&E Division (NRaD) over the past six years. Several other organizations also have participated in these experiment, notably the Naval Research Laboratory (both SWellEx-3 and SWellEx-96), Lockheed Aeronautical Systems Company (SWellEx-3), Tracor (SWellEx-3), and Science Applications International Corporation (SWellEx-3).

Figure 1 shows a map of the study site with contours of the ocean bottom bathymetry and the locations of sources and receivers in the two experiments superimposed. The site is bounded to the west by the shallow underwater ridge called the Cortes Bank and to the east by Point Loma, a peninsular land mass just to the west of downtown San Diego. The southern tip of Point Loma marks the entrance to San Diego Harbor. Most of the shipping traffic associated with the har-

bor is composed of small and medium-sized recreational boats and large Navy ships, although a modest number of commercial ships also visit the port.<sup>10</sup> The heavy circles in the figure indicate the locations of the underwater hydrophone arrays whose data are discussed in this paper. The primary SWellEx-96 position is 7 km to the north of that in SWellEx-3. The figure also shows that water depths near the site vary from 50 m to 200 m. This variable bottom bathymetry provides the opportunity to conduct source tows along tracks with various degrees of range dependence in water depth.

Several types of data acquisition systems have been deployed in the SWellEx experiments. However, most of the data presented in this paper were collected by one type of sensor system, the Marine Physical Laboratory's "SRP" arrays. These underwater acoustic receivers are large-dynamic-range, digital, hydrophone arrays. The large dynamic range is provided by 24-bit A/D converters that are located adjacent to each hydrophone element at the wet end, thereby reducing the possibility of cross-talk and other electronic self-noise. The arrays are constructed in modular fashion of from one to four interchangeable subarrays. Each subarray contains 16 low-frequency hydrophone channels (5–750 Hz) spaced at equal intervals of 1.875 m (equal to half-wavelength spacing at 400 Hz). Included in each subarray are four additional, equally spaced hydrophones whose outputs are fed into nine narrow-band filters (the filters' center frequencies occur in 0.5-kHz increments from 9.0 kHz to 13.0 kHz) to provide acoustic element localization capability. Additional details on the SRP array hardware are provided in Ref. 11.

Data from a second type of sensor system, the "AODS" array deployed in SWellEx-96, also are presented in this paper. The All Optical Deployable System (AODS) array, designed by NRI, manufactured by Litton, and deployed by NRad, is a 32-element, 240-m aperture array. Its data sampling rate is 3277 Hz, slightly more than twice the 1500-Hz rate of the SRP arrays. Although two AODS arrays were deployed in the horizontal on the ocean bottom during SWellEx-96, only the data from the array located closest to FLIP are discussed here.

In the SWellEx experiments, MPL's Floating Instrument Platform (FLIP), a 100-m-long spar buoy, has been the centralized data acquisition platform. R/P FLIP is held in a stable position by mooring it in a three-point configuration to the ocean bottom. During SWellEx-3, a 64-element SRP array was deployed in the vertical direction from FLIP and it covered the lower 120 m of the 200-m deep water column at the site. In SWellEx-96, the vertical array also spanned the lower 120 m, but the water depth extended to 216 m. A second 64-element SRP array, tilted about 45° from the vertical, was used along with a vertical SRP array. This array had an effective vertical aperture of 80 m, spanning the depths of 70 m to 150 m, and approximately the same effective horizontal aperture. It angled downward in the easterly-southeasterly direction so that its horizontal endfire directions were 112° and 292°. These two SRP arrays were cabled back to R/P FLIP, where the data were recorded continuously throughout the 2–3 week duration of the experiments.

In addition, a horizontal AODS array was deployed about 1.9 km south of FLIP. A straight-line fit to the AODS array shape is oriented at an angle of 34° to the east from true north, with a slight bow concave to the E–SE being the most significant deviation from the straight-line fit. This array was fiber-optic cabled back to shore where its data were recorded. Hereafter, the vertical arrays will be referred to as "VLA"'s the tilted array as the "TLA," and the horizontal array as the "HLA."

## B. The events

During the event from SWellEx-3 in 1994 (hereafter referred to as the "Arc 94" event), a 50-m-deep underwater acoustic source tow was conducted first along a radial line away from the VLA, and then along an arc-shaped track at constant range from the array. Figure 1 shows this source track along with the surrounding bottom bathymetry and the position of the array at "FLIP 94." Solid square tick marks are placed along the track at 5-min intervals. The radial part of the track, starting just to the north of the array and covering a distance of about 4 km, occurred over an area of nearly constant 200-m water depth. After completing this radial run out to a range of 6 km, the source ship made a 270° turn to port and headed eastward along an arc for another 4 km, passing into progressively shallower water. The towed source, a J-13 transducer deployed to a nominal depth of 50 m, generated a set of ten tones, at 53, 69, 85, 101, 117, 133, 149, 165, 181, and 197 Hz.

Also shown in Fig. 1 is the ship track for the SWellEx-96 event discussed in this paper, called the "XSlope 96" event. The projecting source was towed at a depth of 55 m, starting 8 km to the south of the array position. The tow ship maintained constant speed on a heading of 12° true, coming within 500 m to the east of the HLA and 1 km to the east of FLIP ("FLIP 96"), and finishing the event about 3 km to the northeast of FLIP. Along the track, the water became gradually shallower as the ship traversed cross-slope over the bottom bathymetry, with a more rapid decrease near the end. At the beginning of the event, the source was near broadside to the effective horizontal aperture of the TLA and passed through its endfire to the east at about the time of closest approach, whereas it was at broadside to the HLA at its nearest approach. The source transmitted 13 "pilot" tones (i.e., tones with large post-processed signal-to-noise ratio) from 49 to 388 Hz. Associated with each of the 13 pilot tones were 4 additional tones, higher in frequency by 3-Hz increments and at progressively lower signal-to-noise ratio, that were generated to examine minimum detectable level issues.<sup>12,13</sup> Only the pilot tone results are discussed in this paper.

## II. DATA PROCESSING

The time series from each of the SRP arrays' elements was windowed with a Kaiser-Bessel window of  $\alpha = 2.5$  and then Fourier transformed using an FFT length of 2048 points, resulting in a frequency resolution of 0.73 Hz (the data sampling rate was 1500 Hz). The FFT length for the HLA data was doubled so that the frequency resolution was

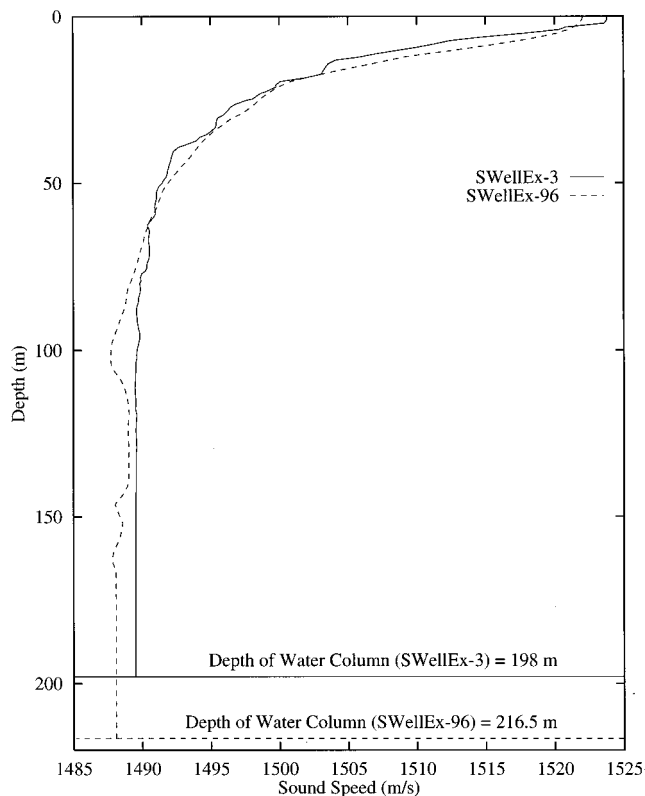


FIG. 2. The sound speed profiles and the water-column depth used in generating the range-independent replica vectors for the Arc 94 event in SWellEx-3 (solid curve) and the XSlope 96 event in SWellEx-96 (dashed curve).

approximately the same as with the SRP data. Sequential FFTs were overlapped by 50%. The time series of complex Fourier coefficients in the frequency bins corresponding to the towed source frequencies, 10 for the SWellEx-3 case and 13 for SWellEx-96, were then extracted and averaged data cross spectral matrices were formed.

The matched field replica vectors at the frequencies of interest for each of the grid points in range, depth, and azimuth were created using the normal mode program Kraken.<sup>14</sup> The range and depth grid for the computations was composed of points equally spaced by approximately 20 m in range and 2 m in depth. For azimuth, angles from array end-fire to endfire ( $0^\circ$  to  $180^\circ$ ) in 1-deg increments were covered. Range-independent geoacoustic models were used as input to

Kraken for data processing in both experiments, even though appreciable water depth changes occurred along the source tow tracks, particularly in SWellEx-3 (as shown in Fig. 1). The two sound speed profiles used in creating the replica vectors for the experiments are shown in Fig. 2. The additional geoacoustic parameters for the ocean bottom are listed in Table I.

The sediment layer at the SWellEx site is composed of undifferentiated and unconsolidated sediments and is derived from the Cretaceous sandstone and shale and Tertiary mudstone and shale sedimentary rocks in the area.<sup>15</sup> The shear wave velocity is assumed to be zero everywhere in the medium. This assumption is justified by the fact that the actual shear wave velocity in these unconsolidated sediments is very small (only 100–200 m/s) and so the effects of bottom shear on the underwater acoustic field can be accounted for by a small increase in the effective compressional attenuation of the bottom.<sup>16</sup> For both experiments, the actual water depth measured at the arrays' locations during the experiments is used in generating the replica vectors. Also, the water column sound speed was obtained from a CTD cast taken near R/P FLIP and near the time when the specific source tow events were conducted. Sediment thickness was extracted from published isopach maps, and published and unpublished seismic reflection profiles. The other geoacoustic parameters were obtained from an NRaD geoacoustic data base<sup>15</sup> that was based on sediment grain size distributions measured in the 1950's<sup>17</sup> and empirical equations.<sup>18</sup> No effort was made to improve these initial geoacoustic models nor to account for distortions in array shape, except for array tilt, in order to improve the MFP correlations presented here. (Other published work on MFP performance at the SWellEx site<sup>19</sup> has used refined geoacoustic parameters obtained from matched field inversions of data from a 1993 experiment in the same area.<sup>16</sup>)

The correlation values of the data and replica vectors at each range, depth, and azimuth grid point and at each time period, using the Bartlett processing approach (e.g., Ref. 20), were incoherently averaged across frequency. Representing the time series of the vectors of complex Fourier coefficients in the frequency bins corresponding to the  $N$  towed source frequencies as  $p(f_i, t)$ ,  $i = 1, 2, \dots, N$ , where  $N$  equals 10 for SWellEx-3 and 13 for 1996, and the replica vectors at each

TABLE I. Geoacoustic parameters for replica vector calculations.

	Arc 94, SWellEx-3	XSlope 96, SWellEx-96
Water depth (m)	198	216.5 (VLA, TLA) 213 (HLA)
Sound speed at top of sediment layer (m/s)	1554	1572.3
Sound speed at bottom of sediment layer (m/s)	1575	1593.0
Sediment layer thickness (m)	30	23.5
Sediment density (g/cc)	1.74	1.76
Sediment attenuation (dB/kmHz)	0.2	0.2
Sound speed at top of subbottom (m/s)	1861	1881
Sound speed gradient in subbottom (1/s)	1.705	1.705
Subbottom attenuation (dB/kmHz)	0.04	0.06
Subbottom density (g/cc)	2.0	2.1

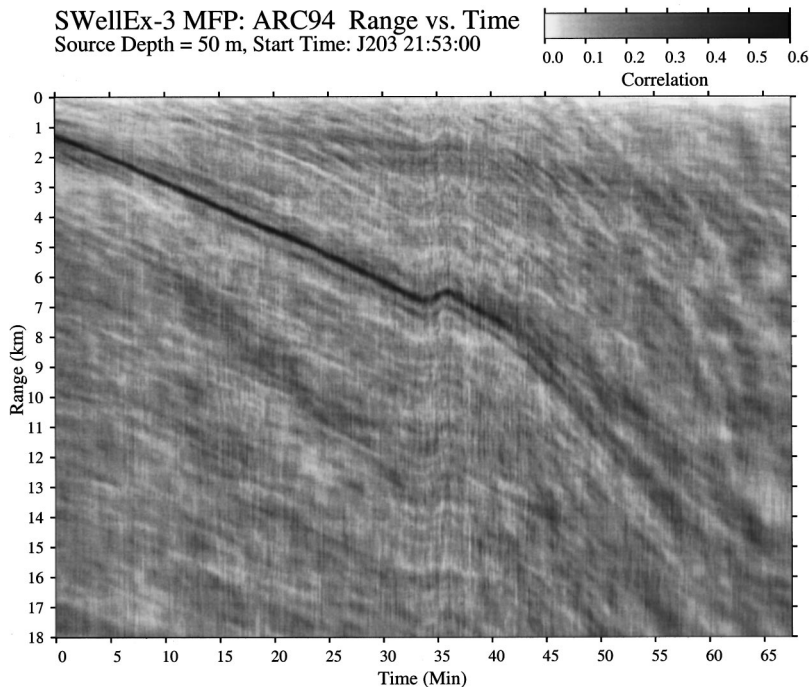


FIG. 3. Gray-scale plot of the MFP-derived range versus time during the Arc 94 event at a fixed source depth of 50 m.

of the grid points in range, depth, and azimuth as  $\hat{p}(f_i, \hat{r}, \hat{z}_s, \hat{\theta})$ ,  $i = 1, 2, \dots, N$ , then:

$$\begin{aligned} \bar{C}^2(\hat{r}, \hat{z}_s, \hat{\theta}, t) &\equiv \frac{1}{N} \sum_{i=1}^N \frac{|p(f_i, t) \hat{p}^*(f_i, \hat{r}, \hat{z}_s, \hat{\theta})|^2}{|p(f_i, t)|^2 |\hat{p}(f_i, \hat{r}, \hat{z}_s, \hat{\theta})|^2} \\ &= \frac{1}{N} \sum_{i=1}^N \frac{\hat{p}^*(f_i, \hat{r}, \hat{z}_s, \hat{\theta}) S_p(f_i, t) \hat{p}(f_i, \hat{r}, \hat{z}_s, \hat{\theta})}{|p(f_i, t)|^2 |\hat{p}(f_i, \hat{r}, \hat{z}_s, \hat{\theta})|^2}. \end{aligned} \quad (1)$$

The quantity  $S_p(f_i, t)$  represents the data cross spectral matrix, equal to the outer product of the data vector,  $p(f_i, t)$ . Finally, the maximum  $\bar{C}^2$  value over all ranges, depths, and azimuths at each time period is selected and plotted.

### III. RESULTS

In this section, the results from SWellEx-3 will be presented first, followed by those from SWellEx-96.

A gray-scale plot of the MFP-estimated range to the towed source as a function of time during the Arc 94 event is shown in Fig. 3. This plot was created by taking a slice at a fixed source depth of 50 m from each of the range-depth ambiguity surfaces for each processing period, after incoherently averaging across frequency, and then vertically stacking these slices. For the first 30–35 min of the plot, the range steadily increases from 1 to 6.5 km as the source was being towed along the radial part of the track. For this radial section, the MFP results accurately represent the true source track since the range-independent geoacoustic model used in calculating the replica vectors approximates the actual environmental parameters. Shortly after the turn on to the arc part of the track, about 35 min into the plot, the MFP track disintegrates due to the bathymetry mismatch as the source travels into progressively shallower water.

However, a quite different picture emerges if the source depth is allowed to vary. That is, instead of extracting the range-depth ambiguity slices at a fixed source depth, Fig. 4 was obtained by extracting the slices at that source depth with the maximum MFP correlation value. The MFP-derived track, rather than breaking up after the turn 35 min into the event, now indicates that the range to the source continued to increase, even though the true range to the source remained nearly constant along the arc part of the track. This remarkable effect is illustrated more clearly in Fig. 5. The range of the maximum MFP correlation, at the source depth with the maximum correlation, is plotted as a function of time (as individual diamonds) along with the true source range (plotted as a line of short dashes). The other two lines in the figure (plotted as a line of dots and as a line of long dashes) that follow approximately the MFP range estimates are predictions from the simple model and will be discussed in the next section. As the water depth under the source becomes shallower, decreasing by over 100 m near the end of the event, the source appears to the MFP processor to continue to increase in range, exceeding the true range by as much as 10 km.

Because of the surprising MFP results, plane wave beamforming was performed with the vertical array data to determine how the vertical arrival structure changed with decreasing water depth along the arc. The results are shown in Fig. 6. A phase speed of 1500 m/s was used in the calculations. No abrupt change in arrival structure occurs after the turn onto the arc 35 min into the plot. Rather, the vertical arrival structure continues to evolve with time in the same way as if the source continued out in range along the radial part of the track.

The corresponding depth of the maximum MFP correlation, at the source range with the maximum correlation, is plotted as a function of time in Fig. 7. As the water depth



SWellEx-3 MFP: ARC94 Range vs. Time  
Peak Source Depths Used, Start Time: J203 21:53:00

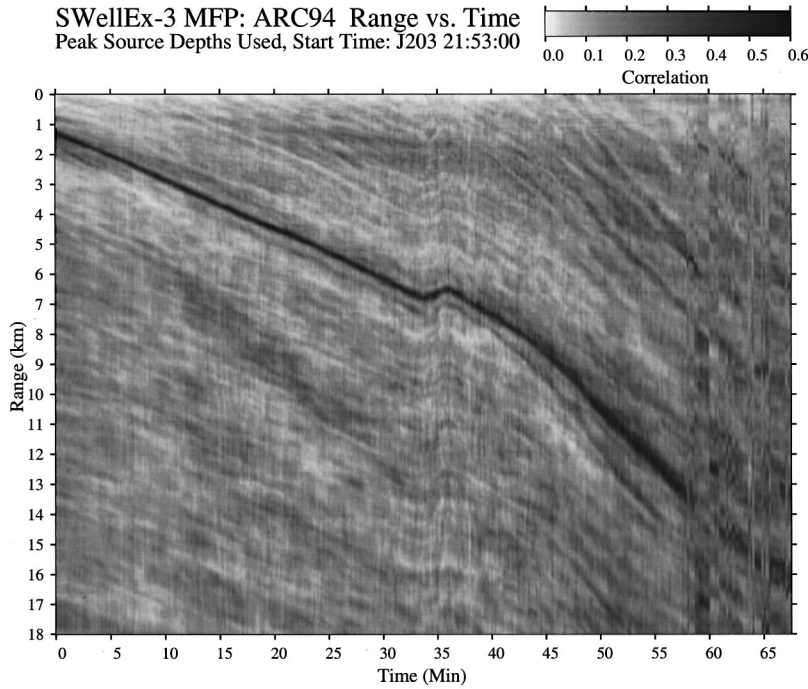


FIG. 4. Gray-scale plot of the MFP-derived range versus time along the Arc 94 track at that source depth with the peak MFP correlation.

under the source (also shown in the figure) became shallower after the turn, the MFP-derived source depth gets progressively deeper. That is, even though the true source depth remained at a nominal value of 50 m, the MFP results indicate that it descended to depths greater than 100 m because of mismatch in bottom bathymetry. The MFP peak remains surprisingly coherent in the presence of this large mismatch. The dotted line that closely follows the MFP peaks is the prediction from the simple model discussed in Sec. IV.

In addition to these results generated with replica vectors from a range-independent geoacoustic model appropriate for the radial part of the track, results for the Arc 94 event also were created using replica vectors produced by an adia-

batic normal mode code using a range-dependent environmental model as input. The range dependence of the environment was determined by the range-varying bottom bathymetry between the array and the source when the source was at a heading of  $21^\circ$  T from the array. All other geoacoustic properties were set equal to the range-independent values used previously. The time when the source was at a  $21^\circ$  T heading was 47 min after the start of the event. At this time, the water depth below the source was 115 m. The resulting MFP-derived peaks in range versus time at the peak source depth are shown in Fig. 8 and the corresponding peaks in depth versus time are presented in Fig. 9. Included on the plots are the true source range and

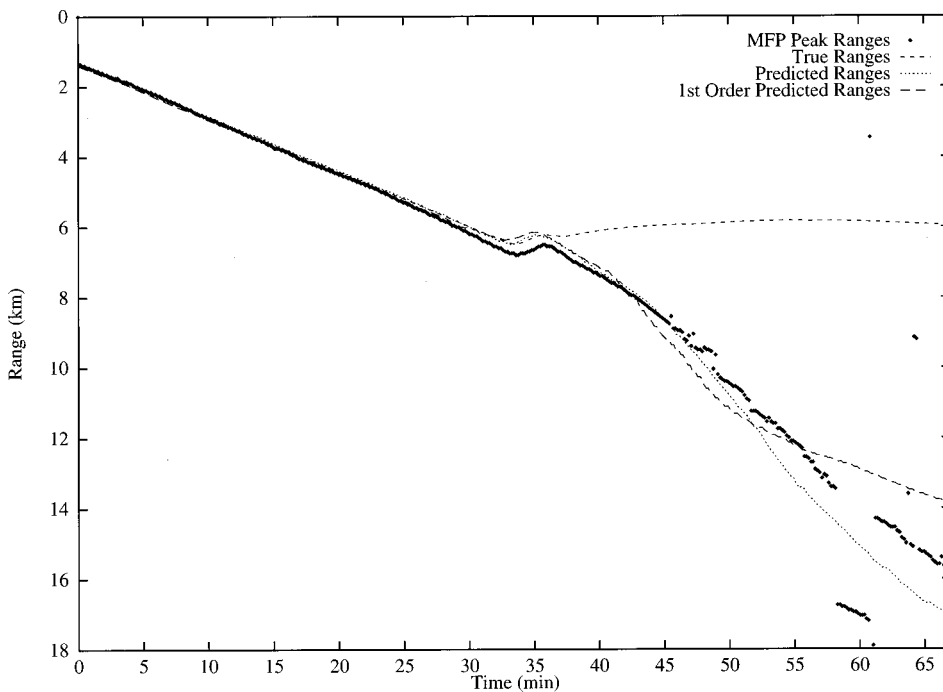


FIG. 5. MFP range-versus-time correlation peaks (plotted with diamonds) during the Arc 94 event. The true range to the source from the navigation data is shown with a medium dashed line. Also plotted are predictions from the simple analytical model; the dotted line is from Eq. (27) and the large dashed line is from Eq. (28).

SWellEx-3, ARC 94 Event, JD 203 21:53  
 Vertical Beamformer Response at 372.07 Hz (1500 m/s)

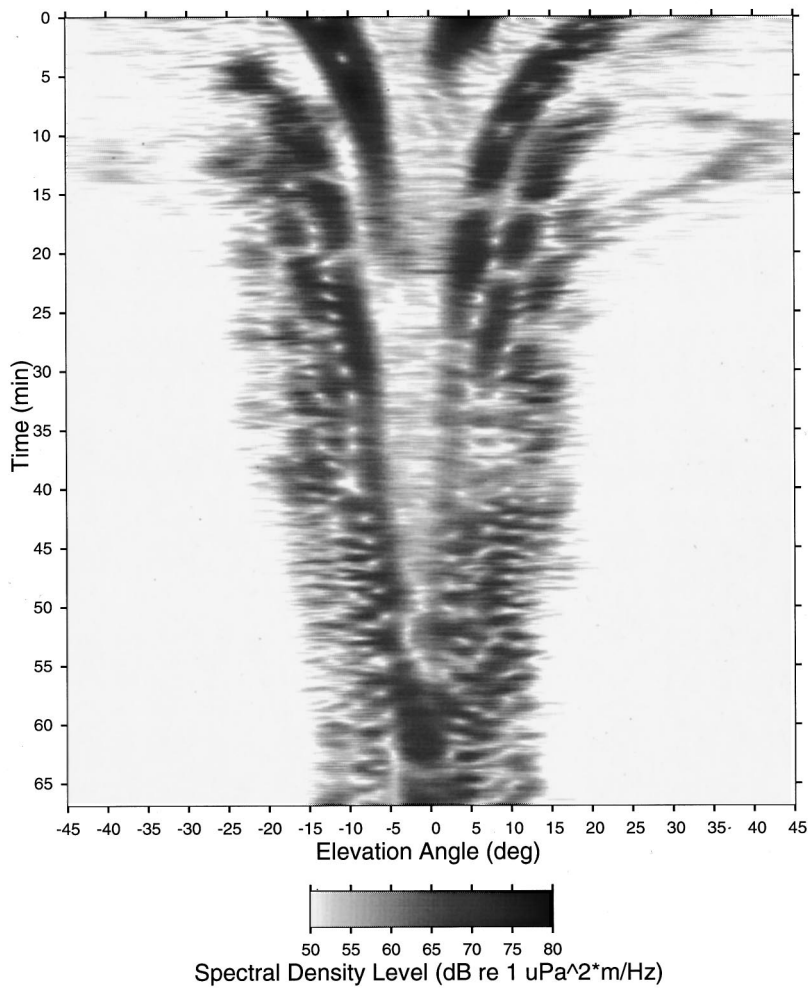


FIG. 6. Gray-scale plot of the vertical plane wave beamforming output versus time at 372 Hz along the Arc 94 track. The beamformer spectral density levels are normalized by frequency and inverse wavelength.

depth. The MFP results properly localize the source at the 47-min time period, as expected. However, before this time, the location estimates are too short in range and too shallow in depth, i.e., they are offset in the other direction from those

in Figs. 5 and 7. Again, the MFP peaks remain quite coherent, even though the bathymetry mismatch between the replica field (where the water depth becomes progressively shallower with range to a depth of 115 m at about 6 km) and the

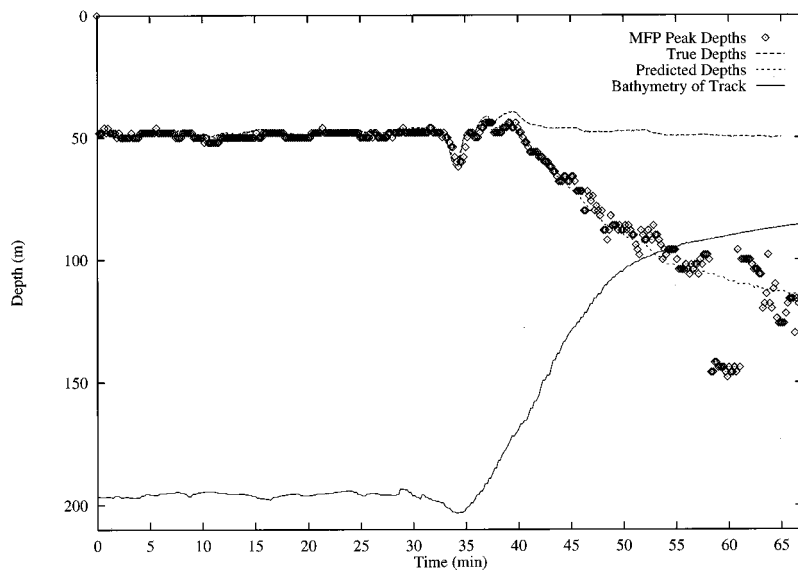


FIG. 7. MFP-derived depth-versus-time peaks (diamonds) during the Arc 94 event at the range with the peak correlation. Also shown is the true source depth (medium dashed line), the actual bottom bathymetry below the source during the event (solid line), and the simple model prediction given by Eq. (10) (dotted line).

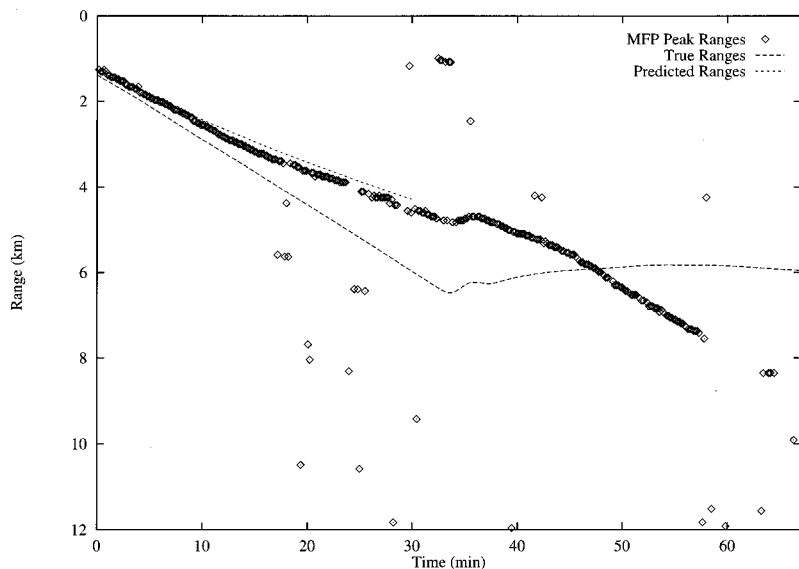


FIG. 8. MFP range-versus-time correlation peaks (diamonds) along the Arc 94 event using range-dependent bottom bathymetry in the replica vector calculations. The bottom bathymetry used was that between the source and receiver when the source was at a heading of  $21^\circ$  true from the array, which occurred 47 min into the event. The true range to the source is plotted with a dashed line and the simple model prediction, from Eq. (A5) in the Appendix, is plotted with a dotted line.

actual data (where for the first 30 min of the event the water depth is relatively range-independent at nearly 200-m depth) is significant.

The data collected in SWellEx-96 permit an evaluation of the relative performance of line arrays deployed in three different geometries, vertical, tilted at  $45^\circ$ , and horizontal. The focus here is on potential differences between these three geometries in the MFP-derived range and depth offsets due to bathymetry mismatch. Any such differences would suggest that different array geometries have different sensitivities to environmental conditions and/or use different pieces of information in the acoustic field to obtain source position estimates. In addition, for those array geometries that permit estimates of the source azimuth to be made, i.e., the TLA and the HLA, the impact of bathymetry mismatch on these estimates can be observed. SWellEx-96 results addressing these two issues now will be presented in turn.

Figures 10–12 show the correlation peaks in range at variable depth as a function of time during the XSlope 96

event for the VLA, TLA, and HLA, respectively. The true range of the source from each array also is shown. (Recall from Sec. I and Fig. 1 that the VLA and TLA were deployed from FLIP whereas the HLA was deployed 1.9 km to the south.) Because the water depth mismatch between the range-independent replica environmental model and the actual bathymetry in this event is much smaller than in Arc 94, the MFP range offsets are much smaller. However, for each of the three geometries, the offset is of the same order in distance and in the same direction, i.e., MFP overestimates the range, particularly near the end of the track. The curves that provide close fits to the MFP peaks are predictions from the simple model to be discussed in the next section.

The corresponding correlation peaks in depth as a function of time for each of the three array geometries are displayed together in Fig. 13. The bathymetry of the ocean bottom under the source is given in the lowermost curve. The depth offsets are almost identical for the three cases; the MFP depth estimates become increasingly deeper as the

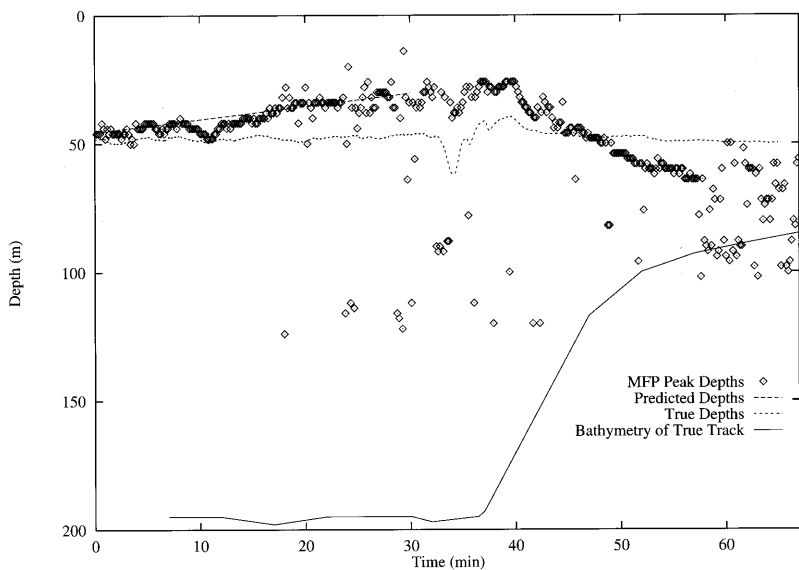


FIG. 9. The corresponding MFP depth-versus-time peaks (diamonds) using the same range-dependent replica vectors as in Fig. 8. The true source depth is plotted as a dotted line, the water depth below the source during the event is plotted with a solid line, and the model prediction, from Eqs. (A1) and (A5), is shown as a dashed line.

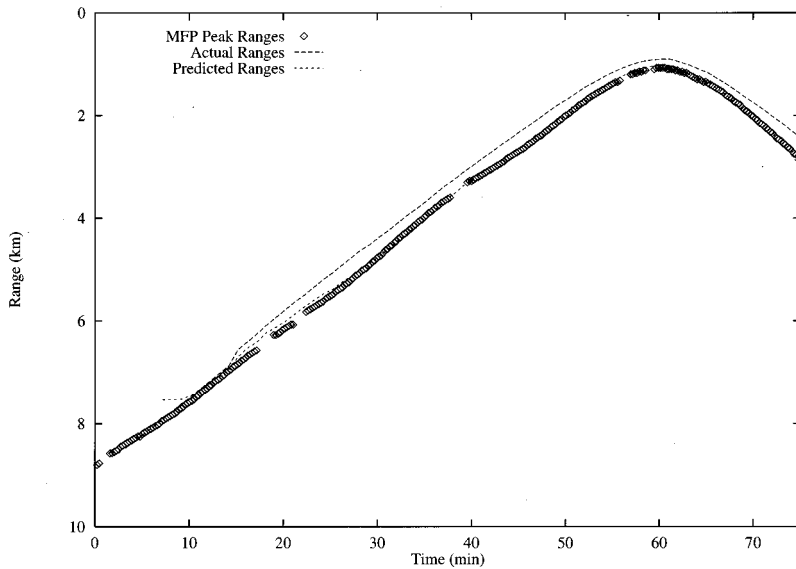


FIG. 10. Range-versus-time peaks, at the source depth with peak correlation, using the VLA data during the XSlope 96 event. The true source/receiver range is plotted with a dashed line and the model prediction using a single straight-line fit to the bathymetry [Eq. (28)] is plotted with a dotted line. The six gaps that appear in the MFP-derived track, as well as the gaps in all remaining figures, are a result of the towed source transmitting a different waveform than the 13-tone comb.

source traverses into shallower water. Note that just before the point of closest approach to the HLA, around 40 min into the event, the MFP peaks in range and depth for the HLA become scattered. The reason for this scatter is that the source also was at broadside to the HLA at this time. Because of this fact, the acoustic field varies little across the array aperture, resulting in a poor depth estimate.

Therefore, the results in these last four figures show that the MFP-derived source range and depth offsets due to bathymetry mismatch are independent of line array geometry. Turning now to the second issue of the effects of bathymetry mismatch on the estimated source bearing, the TLA and HLA source azimuth versus time results are presented in the final two figures. In Fig. 14, the boxes indicate the azimuth with the peak MFP correlation from the TLA as a function of time during the XSlope 96 event. Plotted as a dotted line is the actual source azimuth versus time with respect to this array, obtained from the navigation data. The  $180^\circ$  ambiguity in the MFP results about endfire to the horizontal aperture of the TLA, which occurs at an azimuth of

$112^\circ$  (58 min into the event), has been removed from the plot. Agreement is quite good between the MFP results and ground truth except near endfire where a bias of  $10^\circ$  to  $20^\circ$  away from endfire occurs.

Also presented in Fig. 14 are azimuth versus time results for the HLA during the XSlope 96 event along with ground truth. Both the MFP-derived azimuths (plotted as diamonds) as well as the plane wave beamformer results (plotted as plusses) are shown. The two processors follow the actual track quite well except for the breakdown in the plane wave beamformer results near endfire ( $34^\circ$ ), as expected. Also, the small offset between the plane wave and the MFP results that occurs between 40 min and 50 min, around the time of closest approach, is caused by neglect of wavefront curvature in the plane wave beamformer. The output from curved wavefront beamforming focused at a 1-km range is nearly identical to the MFP track. Notably, the MFP results transition smoothly through endfire without the significant offset seen in the TLA azimuth data. Therefore, the azimuthal bias observed in the TLA results appears to be associated with that

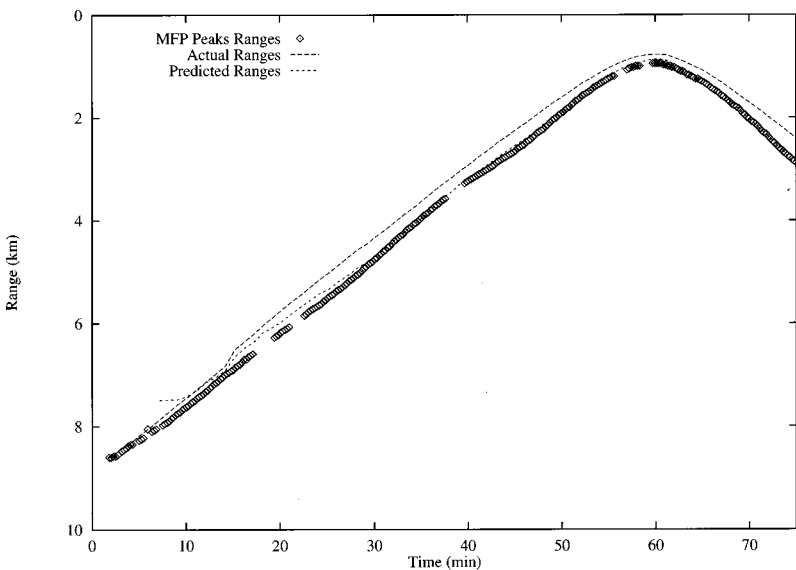


FIG. 11. The corresponding MFP range-versus-time peaks using the TLA data during the XSlope 96 event. All other aspects of this plot are the same as in Fig. 10.

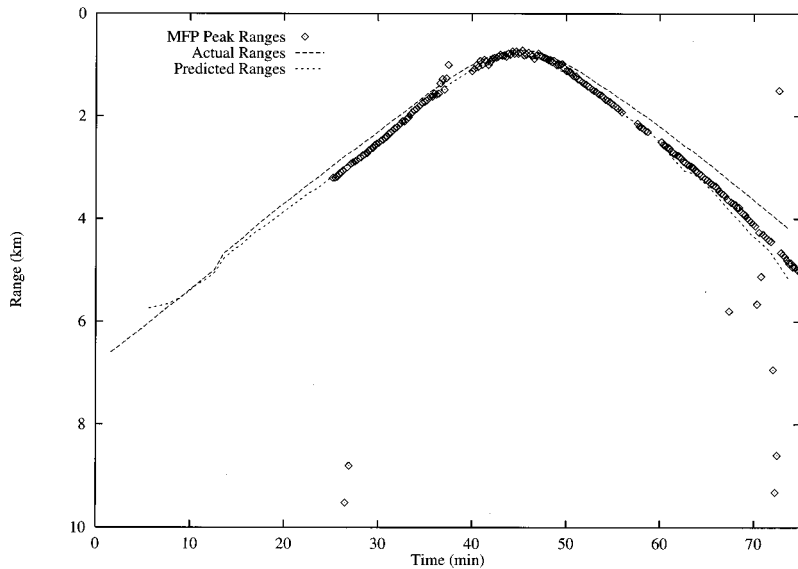


FIG. 12. The corresponding MFP range-versus-time peaks using the HLA data during the XSlope 96 event. All other aspects of this plot are the same as in Figs. 10 and 11. Note that the HLA was not recording data until 25 min after the start of the event.

array geometry and not generally with bathymetry mismatch.

To examine the azimuthal bias of the TLA near endfire in greater detail, the 7-min period from 54 min to 61 min into the event encompassing the passage through endfire was reprocessed in various ways. Figure 15 shows the effects of changes in the range-independent environmental model used to calculate the replica vectors. For reference, the MFP data points from the TLA in Fig. 14 (boxes) are plotted again, this time with diamonds. The results from both decreasing and increasing the water depth by 5 m are plotted with plus signs and boxes, respectively, and those from using another sound speed profile taken at a different time on the same day are shown as “×”s. The two data gaps centered at 56 min and 59 min are the result of the towed source broadcasting a different signal than the 13-tone comb for these short periods. Clearly, none of these changes in the environment, including those that change the amount of bathymetry mismatch, have an appreciable effect on the azimuthal bias near endfire. However, for all these runs, the assumed tilt of the

array was fixed at  $47^\circ$  from vertical. Additional runs with a  $47^\circ$  tilt were made where the duration of the data used to estimate the data cross spectral matrix was changed (from 5.5 s to 62.8 s), the FFT length was changed (from 2048 to 8192), and variations were made in the number of modes at each frequency that were used in the replica field calculation. The azimuthal results (as well as those for source range and depth) using these various processing parameters show no significant differences from those for a  $47^\circ$  array tilt in Fig. 15. Rather, the explanation for the azimuthal offset near endfire is that mismatch exists in the tilt of the TLA rather than because of bathymetry mismatch or other possible effects. Plotted in Fig. 15 are the MFP azimuthal results for an assumed TLA tilt of  $44^\circ$  (asterisks) which shows much closer agreement with the actual source azimuth. Further discussion of this topic is delayed until the next section where a prediction of the amount of azimuthal bias due to errors in array tilt is derived.

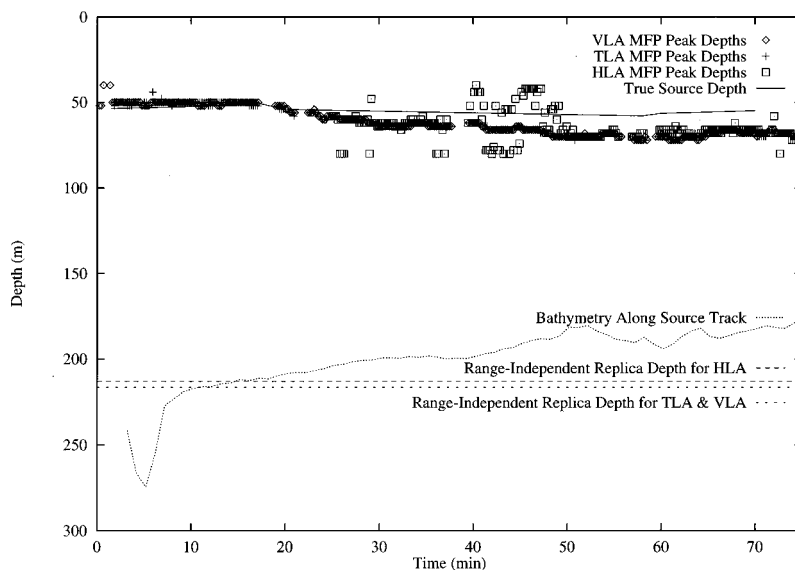


FIG. 13. The MFP depth-versus-time peaks, at the source range with peak correlation, during the XSlope 96 event from the VLA (diamonds), the TLA (plusses), and HLA (squares) data. The actual water depth under the source during the event (dotted line) and the range-independent water depth used in the replica vector calculations for the VLA and TLA (large-dashed line) and for the HLA (small-dashed line) also are shown. In addition, the true depth of the source measured by a depth gauge is plotted as a solid curve.

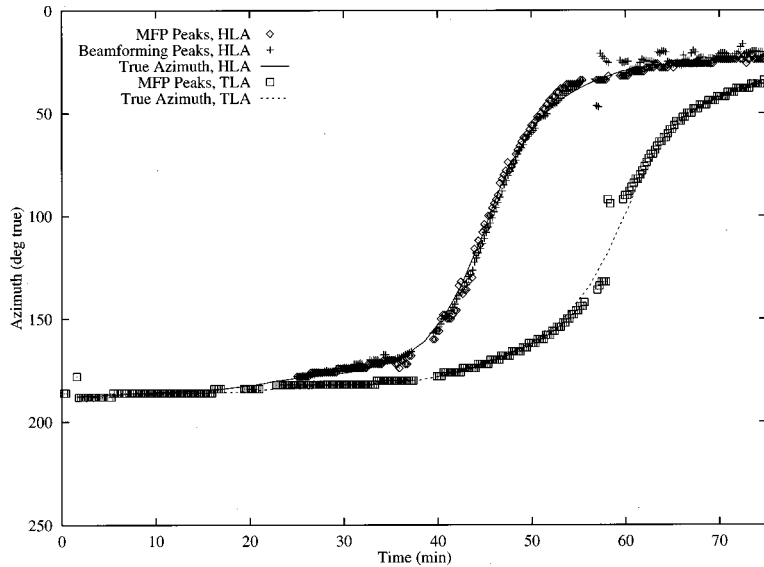


FIG. 14. The azimuth-versus-time results estimated from the TLA and the HLA data during the XSlope 96 event. Those for the TLA, plotted as boxes, were derived using MFP with an assumed tilt from vertical of  $47^\circ$  in the replica vector calculations. For comparison, the true azimuth from this array to the source is shown as a dotted curve. The results from the HLA data of using both MFP (diamonds) and plane wave beamforming (pluses) are shown along with the actual bearing to the source (solid curve).

#### IV. SIMPLE MODEL

##### A. Adiabatic normal mode model development

To explain the offset of the MFP-derived source location from the true location with VLA data in SWellEx-3, a simple analytical model based on normal modes in an ideal waveguide was developed.<sup>21</sup> This model explains why a reduction in range-depth ambiguity is achieved by averaging across frequency, even in the presence of some forms of mismatch. It is based on matching a “true” field with a replica vector field, where both true and replica fields are calculated using adiabatic mode theory in homogeneous, ideal waveguides bounded by pressure-release surfaces on both top and bottom. The densities and water sound speeds in the two model waveguides also are identical; the only difference (mismatch) is in the range dependence of the waveguide depths. The development of this model, extended to the MFP estimation of source azimuth with tilted and horizontal arrays, will now be presented.

Consider a range-independent waveguide with a constant sound speed,  $c$ , bounded by pressure-release surfaces on top, at  $z=0$ , and bottom at depth  $z=D$ . The pressure field measured by the  $q$ th omnidirectional receiver at depth  $z_r^q$  due to a single-tone point source in such a waveguide at a range of  $\hat{r}_q(\hat{\theta})$  and depth  $\hat{z}_s$  can be written as:

$$\hat{p}(\hat{r}_q(\hat{\theta}), z_r^q, \hat{z}_s) = \sum_{n=0}^{N-1} \frac{\hat{A}}{[\hat{k}_n \hat{r}_q(\hat{\theta})]^{1/2}} \hat{\psi}_n(\hat{z}_s) \psi_n(z_r^q) \times \exp[i\hat{k}_n \hat{r}_q(\hat{\theta})], \quad (2)$$

where the normal modes (eigenfunctions) are:

$$\hat{\psi}_n(z) = \sin(\hat{\gamma}_n z) \quad (3)$$

and the corresponding vertical wave number (eigenvalue) for mode  $n$  is:

$$\hat{\gamma}_n = \frac{n\pi}{D}. \quad (4)$$

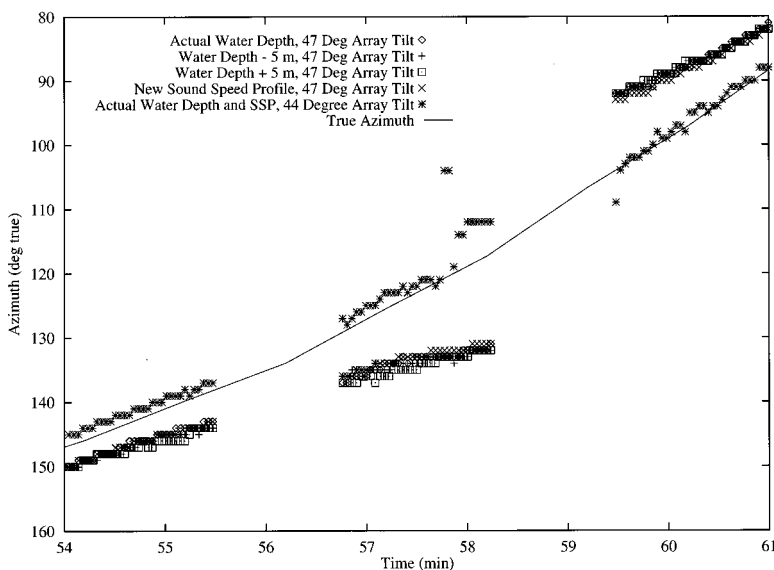


FIG. 15. An expanded view of the MFP-derived azimuth versus time using the TLA data for a 7-min period around when the source passed through array endfire during the XSlope 96 event. The data from Fig. 14, where the range-independent water depth used in the replica vector calculations was equal to that measured at the array location and the array tilt was assumed to be  $47^\circ$ , are replotted with diamonds. The MFP results obtained after decreasing and increasing by 5 m the range-independent water depth used in the replica vector calculations, but keeping the assumed tilt at  $47^\circ$ , are plotted with pluses and squares, respectively. Additional results for an assumed  $47^\circ$  tilt using a different water sound speed profile taken at another time during the experiment along with the actual water depth measured at the array position are plotted with “ $\times$ ”’s. All of these results differ little from one another. In contrast, the MFP results of using an array tilt of  $44^\circ$ , and the actual water depth at the array, are plotted with asterisks and show much better agreement with the true bearing to the source, shown as the solid curve.

The mode horizontal wave number is given by  $\hat{k}_n = (k^2 - \hat{\gamma}_n^2)^{1/2}$ . The source–receiver range for the  $q$ th element is:

$$\hat{r}_q(\hat{\theta}) = \hat{r}_o + \Delta r_q \cos(\hat{\theta}), \quad (5)$$

with  $\hat{\theta}$  representing the angle of the source from endfire of the horizontal projection of the tilted receiving array,  $\hat{r}_o$  the source–receiver range to a reference element in the array, and  $\Delta r_q$  being the horizontal distance of the  $q$ th element from the reference element. The expression given in Eq. (5) assumes that the source is in the array far field so that  $\hat{\theta}$  is constant along the array. Equation (2) is used to calculate the replica vectors for a tilted array. The hat (“^”) symbol is used to indicate the properties to be estimated or modeled in the MFP calculation. The term  $\hat{A}$  represents a complex frequency-dependent amplitude term that is independent of mode number, source depth, azimuth, and range.

In weakly range-dependent environments, the coupling of energy between modes can be ignored and the “adiabatic” mode approximation can be used. In this approximation, the assumed range-independent normal mode eigenfunction evaluated at the presumed source depth,  $\hat{\psi}_n(\hat{z}_s)$ , is replaced by the actual modal eigenfunction at the source location,  $\psi_n^s(z_s) = \sin(\gamma_n^s z_s)$ , where  $z_s$  is the true source depth. The true mode vertical wave number at the source location,  $\gamma_n^s$ , is given by Eq. (4) with  $D$  replaced by  $d_s$ , the true water depth at the source location. Also,  $\hat{k}_n \hat{r}_q$  is replaced by  $\int_0^{R_q} k_n dr$  with  $R_q(\theta) = R_o + \Delta r_q \cos(\theta)$ . The source azimuth,  $\theta$ , is the actual horizontal arrival angle as perceived by the receiving array, i.e., it may include effects of horizontal refraction. In the following, the notation  $\int k_n dr$  will be understood to represent the definite integral  $\int_0^{R_q} k_n dr$ . Equation (2) then becomes:

$$p(R_q, z_r^q, z_s) = \sum_{n=0}^{N-1} \frac{A}{(\int k_n dr)^{1/2}} \psi_n^s(z_s) \psi_n(z_r^q) \times \exp\left[i \int k_n dr\right]. \quad (6)$$

Equation (6) is used to calculate the “true” pressure field for predicting the MFP results.

The matched field output to be maximized by the appropriate selection of  $\hat{r}$ ,  $\hat{z}_s$ , and  $\hat{\theta}$  is an amplitude squared and normalized version of:<sup>20</sup>

$$\text{MFP output} \equiv \sum_q p(R_q, z_r^q, z_s) \hat{p}^*(\hat{r}_q, z_r^q, \hat{z}_s), \quad (7)$$

so that, matching the replica vectors calculated using Eq. (2) with the “true” pressure field given by Eq. (6) yields:

$$\begin{aligned} & \sum_q p(R_q(\theta), z_r^q, z_s) \hat{p}^*(\hat{r}_q(\hat{\theta}), z_r^q, \hat{z}_s) \\ &= \sum_q \sum_{n=0}^{N-1} \frac{A \hat{A}}{(\hat{k}_n \hat{r}_q \int k_n dr)^{1/2}} [\psi_n^s(z_s) \hat{\psi}_n(\hat{z}_s)] \psi_n^2(z_r^q) \\ & \quad \times \exp\left(i \int k_n dr - i \hat{k}_n \hat{r}_q\right) + \text{cross terms.} \end{aligned} \quad (8)$$

Equation (8) shows that matching in depth involves the properties of the medium only at the source and receiver locations, whereas matching in range involves an integrated effect over the whole range between source and receiver. As will be shown below, the match in azimuth depends only on the properties at the receiver location. With the model used here, the modal vertical wave numbers are a function only of water depth and do not depend on frequency.

Upon examination of the term inside the brackets in Eq. (8), it is immediately obvious that to match in depth, the requirement is:

$$\gamma_n^s z_s = \hat{\gamma}_n \hat{z}_s, \quad (9)$$

or, after plugging in for the expressions for the vertical wave numbers:

$$\hat{z}_s = \frac{D}{d_s} z_s. \quad (10)$$

Note that the expression in Eq. (10) is independent of mode number and frequency, so that it represents the best match in depth for all modes and frequencies.

Determining the estimated source range (and from that the source azimuth) that provides the best match between true and replica fields is not as straightforward as it was with depth. The reason has to do with the differences in the type of information in the acoustic field that the matched field processor uses to obtain source range and depth. Whereas the prediction of the source depth is obtained by matching true and replica fields mode-amplitude-by-mode-amplitude over depth, the range estimate is based on matching modal spatial interference patterns. Therefore, the cross terms in Eq. (8) contain the information required to estimate range. To account for these modal interaction terms, a specified reference mode (whose quantities are indicated by a subscript “ $l$ ”), can be factored out of the the range-independent and adiabatic normal mode expressions in Eqs. (2) and (6). Ignoring cylindrical spreading and focusing on those terms that depend on range, the MFP output is then:

$$\begin{aligned} \sum_q p(R_q) \hat{p}^*(\hat{r}_q) &= \sum_q B_l \hat{B}_l \exp\left[i \int k_l dr - i \hat{k}_l \hat{r}_q\right] \\ & \quad \times \left\{ 1 + \sum_{n=1}^{N-1} \frac{B_n}{B_l} \exp\left[i \int \Delta k_{n,l} dr\right] \right. \\ & \quad + \sum_{n=1}^{N-1} \frac{\hat{B}_n}{\hat{B}_l} \exp[-i \Delta \hat{k}_{n,l} \hat{r}_q] \\ & \quad + \left. \left( \sum_{n=1}^{N-1} \frac{B_n}{B_l} \exp\left[i \int \Delta k_{n,l} dr\right] \right) \right. \\ & \quad \times \left. \left( \sum_{n=1}^{N-1} \frac{\hat{B}_n}{\hat{B}_l} \exp[-i \Delta \hat{k}_{n,l} \hat{r}_q] \right) \right\}. \end{aligned} \quad (11)$$

The wave number difference quantities are defined as  $\Delta k_{n,l} \equiv k_n - k_l$  and  $\Delta \hat{k}_{n,l} \equiv \hat{k}_n - \hat{k}_l$ . Also, the terms in the mode expansion that are independent of range have been lumped together into mode-dependent amplitudes,  $B_n$  and  $\hat{B}_n$ . In maximizing the MFP output with respect to range, the first three terms inside the braces in Eq. (11) can be ignored.

Expanding the fourth term involving the product of two sums then:

$$\sum_q p(R_q) \hat{p}^*(\hat{r}_q) = \sum_q \exp \left[ i \int k_l dr - i \hat{k}_l \hat{r}_q \right] \times \left[ \sum_{n=1}^{N-1} B_n \hat{B}_n \exp \left[ i \int \Delta k_{n,l} dr - i \Delta \hat{k}_{n,l} \hat{r}_q \right] + \text{cross terms} \right], \quad (12)$$

where the cross terms have the form:

$$B_j \hat{B}_n \exp \left( i \int \Delta k_{j,l} dr - i \Delta \hat{k}_{n,l} \hat{r}_q \right) + B_n \hat{B}_j \exp \left( i \int \Delta k_{n,l} dr - i \Delta \hat{k}_{j,l} \hat{r}_q \right). \quad (13)$$

From Eqs. (12) and (13), the predicted MFP range is determined by the simultaneous match of the following two equations:

$$\int \Delta k_{n,l} dr = \Delta \hat{k}_{n,l} \hat{r}_q, \quad (14a)$$

$$\int k_l dr = \hat{k}_l \hat{r}_q + h 2 \pi, \quad (14b)$$

where  $h$  is an integer.

To obtain an expression for the predicted MFP azimuth, both sides of the expressions in Eqs. (14a) and (14b) are divided into the sum of two parts; one pertaining to the range between the source and the reference receiver (from 0 to  $R_o$ ) and the second associated with the horizontal distance across the tilted array aperture (from  $R_o$  to  $R_q$ ). Using the expressions for  $\hat{r}_q(\hat{\theta})$  and  $R_q(\theta)$  given earlier, then:

$$\int_0^{R_o} \Delta k_{n,l} dr + \int_{R_o}^{R_q} \Delta k_{n,l} dr = \Delta \hat{k}_{n,l} \hat{r}_o + \Delta \hat{k}_{n,l} \Delta r_q \cos(\hat{\theta}), \quad (15a)$$

$$\int_0^{R_o} k_l dr + \int_{R_o}^{R_q} k_l dr = (\hat{k}_l \hat{r}_o + h 2 \pi) + \hat{k}_l \Delta r_q \cos(\hat{\theta}). \quad (15b)$$

These equations can be satisfied by setting

$$\int_0^{R_o} \Delta k_{n,l} dr = \Delta \hat{k}_{n,l} \hat{r}_o \quad (16)$$

and

$$\int_{R_o}^{R_q} k_l dr = \hat{k}_l \Delta r_q \cos(\hat{\theta}). \quad (17)$$

This approach results in the separation of the MFP estimate for source range and source azimuth. In doing so, the integer  $h$  disappears. The reason is that the terms,  $\int_0^{R_o} k_l dr$  and  $\hat{k}_l \hat{r}_o + h 2 \pi$  [re: the left-hand terms on both sides of Eq. (15b)], now are independent of array element number,  $q$ . Therefore, the exponentials involving these terms can be taken outside the sum over  $q$  in Eq. (12) for the MFP output, where they disappear because only the amplitude of the out-

put is relevant; absolute phase information is discarded. An ambiguity of  $2\pi$  does not exist for the range integral from  $R_o$  to  $R_q$  in Eq. (17) since the interelement spacing of the tilted array is presumed to be sufficient to prevent spatial aliasing. Therefore, Eq. (17) can be rearranged to give:

$$\cos(\hat{\theta}) = \frac{1}{\hat{k}_n \Delta r_q} \int_{R_o}^{R_q} k_n dr, \quad (18)$$

where the upper limit of the integration is  $R_q(\theta) = R_o + \Delta r_q \cos(\theta)$ . Equation (18) shows that any potential MFP azimuthal bias is due solely to mismatch across the horizontal aperture of the array. If the range dependence of the mode wave numbers across this aperture is negligible, then the MFP-derived estimate of azimuth becomes simply:

$$\cos(\hat{\theta}) = \frac{k_n}{\hat{k}_n} \cos(\theta). \quad (19)$$

This equation is an expression of Snell's law. Using a Taylor series expansion and keeping the first term (to the order of the cube of the ratio of the vertical wave number to the medium wave number), then

$$\cos(\hat{\theta}) = \left( 1 - \frac{1}{2} \frac{1}{k^2} (\gamma_n^2 - \hat{\gamma}_n^2) \right) \cos(\theta). \quad (20)$$

The other part of the separation that gives the MFP source range estimate, i.e., Eq. (16), can be rewritten as:

$$\hat{r}_o = \frac{\int \Delta k_{n,l} dr}{\Delta \hat{k}_{n,l}}, \quad (21)$$

where the integral now is over the range from the source to the reference array element.

To proceed further, the integral over range of wave number differences given in Eq. (21) must be evaluated. Rewriting Eq. (5) for the mode wave numbers as  $k_n = k(1 - (\gamma_n^2/k^2))^{1/2}$  and using a Taylor series expansion gives:

$$\int \Delta k_{n,l} dr = -\frac{1}{2} \frac{1}{k} \int (\gamma_n^2 - \gamma_l^2) dr - \frac{1}{8} \frac{1}{k^3} \times \int (\gamma_n^4 - \gamma_l^4) dr - \dots - \frac{\Pi |2(m-1)-1|}{2^m m!} \frac{1}{k^{2m-1}} \times \int (\gamma_n^{2m} - \gamma_l^{2m}) dr - \dots. \quad (22)$$

Plugging in the expression for the vertical wave numbers from Eq. (4) and letting  $d(r)$  represent the range dependence of the water depth, then the  $m$ th term of Eq. (22), where  $m = 1, 2, 3, \dots$ , is:

$$\frac{\Pi |2(m-1)-1|}{2^m m!} \frac{\pi^{2m}}{k^{2m-1}} (l^{2m} - n^{2m}) \int \frac{1}{d(r)^{2m}} dr. \quad (23)$$



The integral in Eq. (23) can be evaluated easily if the bottom bathymetry between source and receiver is modeled by connected straight-line segments. Each of the  $M$  segments is defined by a starting range,  $x_{i-1}$ , and ending range,  $x_i$ , so that the true range,  $R_o$ , is:

$$R_o = \sum_{i=1}^M (x_i - x_{i-1}). \quad (24)$$

The water depths at the starting and ending ranges of segment  $i$  are  $d_{i-1}$  and  $d_i$ , respectively. Letting the bottom slope within segment  $i$  be:

$$\alpha_i \equiv \frac{d_i - d_{i-1}}{x_i - x_{i-1}} \quad (25)$$

so that the range-dependent bottom depth within the segment is  $d(r) = d_{i-1} + \alpha_i(r - x_{i-1})$ , then:

$$\int \frac{1}{d(r)^{2m}} dr = \sum_{i=1}^M \frac{1}{\alpha_i} \frac{1}{2m-1} \times \left( \frac{1}{d_{i-1}^{2m-1}} - \frac{1}{(d_{i-1} + \alpha_i(x_i - x_{i-1}))^{2m-1}} \right). \quad (26)$$

For a segment where  $\alpha_i = 0$ , the integral over that segment equals  $(x_i - x_{i-1})/d_{i-1}^{2m}$ , an expression which either can be determined directly from the left side of Eq. (26) or from the right side by taking the limit as  $\alpha_i$  approaches 0.

The expressions in Eqs. (21), (22), (23), and (26) can be combined to give the final result for the apparent source range from this MFP model. Keeping just the first term in the expansion gives:

$$\hat{r}_o = \frac{1}{\hat{k}_n - \hat{k}_l} \int_0^{R_o} (k_n - k_l) dr \approx D^2 \int_0^{R_o} \frac{1}{d(r)^2} dr \approx D^2 \sum_{i=1}^M \frac{x_i - x_{i-1}}{d_i d_{i-1}}, \quad (27)$$

which is valid to order  $(\gamma_n/k)^3$ . As with the match in depth given in Eq. (10), the expression in Eq. (27) is independent of mode number so that it satisfies all mode differences simultaneously and is independent of frequency. However, the next higher order term in the expansion, to order of  $(\gamma_n/k)^5$ , is dependent on mode number squared and inverse frequency squared. It therefore predicts the ultimate breakup of the MFP peak. However, the natural spatial filtering of the waveguide (because of the stronger interaction of the higher order modes with the attenuating ocean bottom) tends to counteract its influence.

A special case is that of a bottom bathymetry model composed of just one segment. In this case, Eq. (27) becomes:

$$\hat{r}_o = \frac{D}{d_s} R_o, \quad (28)$$

which has exactly the same form as Eq. (10) for the apparent source depth.

Returning to the expression for the MFP-derived azimuth in Eqs. (19) and (20), a mismatch in depth at the receiver location can be introduced by letting the true depth be  $D + \Delta d$ . Then:

$$\begin{aligned} \cos(\hat{\theta}) &\approx \left[ 1 - \frac{1}{2} \left( \frac{n\pi}{kD} \right)^2 \left( \frac{1}{(1 + \Delta d/D)^2} - 1 \right) \right] \cos(\theta) \\ &\approx \left[ 1 + \left( \frac{n\pi}{kD} \right)^2 \frac{\Delta d}{D} \right] \cos(\theta). \end{aligned} \quad (29)$$

The MFP azimuthal bias is dependent on mode number squared and inverse frequency squared (through  $k$ ). Therefore, bathymetry mismatch at the receiver location typically results in degradation in broadband MFP correlation rather than a bias offset as with source range and depth. As mentioned previously, mismatch in bathymetry at other places along the source-to-receiver path are predicted to have no effect on the MFP-derived azimuth. This statement does not take into account horizontal refraction effects, which have been ignored in this development.

However, a source of mismatch does exist that leads to significant offsets in the MFP-derived azimuth without appreciable degradation in correlation. It is the error in the assumed tilt from vertical of an array. To show this, note first that, to this point in the development, the quantity,  $\Delta r_q$ , representing the horizontal distance of array element  $q$  from the reference element, has been taken as a known parameter. If it is known inexactly, and the assumed value is indicated as  $\Delta \hat{r}_q$ , then Eq. (18) gives in the case of range independence across the tilted array aperture:

$$\cos(\hat{\theta}) = \frac{k_n \Delta r_q}{\hat{k}_n \Delta \hat{r}_q} \cos(\theta). \quad (30)$$

As pointed out before, the ratio between  $k_n$  and  $\hat{k}_n$  is dependent upon mode number squared and inverse frequency squared and thus causes, in general, a degradation in correlation. If the environmental mismatch over the receiving array aperture is negligible, then:

$$\cos(\hat{\theta}) = \frac{\Delta r_q}{\Delta \hat{r}_q} \cos(\theta). \quad (31)$$

Nonlinear distortions in the array shape, e.g., a catenary, in general result in mismatch that depends upon the position of the array element along the cable, leading to MFP correlation degradation. However, because of the large tensions applied along the array cables in the SWellEx deployments, these types of distortions are of second order in importance and will not be considered further here. In contrast, small errors in the assumed tilt of the array from vertical can lead to large estimated source position offsets. To consider this effect, let  $a$  represent the interelement spacing of an equally spaced line array and  $\tau$  the true array tilt from vertical. Then:

$$\Delta r_q = qa \sin(\tau). \quad (32)$$

If  $\hat{\tau}$  is the assumed array tilt, then Eq. (31) gives:

$$\cos(\hat{\theta}) = \frac{\sin(\tau)}{\sin(\hat{\tau})} \cos(\theta). \quad (33)$$

This expression for the offset in the apparent azimuth of the source from its true azimuth is independent of mode number, so that it satisfies all modes simultaneously, and is independent of frequency, so that broadband processing gain is maintained. Therefore, mismatch in array tilt has the same effect on source azimuth as does bottom bathymetry mismatch on the source range and depth.

Equation (33) indicates that the effect of tilt mismatch on the azimuthal offset is greatest when  $\cos(\theta)$  is greatest, i.e., when the source is near endfire to the horizontal aperture. It also shows that a given amount of mismatch has an increasing effect as the nominal tilt from vertical decreases. That is, if the tilt mismatch is a small value, say  $\epsilon$ , then Eq. (33) approximately equals:

$$\cos(\hat{\theta}) \approx \left[ 1 + \epsilon \frac{1}{\tan(\hat{\tau})} \right] \cos(\theta). \quad (34)$$

Since the tangent function is a monotonically increasing function over the interval from  $0^\circ$  to  $90^\circ$ , then the contribution of the mismatch term inside the brackets in Eq. (34) decreases as the tilt increases. This result shows that errors in tilt have negligible effect on the source azimuth obtained from a horizontal line array, whereas an array oriented at an appreciable angle away from the horizontal, such as the TLA deployed at  $45^\circ$  in SWellEx-96, can suffer significant azimuthal offsets. Obviously, a line array deployed near the vertical cannot provide reliable azimuthal estimates.

Several sets of simulation runs were conducted to provide verification of the basic predictions of this simple model. They show that bathymetry mismatch causes MFP-derived depth and range offsets that are predictable by Eqs. (10) and (27), respectively, and that are independent of line array geometry. Also, the simulations show that no appreciable offsets in azimuth typically occur with this type of mismatch, even when the mismatch occurs at the receiver location, as predicted by Eq. (29). On the other hand, mismatch in array tilt leads to offsets in MFP estimates of source azimuth that follow Eq. (33), without significantly affecting the corresponding MFP estimates of range and depth. Of much greater importance, however, is the fact that the simple model successfully predicts the MFP results using real ocean acoustic data, as demonstrated in the next part of this section.

## B. Comparison of model predictions to results

Because of the nature of the adiabatic approximation, the simple model just presented predicts that the MFP offset in depth is determined by bathymetry mismatch at the source location and at the location of the receiver [only mismatch at the source location is taken into account in Eq. (10)], whereas the MFP offset in range is the result of the integrated mismatch between source and receiver. The two equations predicting the MFP depth and range offsets, i.e., Eqs. (10) and (27), are independent of mode number. Therefore, they predict that the MFP peak will remain coherent even in the presence of significant water depth mismatch. In addition, both equations are independent of frequency so that processing gain can be obtained by averaging across frequency. In contrast, the predicted MFP offset in azimuth due

to bathymetry mismatch depends only on mismatch at the receiver location. The term predicting this azimuth offset is dependent both upon mode number squared and inverse frequency squared, *re*: Eq. (29), so that bathymetry mismatch at the receiver usually will cause simply a degradation in MFP correlation. However, mismatch in array tilt has the same effect on azimuth as does bathymetry mismatch on range and depth, i.e., its predominant result is a shift of the apparent source location away from the true position and not significant degradation in broadband correlation.

Figure 5 shows a comparison of the prediction of the MFP range given by Eq. (27) with the actual MFP result for the Arc 94 event. Also plotted is the prediction of Eq. (28) where the bottom bathymetry is modeled by a single straight line between source and receiver (the curve labeled “1st Order Predicted Ranges”). Both predictions provide reasonably good approximations to the actual MFP range results. Deviations of the predictions from the actual results that occur after 52 min into the event may be caused by energy coupling between the modes, the higher order terms in the range integral in Eq. (22), and/or bottom penetration.

The fit of the simple model prediction for the MFP depth offset to the actual data along the Arc 94 track is presented in Fig. 7. The correspondence between the prediction and the actual MFP results is surprisingly good, even when bathymetry mismatch at the source location exceeds 100%. Therefore, the model provides quite accurate predictions of the MFP range and depth offsets due to bathymetry mismatch for data from a vertical line array.

The results to this point are obtained by matching with replica vectors from a range-independent environment where the assumed fixed water depth is greater than or equal to the actual water depth. The predicted offsets in MFP range and depth when the replica field is generated using a mismatched range-varying environment are shown in Figs. 8 and 9, respectively. The model predictions for this case are developed in the Appendix. For simplicity, the predictions are made only during the radial part of the Arc 94 track, where the actual water depth is approximately range independent. Again, the model fits to the actual MFP results are quite good.

These fits pertain to the MFP range and depth offsets obtained with data from a vertical array. Such an array is aligned with one of coordinate axes of symmetry of the acoustic waveguide. On the other hand, an array that is tilted significantly from the vertical breaks this symmetry, thereby potentially leading to coupling between the MFP offsets in range and depth, and deviations from the simple model predictions. However, the MFP range estimates from the VLA, TLA, and HLA arrays during the XSlope 96 event in SWellEx-96 are well modeled by the analytical predictions from Sec. IV B, as seen by the curve fits in Figs. 10, 11, and 12. These range offset predictions are calculated using a single straight-line approximation to the range-varying bathymetry between source and receiver [Eq. (28)]. The predictions from Eq. (28) agree extremely well with the results for all three array geometries. (Note that the predictions are not made until beginning about 7 min after the start of each figure because the source initially passed over a bathymetric

trough—*re*: the dotted bathymetry curve in Fig. 13—prior to this time, thereby preventing the use of a single straight-line approximation to the bathymetry.) This agreement, and the agreement between the three source depth estimates shown in Fig. 13, indicates that the MFP range and depth offsets due to bathymetry mismatch are independent of array geometry. A corollary to this conclusion is that the information contained in the acoustic field that a matched field processor uses to obtain source range and depth estimates is the same for all line array geometries, even when the array configuration breaks the waveguide symmetry.

For the MFP estimate of source azimuth, the model predicts that bathymetry mismatch does not, in general, lead to an offset. However, the results from the TLA for the XSlope 96 event display a definite bias away from endfire, *re*: Figs. 14 and 15. As Fig. 15 illustrates, this bias is not affected appreciably by *changes* in bathymetry mismatch, consistent with the model predictions. The cause of this azimuthal offset appears to be an error in the assumed tilt of the array. Taking this statement to be true, then Eq. (33) provides a means of inverting directly for the true array tilt using the observed azimuthal offset. That is, as the source passes through endfire, then  $\cos(\theta)$  equals unity and the azimuthal offset of approximately  $18^\circ$  away from the true endfire bearing of  $112^\circ$  T seen in Figs. 14 and 15, along with the assumed array tilt of  $47^\circ$  used in the replica field calculations, suggests that the true array tilt is closer to  $44^\circ$ . Figure 15 shows a comparison of the MFP azimuth results using a tilt of  $47^\circ$  and a tilt of  $44^\circ$  (plotted as asterisks) over the 7-min period of the XSlope 96 event as the source passed through endfire to the TLA. The solid curve is the actual source bearing over time from the navigation data. The azimuthal offset is nearly eliminated by using the smaller array tilt. Actually, it appears that the assumed tilt now is very slightly smaller than the true tilt because of the small offset in the other direction toward endfire, leading to the sequence of six asterisks at a constant  $112^\circ$  azimuth just after 58 min. These azimuth results are consistent with the simple model predictions. The model predictions also were verified by simulation, not presented here.

Finally, a prediction given by Eq. (33) is that errors in the tilt of horizontal arrays cause a much smaller MFP azimuth offsets than those for an array with a tilt around  $45^\circ$ . The accurate azimuth results for the HLA shown in Fig. 14 are robust with respect to small errors in array tilt (e.g., a monotonic variation in bottom depth across the array aperture), even near endfire.

### C. Physical interpretation

The development of the simple model illustrates that the MFP processor uses different pieces of information in the acoustic field to estimate source depth, range, and azimuth. That is, to obtain source depth, mode eigenfunction amplitudes are matched, whereas for estimates of range, the mode wave number differences, i.e., the mode interference patterns, are used. The azimuth of the source is derived by comparing the evolution of the wave field phase at a given frequency across the effective horizontal aperture of the array.

The predictions of the simple model based on these principles can be given simple physical interpretations. For the source depth prediction in Eq. (10), no matter what the water depth, a mode must satisfy the pressure-release boundary condition at the surface and at the bottom. Therefore, as a mode travels from shallow into deeper water (or vice versa), the top of its eigenfunction is “pinned” to the ocean surface as its “tail” is pulled downward (or pushed upward) in order to continually satisfy the bottom condition. The degree of stretching (or compression) is the same for all modes and causes the modal nodes and antinodes to get increasingly deeper (or shallower). The increase in the apparent source depth along the arc part of Arc 94 is the result of this mode stretching process.

The terms involved in matching in range are initially composed of two parts, a high-spatial-frequency part, i.e.,  $\exp[i\int k_l dr]$ , and a low-frequency part,  $\exp[i\int \Delta k_{n,l} dr]$ . The results indicate that the predicted MFP range is determined by the ratio of modal wave number differences rather than the ratio of modal wave numbers themselves. In effect, the information on the range of the source is in the low-frequency part, the amplitude modulation, rather than in the high-frequency carrier. If the  $2\pi$  ambiguity in phase could be removed somehow, then an MFP processor could be designed to account for absolute phase and this situation would change. A mode’s wave number becomes smaller as the mode travels into shallower water. However, because the higher order modes interact with the bottom more strongly, their modal wave numbers decrease more rapidly than those of the lower order modes [i.e.,  $\partial k_n / \partial d = (\pi^2/d^3) \times (n^2/k_n)$ ,  $d$  being the range-dependent bottom depth]. Therefore, the modal wave number differences increase when the wave field moves into shallower water even though the individual wave numbers decrease. In other words, the amplitude envelope variations increase in frequency whereas the carrier frequency decreases. This increase in the wave number differences leads to the appearance of the source at significantly greater ranges along the arc part of the Arc 94 track.

The MFP estimation of source azimuth is performed in a way similar to that of a plane wave beamformer. However, rather than using a single phase velocity to predict the phase evolution across the array aperture, the MFP processor assigns a different phase velocity to each mode. Only mismatch at the receiver location affects the estimate of source azimuth since the processor uses the phase differences between array elements, rather than the individual phases themselves. Bathymetry mismatch typically leads to degradation in correlation rather than an offset because the effect of water depth changes on the mode phase velocities is a function of mode number and frequency (*re*: the equation in the previous paragraph). With source azimuth, then, the mode wave numbers themselves (i.e., the high-frequency spatial part) are the important quantity versus the wave number *differences* used in the estimate of range. Conversely, mismatch in array tilt affects the apparent horizontal range over which the mode phases evolve rather than the mode phase velocities, and so is the same for all modes and frequencies. The fractional change in the horizontal aperture, rather than the absolute

amount of change, is the cause of the azimuthal offset so that the offset is independent of the array's total aperture.

## V. SOME APPLICATIONS AND EXTENSIONS

The simple model developed in the previous section can be applied in various ways to address issues important to matched field processing. Some of these applications are presented in this section.

### A. Effects of ocean surface waves

If a single spatial frequency component of an ocean surface wave field is superimposed on a range-independent waveguide of depth  $D$ , then the water depth can be expressed as:

$$d(r, t) = D + A \sin[\beta r + \phi(t)]. \quad (35)$$

Applying the prediction for the MFP peak in range given in Eqs. (22) and (23), keeping just the first term ( $m=1$ ), assuming that  $D \gg A$ , and that the ocean surface remains "frozen" over the period of time it takes for the sound to propagate, then:

$$\hat{r}(t) - R \approx 2 \frac{A}{\beta D} [\cos(\beta R + \phi(t)) - \cos(\phi(t))]. \quad (36)$$

This result shows that if an integral number of wavelengths exist between source and receiver, i.e.,  $\beta R = n2\pi$ , then the estimated MFP range equals the true range. However, if an additional half-wavelength is included, i.e.,  $\beta R = (2n+1)\pi$ , the MFP range error has a maximum amplitude of  $4A/(\beta D)$  and oscillates in time as  $\cos(\phi(t))$ .

The impact of water depth mismatch on shallow water MFP for source localization has been investigated previously using numerical simulation.<sup>22</sup> This study found that for an unperturbed, range-independent water layer of 100-m depth overlying a fluid half-space, a source range of 4 km, a vertical receiving array of 21 elements with 2.5-m spacing centered in the water layer, and a source frequency of 156 Hz: (1) the MFP output for source location "...varies in a systematic way," with increases in water depth causing the source to appear closer, and decreases in depth resulting in the source appearing to be farther away; (2) the MFP source range error due to water depth mismatch is independent of source depth (*re*: Fig. 5 of Ref. 22); and (3) the range error appears to be a linear function of water depth perturbations, with water depth changes of  $\pm 3$  m resulting in source range errors of  $\mp 250$  m (*re*: Fig. 5 of Ref. 22). These results in range error can be predicted either by evaluating the range integral over inverse depth squared directly, or by setting  $\phi(t) = \pi/2$  in Eq. (36) and taking the limit as the spatial frequency,  $\beta$ , goes to zero. The result is:

$$\hat{r} - R \approx -2 \frac{A}{D} R, \quad (37)$$

which predicts source range errors of  $\mp 240$  m that are independent of source depth for  $\pm 3$  m water depth mismatches, in good agreement with the simulation results. Note that the minimum variance processor was used in Ref. 22.

The theoretical predictions of the bias in MFP range and depth localization observed in the simulation results of Ref. 22 were originally presented in Ref. 23. This paper considers range-independent waveguides and accounts for sediment-type mismatch, through the same use of the concept of effective depth as used in the next section, and sound speed mismatch. Subsequent work by G. B. Smith (e.g., Ref. 24) also has examined the effects of various types of mismatch on MFP performance from an analytical point of view. Additional relevant papers on the issue of bathymetry mismatch include, among others, Refs. 25, 26, and 27.

Note that Eq. (37) differs from the expression for the single-straight-line bottom bathymetry model given in Eq. (28) since the latter equation assumes that no depth mismatch exists at the receiver location.

### B. Penetrability of the ocean bottom

The simple model can be extended to examine the impact of mismatch in geoacoustic properties of the ocean bottom on MFP source localization. This examination simultaneously provides insight into the use of MFP for inversions of geoacoustic properties of the ocean bottom. The environmental model under consideration is that of a homogeneous fluid water layer of density  $\rho_w$  and sound speed  $c_w$  overlying an elastic half-space with density  $\rho_s$ , compressional wave speed  $c_p$ , and shear wave speed  $c_s$ . The concept of effective depth, introduced by D. E. Weston<sup>28</sup> and extended to include shear wave effects by Chapman *et al.*,<sup>29</sup> provides the basis for this extension. The idea is that the phase change imparted by the reflection of a plane wave from a fluid-elastic interface is equal to that from a pressure-release boundary a distance  $\Delta H$  below the true interface. Since  $\Delta H$  is almost independent of the angle of incidence, then modal propagation in a waveguide having a penetrable bottom and water depth  $D$  can be approximated by propagation in an ideal waveguide bounded by two pressure-release boundaries (identical to the environmental model used in the development in Sec. IV) and having a water depth of  $D + \Delta H$ . From Eq. (17) of Ref. 29,  $\Delta H$  is given by:

$$\Delta H = [1 - 2(c_s/c_w)^2]^2 (\rho_s/\rho_w) c_w / [2\pi f \sin(\cos^{-1}(c_w/c_p))]. \quad (38)$$

Note that the change in effective water depth is inversely proportional to frequency. Therefore, at a single frequency, the impact of geoacoustic parameter mismatch on source localization is identical to the impact of water depth mismatch discussed in the previous section. For example, by replacing  $A$  in Eq. (37) with an expression for the differences in  $\Delta H$  using Eq. (38), the error in the single-frequency MFP estimate of source range due to a range-independent bottom geoacoustic parameter mismatch can be obtained. This result also suggests that the concept of "focalization," i.e., the process of simultaneously localizing a source and determining geoacoustic properties,<sup>4</sup> should be done at more than one frequency in order to distinguish unknown water depth changes from changes in geoacoustic properties.

The use of water-column MFP for inverting for geoacoustic properties of the bottom requires that the water-

column sound field contain information on these properties. A quantitative measure of the sensitivity is given by the logarithmic derivative. Equation (38) provides a simple analytical method for calculating the logarithmic derivative of the change in effective depth with respect to changes in bottom density, shear wave speed, and compressional wave speed. They are:

$$\frac{\partial \ln(\Delta H)}{\partial \ln(\rho_s)} = \frac{\partial \Delta H / \Delta H}{\partial \rho_s / \rho_s} = 1, \quad \frac{\partial \Delta H / \Delta H}{\partial c_s / c_s} = -\frac{8}{(c_w / c_s)^2 - 2}, \quad (39)$$

$$\frac{\partial \Delta H / \Delta H}{\partial c_p / c_p} = -\frac{(c_w / c_p)}{[1 - (c_w / c_p)^2]^{1/2}} \frac{1}{\tan(\cos^{-1}(c_w / c_p))}.$$

The expressions for these unitless quantities in Eq. (39) are a measure of the sensitivity of the sound field to changes in each of the geoacoustic parameters. Conversely, the inverse of these quantities is a measure of the sensitivity of the geoacoustic parameter inversion to unknown water depth changes. The negative signs in the logarithmic derivatives for shear and compressional wave speeds signify that increases in either of these bottom wave speeds result in decreases in the effective depth of the waveguide. By calculating these quantities using various realistic bottom parameters, one can see that the water-borne sound field becomes more sensitive to changes in a bottom seismic wave speed when that seismic speed approaches that of the water sound speed.

Historical data collected at the SWellEx site suggests that the sediment density is about 1.7 g/cc, the shear speed about 170 m/s, and the compressional speed about 1570 m/s. Using these values, the logarithmic derivatives in Eq. (39) have the values of 1, -0.11, and -9.1 for density, shear speed, and compressional speed, respectively. That is, the sound field should be quite sensitive to the bottom compressional wave speed and about an order of magnitude less so to the bottom density. However, because of the lack of sensitivity to changes in bottom shear wave speed, nothing of relevance can be said about this parameter. Conversely, errors in the assumed bottom shear wave velocity used as input to the calculation of the replica vectors will have no significant impact on source localization at the SWellEx site. Note that accounting for the penetrability of the ocean bottom corrects the tendency of the simple model predictions to overestimate the MFP range results in Fig. 5 near the end of the track.

The frequency dependence of  $\Delta H$  allows for bottom geoacoustic parameter mismatch to be separated from water depth mismatch in broadband MFP. However, a whole family of values of  $c_p$ ,  $\rho_s$ , and  $c_s$  combine to give the same value of  $\Delta H$ , so that additional information must be collected to uniquely determine each of these parameters. This form of the coupling between the bottom parameters holds when a model of a homogeneous fluid layer over a homogeneous halfspace is valid; appreciable gradients with depth in these bottom properties can cause both mode dependence and a more complicated frequency dependence to the mismatch.

### C. Effects of an offset in frequency

The effects of an offset in frequency can be derived easily from the simple model. Returning to Eq. (22), and introducing a frequency offset, i.e.,  $\hat{\omega} = \omega + \Delta\omega$ , rather than a mismatch in bottom bathymetry, then the first term in the Taylor series expansion gives:

$$\begin{aligned} \int_{R_0}^{\hat{r}_0} (\gamma_l^2 - \gamma_n^2) dr &\approx \frac{\Delta\omega}{\omega} \int_0^{R_0} (\gamma_l^2 - \gamma_n^2) dr \\ &\Rightarrow \int_{R_0}^{\hat{r}_0} \frac{1}{d(r)^2} dr \approx \frac{\Delta\omega}{\omega} \int_0^{R_0} \frac{1}{d(r)^2} dr, \end{aligned} \quad (40)$$

where Eq. (23) also has been used. If the actual bottom bathymetry is approximately range-independent, then:

$$\hat{r}_0 - R_0 = \frac{\Delta\omega}{\omega} R_0 \quad \text{or} \quad \frac{\Delta r}{r} = \frac{\Delta\omega}{\omega}. \quad (41)$$

Equation (41) is the same expression as that obtained using the waveguide invariant approach in shallow-water waveguides where the invariant,  $\beta$ , equals 1.<sup>27,30,31</sup> This offset of source range due to a change in frequency has been applied in time-reversal experiments to refocus the sound at different ranges in the ocean.<sup>32</sup>

Because of the frequency-independence of the modal vertical wave numbers and eigenfunctions in a homogeneous waveguide [*re*: Eq. (4)], then a frequency offset is predicted to have no effect on the estimate of source depth. For the estimate of azimuth, Eq. (19) can be used to derive the following expression:

$$\cos(\hat{\theta}) = \left( \frac{1}{1 + \frac{v_p^n \Delta\omega}{u_g^n \omega}} \right) \cos(\theta), \quad (42)$$

where  $v_p^n$  and  $u_g^n$  are the mode phase and group velocities, respectively. The ratio of velocities is dependent upon mode number and so frequency offsets typically cause a degradation in correlation for azimuth.

## VI. CONCLUSIONS AND IMPLICATIONS

The ocean acoustic data presented in this paper illustrate that broadband, shallow water, MFP estimates of source range and depth remain quite coherent in the presence of large mismatch in bottom bathymetry. However, the estimates can be offset significantly, sometimes by more than 100%, from their true positions. The offsets are independent of line array orientation. Broadband MFP estimates of source azimuth, for those array geometries with some horizontal aperture, also display large offsets, but the cause of mismatch in this case is in the vertical tilt of the array. Arrays with smaller nominal tilts from vertical show larger azimuthal offsets for a given degree of tilt mismatch. Changes in bathymetry mismatch even at the receiver location do not significantly affect the observed source azimuthal offsets.

The predictions from a simple analytical model based on adiabatic normal modes in ideal waveguides fit these MFP results surprisingly well. The prediction of the MFP peak in

depth due to bathymetry mismatch is obtained by matching true and replica fields in depth mode-by-mode, whereas the range prediction is based on matching modal spatial interference patterns in range. Both resulting equations are independent of mode number, at least to lowest order, so that the MFP peak will remain coherent with water depth mismatch. Also, both equations are independent of frequency, implying that processing gain can be obtained by averaging across frequency. Source azimuth is obtained by comparing the evolution of the wave field phase at a given frequency across the effective horizontal aperture of the array. The term arising from the effect of bathymetry mismatch on source azimuth is dependent upon mode number squared and inverse frequency squared, so that this type of mismatch typically results in degradation in MFP correlation. However, mismatch in array tilt results in an expression that is independent of mode number and frequency, and so predicts an offset of the source azimuth estimate from the true azimuth when this form of mismatch is present. Because of the nature of the adiabatic approximation, the prediction of the MFP depth estimate depends only on the properties at the source and receiver locations, whereas the prediction in range depends on the integrated properties between source and receiver. Source azimuth is affected only by bathymetry mismatch at the receiver location.

The good match between the predictions and the MFP results in this paper rely on the fact that the waveguide boundaries in this shallow water environment are formed primarily by reflection from the sea surface and bottom. Therefore, the waveguide thickness is independent of frequency and mode number. For waveguides formed by refraction, such as the deep ocean sound channel, the effective thickness depends on frequency and particularly on mode number. In this case, the predictions of the simple model no longer hold. For the same reasons, they cannot be expected to hold in shallow water environments where significant bottom penetration and upward refraction within the bottom occurs.

The SWellEx broadband MFP results, and the physical insight that the simple analytical model provides, leads to interesting implications for the application of MFP. Some of these are:

- (1) MFP mismatch can lead to large shifts in location of the broadband MFP peak, as well as degradation in correlation. For various types of mismatch, these two effects contribute to varying degrees.
- (2) Azimuthally dependent bathymetry in many cases will provide only poor azimuthal resolution for vertical hydrophone arrays.
- (3) Synthetic aperture studies may be affected by unknown water depth changes by introducing error in the synthetic element positions.
- (4) Line arrays tilted about  $45^\circ$  from vertical provide estimates of source range and depth that compare favorably with those from an equivalent-aperture vertical array. However, a tilted array also gives estimates of source azimuth.
- (5) Line arrays that are angled significantly away from the

horizontal are sensitive to tilt mismatch. However, the measured offset of the estimated source azimuth from the true azimuth of a known source can be used to directly invert for the true array tilt.

- (6) Water depth errors can cause a frequency dependence of the bottom geoacoustic parameters in matched field inversions.
- (7) Conversely, "focalization," i.e., the simultaneous estimate of source position and geoacoustic properties, should be done broadband.
- (8) The concept of "effective depth" may permit simple inversions for geoacoustic parameters with broadband MFP data, but only for a family of parameter values, and not for values of each parameter uniquely.
- (9) Mismatch in frequency behaves in a way similar to that of bathymetry mismatch for estimates of source range (offset) and azimuth (degradation in correlation).
- (10) The use of range-independent geoacoustic models allow for shallow water, broadband MFP results to be generated rapidly, and can be easily corrected for bathymetry mismatch.

## ACKNOWLEDGMENTS

Dick Harriss at the Marine Physical Lab worked through all the ocean engineering details in deploying the hydrophone line arrays from FLIP. Joe Rice at NRaD performed the demanding task of test director. In addition, special thanks go to the crew of R/P FLIP and to Bill Gaines whose seagoing support of these experiments has been tremendous. Paul Baxley at NRaD made many important contributions to the analysis of the SWellEx data, including his work on matched field inversion for geoacoustic parameters (Ref. 16). A careful reading of the draft of this paper and useful critical comments were provided to us by R. T. Kessel at the University of Victoria. Additional helpful discussions were held with W. A. Kuperman at MPL, with N. R. Chapman at the University of Victoria, and with G. B. Smith at NRL Stennis. The SWellEx-3 results discussed in this paper were originally presented at the Fall, 1994, Matched Field Processing Workshop in Victoria, Canada, in a Marine Physical Laboratory Technical Memorandum (Ref. 21), at the Spring, 1995 Acoustical Society meeting (Ref. 33), and at the International Conference on Acoustics in Trondheim, Norway (Ref. 34). The SWellEx-96 results were discussed at the Spring, 1997 Acoustical Society meeting (Ref. 35). This work was supported by the Space and Naval Warfare Systems Center under Contract No. N66001-96-C-6000 and the Office of Naval Research, Code 321, under Contract No. N00014-93-D-0141 (DO #6).

## APPENDIX: SIMPLE MODEL PREDICTIONS FOR REPLICA VECTORS FROM A RANGE-DEPENDENT ENVIRONMENT

This Appendix extends the simple model predictions in Sec. IV A to cases where the replica field is calculated using a range-dependent environmental model. As before, the adiabatic normal mode approximation is used for both replica and true fields.

Assuming, as in Sec. IV, that no mismatch occurs at the receiver location, then Eq. (9) for the matching in depth still holds and becomes:

$$\hat{z}_s = \frac{\hat{d}_s}{d_s} z_s. \quad (\text{A1})$$

Since the replica vectors are calculated from a range-dependent environment, then  $\hat{d}_s$  is dependent upon the estimated range,  $\hat{r}_0$ . Therefore, in contrast to the range-independent replica case, the match in range must be solved first before Eq. (A1) can be evaluated.

Considering the matching of modal interference patterns, Eq. (16) becomes:

$$\int_0^{R_0} \Delta k_{n,l} dr = \int_0^{\hat{r}_0} \Delta \hat{k}_{n,l} dr. \quad (\text{A2})$$

Equations (22), (23), and (26) can be applied to evaluate the integrals on both sides of Eq. (A2). Keeping just the first term in the expansion [to order  $(\gamma_n/k)^3$ ], then:

$$\sum_{i=1}^M \frac{x_i - x_{i-1}}{d_i d_{i-1}} \approx \sum_{j=1}^{\hat{M}} \frac{\hat{x}_j - \hat{x}_{j-1}}{\hat{d}_j \hat{d}_{j-1}}, \quad (\text{A3})$$

where  $R_0 = \sum_{i=1}^M (x_i - x_{i-1})$  and  $\hat{r}_0 = \sum_{j=1}^{\hat{M}} (\hat{x}_j - \hat{x}_{j-1})$ . To apply this equation, the left-hand side is first calculated for a given true source location using the actual bottom bathymetry. The right-hand side is then evaluated with the bottom bathymetry used in the replica calculations to determine the value of  $\hat{N}$  such that:

$$\sum_{j=1}^{\hat{N}-1} \frac{\hat{x}_j - \hat{x}_{j-1}}{\hat{d}_j \hat{d}_{j-1}} \leq \sum_{i=1}^M \frac{x_i - x_{i-1}}{d_i d_{i-1}} \leq \sum_{j=1}^{\hat{N}} \frac{\hat{x}_j - \hat{x}_{j-1}}{\hat{d}_j \hat{d}_{j-1}}. \quad (\text{A4})$$

Linear interpolation is performed in the last range sector, i.e., between ranges  $\hat{x}_{\hat{N}-1}$  and  $\hat{x}_{\hat{N}}$  to obtain  $\hat{x}_{\hat{M}}$  such that Eq. (A3) is satisfied. The value of  $\hat{r}_0$  is gotten by subtracting from  $\hat{x}_{\hat{M}}$  the initial range which, for most applications, equals zero.

If the true bottom bathymetry is approximately range independent with water depth  $D$ , as it is along the radial part of Arc 94, then the left-hand side of Eq. (A3) equals  $R_0/D^2$ . If, in addition, the range-dependence of the bathymetry used in calculating the replica vectors can be modeled by a single straight line with slope  $\hat{\alpha}$ , then:

$$\hat{r}_0 \approx \frac{D}{D + \hat{\alpha} R_0} R_0. \quad (\text{A5})$$

This simplification is used to obtain the predictions shown in Figs. 8 and 9.

<sup>1</sup>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).

<sup>2</sup>J. M. Tran and W. S. Hodgkiss, "Array surveying using matched-field processing," J. Acoust. Soc. Am. **94**, 2851–2858 (1993).

<sup>3</sup>W. S. Hodgkiss, D. E. Ensberg, J. J. Murray, G. L. D'Spain, N. O. Booth, and P. W. Schey, "Direct measurement and matched-field inversion approaches to array shape estimation," IEEE J. Ocean Eng. **21**, 393–401 (1996).

<sup>4</sup>M. D. Collins and W. A. Kuperman, "Focalization: Environmental focusing and source localization," J. Acoust. Soc. Am. **90**, 1410–1422 (1991).

<sup>5</sup>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).

<sup>6</sup>E. J. Sullivan and D. Middleton, "Estimation and detection issues in matched-field processing," IEEE J. Ocean Eng. **18**, 156–167 (1993).

<sup>7</sup>A. Tolstoy, *Matched Field Processing for Underwater Acoustics* (World Scientific, Singapore, 1993).

<sup>8</sup>D. F. Gingras, "Methods for predicting the sensitivity of matched-field processors to mismatch," J. Acoust. Soc. Am. **86**, 1940–1949 (1989).

<sup>9</sup>J. S. Perkins and W. A. Kuperman, "Environmental signal processing: Three-dimensional matched field processing with a vertical array," J. Acoust. Soc. Am. **87**, 1553–1556 (1990).

<sup>10</sup>"Observations of surface vessel activity off San Diego Harbor during the summer of 1994," Tracor Rept. No. T-95-56-0001-U, Tracor Applied Sciences, San Diego, CA (1995).

<sup>11</sup>W. S. Hodgkiss, J. C. Nickles, G. L. Edmonds, R. A. Harriss, and G. L. D'Spain, "A large dynamic range vertical array of acoustic sensors," Full Field Inversion Methods in Ocean and Seismic Acoustics Conference, SAFLANT Undersea Research Centre, La Spezia, Italy (1994).

<sup>12</sup>W. S. Hodgkiss, J. J. Murray, N. O. Booth, and P. W. Schey, "An inter-comparison of low-level signal detectability in a shallow water environment using horizontal, vertical, and tilted vertical arrays," J. Acoust. Soc. Am. **101**, 3047 (1997).

<sup>13</sup>N. O. Booth, P. W. Schey, and W. S. Hodgkiss, "Detection of low-level broadband signals using adaptive matched-field processing," J. Acoust. Soc. Am. **101**, 3047 (1997).

<sup>14</sup>M. B. Porter, "The Kraken normal mode program," user's manual, SAFLANT Undersea Research Centre, La Spezia, Italy (1991).

<sup>15</sup>R. T. Bachman, P. W. Schey, N. O. Booth, and F. J. Ryan, "Geoacoustic databases for matched-field processing: Preliminary results in shallow water off San Diego, California," J. Acoust. Soc. Am. **99**, 2077–2085 (1996).

<sup>16</sup>P. A. Baxley, "Matched-field replica model optimization and bottom property inversion in shallow water," Master's Thesis, University of California, San Diego (1997).

<sup>17</sup>K. O. Emery, W. S. Butcher, H. R. Gould, and F. P. Shepard, "Submarine geology off San Diego, California," J. Geol. **60**, 511–548 (1952).

<sup>18</sup>E. L. Hamilton, "Geoacoustic modeling of the sea floor," J. Acoust. Soc. Am. **68**, 1313–1340 (1980).

<sup>19</sup>N. O. Booth, P. A. Baxley, J. A. Rice, P. W. Schey, W. S. Hodgkiss, G. L. D'Spain, and J. J. Murray, "Source localization with broadband matched-field processing in shallow water," IEEE J. Ocean Eng. **21**, 402–412 (1996).

<sup>20</sup>M. B. Porter, "Acoustic models and sonar systems," IEEE J. Ocean Eng. **18**, 425–437 (1993).

<sup>21</sup>G. L. D'Spain, J. J. Murray, W. S. Hodgkiss, and N. O. Booth, "Predicting the broadband matched field processing results for the arc event during SWellEx-3," MPL Tech Memo 442, Marine Physical Laboratory, Scripps Institution of Oceanography, San Diego, CA (1994).

<sup>22</sup>D. R. Del Balzo, C. Feuillade, and M. M. Rowe, "Effects of water-depth mismatch on matched-field localization in shallow water," J. Acoust. Soc. Am. **83**, 2180–2185 (1988).

<sup>23</sup>E. C. Shang and Y. Y. Wang, "Environmental mismatching effects on source localization processing in mode space," J. Acoust. Soc. Am. **89**, 2285–2290 (1991).

<sup>24</sup>G. B. Smith, M. A. Akundi, and N. D. Gardner, "Source depth dependence of mismatch-induced range and depth localization errors in conventional matched-field processing," J. Acoust. Soc. Am. **97**, 3292 (1995).

<sup>25</sup>S. E. Dosso, "Matched-field inversion for source localization with uncertain bathymetry," J. Acoust. Soc. Am. **94**, 1160–1163 (1993).

<sup>26</sup>P. Zakarauskas, S. E. Dosso, and J. A. Fawcett, "Matched-field inversion for source localization and optimal equivalent bathymetry," J. Acoust. Soc. Am. **100**, 1493–1500 (1996).

<sup>27</sup>G. A. Grachev, "Theory of acoustic field invariants in layered waveguides," Acoust. Phys. **39**, 67–71 (1993).

<sup>28</sup>D. E. Weston, "A Moire fringe analog of sound propagation in shallow water," J. Acoust. Soc. Am. **32**, 647–654 (1960).

<sup>29</sup>D. M. F. Chapman, P. D. Ward, and D. D. Ellis, "The effective depth of a Pekeris ocean waveguide, including shear wave effects," J. Acoust. Soc. Am. **85**, 648–653 (1989).

- <sup>30</sup>S. D. Chuprov, "Interference structure of a sound field in a layered ocean," in *Ocean Acoustics, Current State*, edited by L. M. Brekhovskikh and I. B. Andreevoi (Nauka, Moscow, 1982), pp. 71–91.
- <sup>31</sup>L. M. Brekhovskikh and Y. P. Lysanov, *Fundamentals of Ocean Acoustics*, 2nd ed., New York (Springer-Verlag, New York, 1991), pp. 140–145.
- <sup>32</sup>H. C. Song, W. A. Kuperman, and W. S. Hodgkiss, "A time reversal mirror with variable range focusing," *J. Acoust. Soc. Am.* **103**, 3234–3240 (1998).
- <sup>33</sup>G. L. D'Spain, J. J. Murray, W. S. Hodgkiss, and N. O. Booth, "Mirages in shallow water matched-field processing," *J. Acoust. Soc. Am.* **97**, 3291 (1995).
- <sup>34</sup>G. L. D'Spain, W. S. Hodgkiss, and W. A. Kuperman, "Broadband matched field processing in shallow water," *Proc. 15th Int'l Conf. on Acoustics*, Vol. I, pp. 231–234 (1995).
- <sup>35</sup>G. L. D'Spain, W. S. Hodgkiss, J. J. Murray, P. W. Schey, and N. O. Booth, "Analytical and experimental comparison of azimuth/range/depth bias errors in MFP source localization using vertical and tilted arrays," *J. Acoust. Soc. Am.* **101**, 3047 (1997).



# A new calculation procedure for spatial impulse responses in ultrasound

Jørgen Arendt Jensen

*Department of Information Technology, Building 344, Technical University of Denmark, DK-2800 Lyngby, Denmark*

(Received 24 October 1997; revised 18 February 1999; accepted 25 February 1999)

A new procedure for the calculation of spatial impulse responses for linear sound fields is introduced. This calculation procedure uses the well known technique of calculating the spatial impulse response from the intersection of a circle emanating from the projected spherical wave with the boundary of the emitting aperture. This general result holds for all aperture boundaries for a flat transducer surface, and is used in the procedure to yield the response for all types of flat transducers. An arbitrary apodization function over the aperture can be incorporated through a simple one-dimensional integration. The case of a soft baffle mounting of the aperture is also included. Specific solutions for transducer boundaries made from lines are given, so that any polygon transducer can be handled. Specific solutions for circles are also given. Finally, a solution for a general boundary is stated, and all these boundary elements can be combined to, e.g., handle annular arrays or semi-circle transducers. Results from an implementation of the approach are given and compared to previously developed solutions for a simple aperture, a complex aperture, and a Gaussian apodized circular transducer. © 1999 Acoustical Society of America. [S0001-4966(99)01406-X]

PACS numbers: 43.35.Cg, 43.20.Px [HEB]

## INTRODUCTION

The calculation of linear, acoustic fields is most often based on the spatial impulse response approach as suggested by Tupholme and Stepanishen.<sup>1-3</sup> Here the pulsed pressure field is found from a convolution between the acceleration of the transducer surface and the spatial impulse response. The impulse response has been found for a number of geometries (round flat piston,<sup>2</sup> round concave,<sup>4,5</sup> flat rectangle,<sup>6,7</sup> and flat triangle<sup>8</sup>). The solutions arrived at are often complicated, since it involves the evaluation of the Rayleigh surface integral. The response depends on the relative position of the field point and many special cases exist, which makes both the derivation of the solution difficult and the evaluation of the responses cumbersome. For example, to evaluate the response from a rectangle, four synthetic rectangles are introduced, and when evaluating a triangle, three synthetic triangles are introduced to account for the 15 different possible cases of triangle shape and field point positions. This makes it necessary to use computers for evaluating and interpreting the responses, since the formulas do not readily give a useful perception of the solution.

It would be appropriate to arrive at general solutions for any geometry that would be both easy to derive analytically and fast to evaluate with a computer. This has previously been achieved by dividing the aperture surface into smaller elements like rectangles<sup>9</sup> or triangles,<sup>10</sup> and then summing the response for the sub-elements. Often the transducer must be divided into many elements and only a piecewise approximation to the apodization is obtained. The fitting to the actual surface is also only approximative for round or oval surfaces; even when using a triangular shape.

Spatial impulse responses from bounded and non-apodized apertures always have discontinuities due to their sharp edges, which makes it difficult to keep the full energy and spectral content in a sampled evaluation. Various techniques have been applied for coping with the discontinuities in the spatial impulse response. This has included using very high sampling frequencies, making a time adapted evaluation, or using the integrated response. Computer evaluation is, thus, always necessary, when evaluating spatial impulse responses.

This paper therefore suggests a new procedure for calculating the spatial impulse response in which the computer is involved at an earlier stage in the evaluation of the responses. The response is determined by the crossings of the boundary of the aperture by the spherical wave emitted from the field point. For flat apertures this observation makes it possible to derive a general approach for calculating the spatial impulse response for any aperture geometry and find the response with no approximation. The paper derives impulse responses for apertures described by bounding lines and circles and outlines how the response can be evaluated for apertures bounded by any polynomial in the plane's coordinates  $(x,y)$ . Some of the intersections need not always be calculated, and this is used to devise an optimized algorithm that only needs to find the minimum number of intersections. It is also shown in Sec. III how an arbitrary apodization can be introduced through a previously developed simple one-dimensional integration, and how the solution also can be applied to both the soft baffle and rigid baffle situations. A number of examples from use of the approach are given in Sec. VII. Among these is a comparison between the tradi-

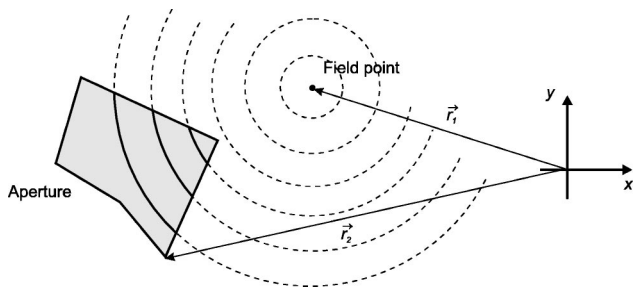


FIG. 1. Intersection of spherical waves from the field point by the aperture, when the field point is projected onto the plane of the aperture.

tional solution for a rectangle and the new approach, which both yield the same response.

### I. BASIC THEORY

A short review of the calculation of spatial impulse responses is given in this section to facilitate the development of the new calculation procedure.

The spatial impulse response is found from the Rayleigh integral given by:<sup>2,11</sup>

$$h(\mathbf{r}_1, t) = \int_S \frac{\delta\left(t - \frac{|\mathbf{r}_1 - \mathbf{r}_2|}{c}\right)}{2\pi|\mathbf{r}_1 - \mathbf{r}_2|} dS \quad (1)$$

when the aperture  $S$  is mounted in an infinite, rigid baffle. Here  $\mathbf{r}_1$  denotes the position of the field point,  $\mathbf{r}_2$  denotes a position on the aperture,  $c$  is the speed of sound, and  $t$  is time. The integral is essentially a statement of Huyghen's principle that the field is found by summing the radiated spherical waves from all parts of the aperture. This can also be reformulated, due to acoustic reciprocity, as finding the part of the spherical wave emanating from the field point that intersects the aperture. The task is, thus, to project the field point onto the plane coinciding with the aperture, and then find the intersection of the projected spherical wave (the circle) with the active aperture as shown in Fig. 1.

Rewriting the integral into polar coordinates gives:

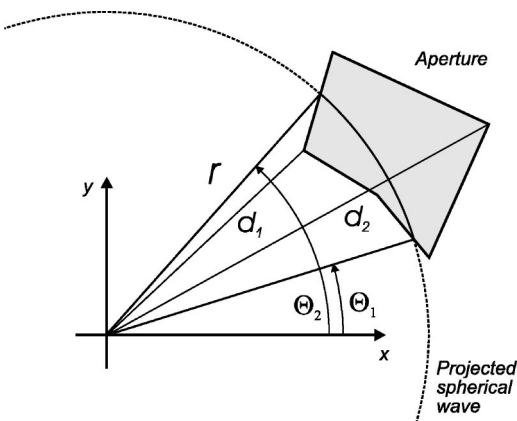


FIG. 2. Definition of distances and angles in the aperture plan for evaluating the Rayleigh integral.

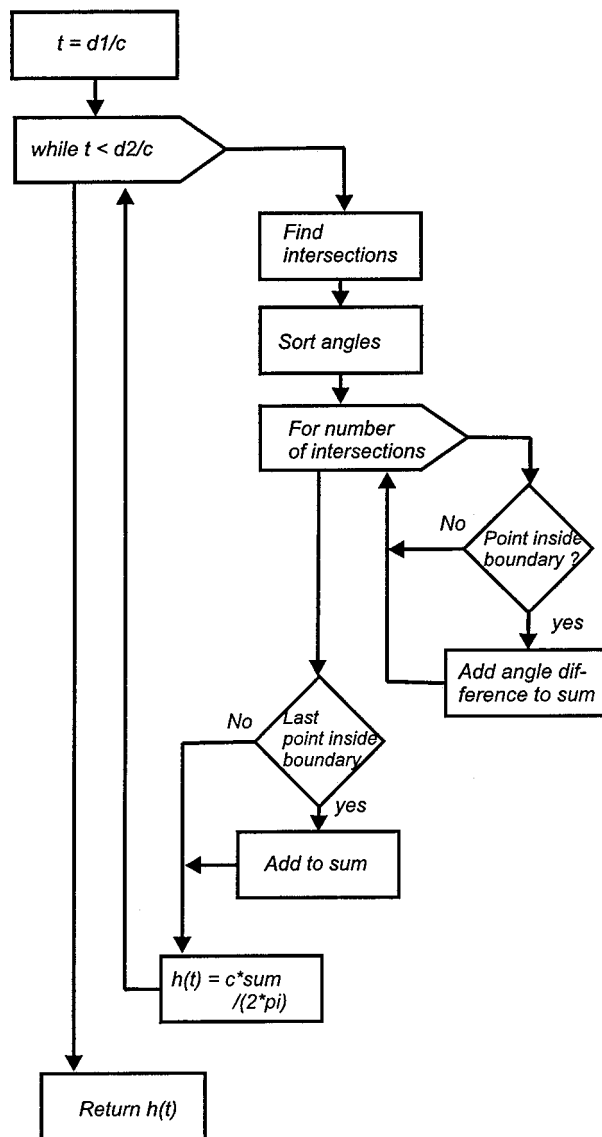


FIG. 3. Flow chart for the simple approach for calculating the spatial impulse response.

$$h(\mathbf{r}_1, t) = \int_{\Theta_1}^{\Theta_2} \int_{d_1}^{d_2} \frac{\delta\left(t - \frac{R}{c}\right)}{2\pi R} r dr d\Theta, \quad (2)$$

where  $r$  is the radius of the projected circle and  $R$  is the distance from the field point to the aperture given by  $R^2 = r^2 + z_p^2$ . Here  $z_p$  is the field point height above the  $x-y$  plane of the aperture. The projected distances  $d_1, d_2$  are determined by the aperture and are the distance closest to and furthest away from the aperture, and  $\Theta_1, \Theta_2$  are the corresponding angles for a given time (see Fig. 2).

Introducing the substitution  $2R dR = 2r dr$  gives

$$h(\mathbf{r}_1, t) = \frac{1}{2\pi} \int_{\Theta_1}^{\Theta_2} \int_{R_1}^{R_2} \delta\left(t - \frac{R}{c}\right) dR d\Theta. \quad (3)$$

The variables  $R_1$  and  $R_2$  denote the edges closest to and furthest away from the field point. Finally using the substitution  $t' = R/c$  gives

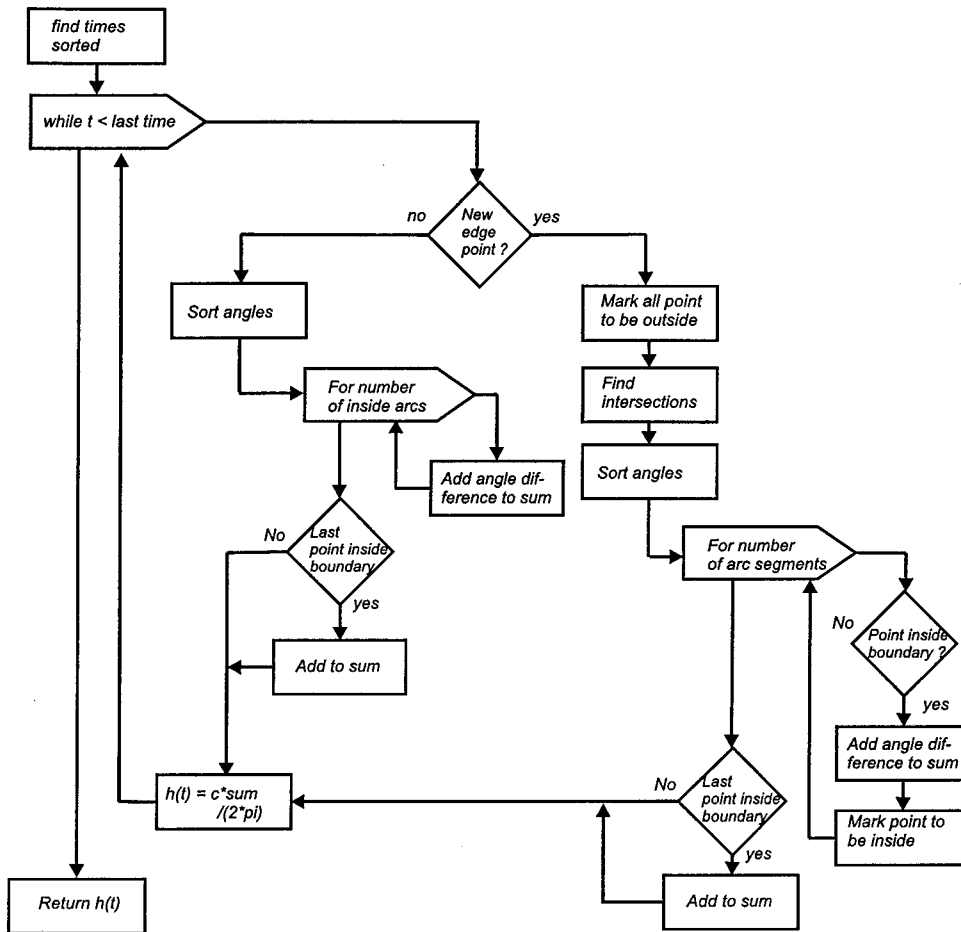


FIG. 4. Flow chart for the optimized approach for calculating the spatial impulse response.

$$h(\mathbf{r}_1, t) = \frac{c}{2\pi} \int_{\Theta_1}^{\Theta_2} \int_{t_1}^{t_2} \delta(t-t') dt' d\Theta. \quad (4)$$

For a given time instance the contribution along the arc is constant and the integral gives

$$h(\mathbf{r}_1, t) = \frac{\Theta_2 - \Theta_1}{2\pi} c \quad (5)$$

when assuming the circle arc is only intersected once by the aperture. The angles  $\Theta_1$  and  $\Theta_2$  are determined by the intersection of the aperture and the projected spherical wave, and the spatial impulse response is, thus, solely determined by these intersections, when no apodization of the aperture is used. The response can therefore be evaluated by keeping track of the intersections as a function of time.

## II. A NEW CALCULATION PROCEDURE

From the derivation in the last section it can be seen that the spatial impulse response in general can be expressed as

$$h(\mathbf{r}_1, t) = \frac{c}{2\pi} \sum_{i=1}^{N(t)} [\Theta_2^{(i)}(t) - \Theta_1^{(i)}(t)], \quad (6)$$

where  $N(t)$  is the number of arc segments that crosses the boundary of the aperture for a given time and  $\Theta_2^{(i)}(t)$ ,  $\Theta_1^{(i)}(t)$  are the associated angles of the arc. This was also noted by Stepanishen.<sup>12</sup> The calculation can, thus, be formu-

lated as finding the angles of the aperture edge's intersections with the projected spherical wave, sorting the angles, and then summing the arc angles that belong to the aperture. Finding the intersections can be done from the description of the edges of the aperture. A triangle can be described by three lines, a rectangle by four, and the intersections are then found from the intersections of the circle with the lines. This makes it possible to devise a general procedure for calculating spatial impulse responses for any flat, bounded aperture, since the task is just to find the intersections of the boundary with the circle.

The spatial impulse response is calculated from the time the aperture first is intersected by a spherical wave to the time for the intersection furthest away. The intersections are found for every time instance and the corresponding angles are sorted. The angles lie in the interval from 0 to  $2\pi$ . It is then found whether the arc between two angles belongs to the aperture, and the angle difference is added to the sum, if the arc segment is inside the aperture. This yields the spatial impulse response according to Eq. (6). The approach can be described by the flow chart shown in Fig. 3.

The only part of the algorithm specific to the aperture is the determination of the intersections and the whether the point is inside the aperture. Section IV shows how this is done for polygons, Sec. V for circles, and Sec. VI for higher-order parametric boundaries.

All the intersections need not be found for all times. New intersections are only introduced, when a new edge or

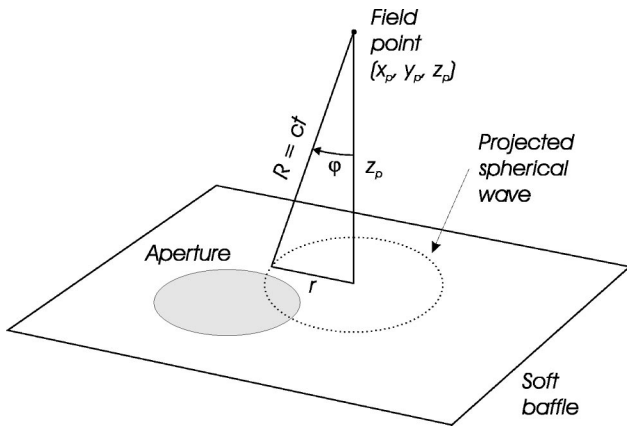


FIG. 5. Definition of angle used for a soft baffle.

corner of the aperture is met. Between times when two such corners or edges are encountered the number of intersections remains constant and only intersections, which belong to points inside the aperture need be found. Note that an aperture edge gives rise to a discontinuity in the spatial impulse response. Also testing whether the point is inside the aperture is often superfluous, since this only needs to be found once after each discontinuity in the response. These two observations can significantly reduce the number of calculations, since only the intersections affecting the response are found. The flow chart for the optimized approach is shown in Fig. 4.

The procedure first finds the number of discontinuities. Then only intersection influencing the response are calculated between two discontinuity points. This can potentially make the approach faster than the traditional approach, where the response from a number of different rectangles or triangles must be calculated.

### III. APODIZATION AND SOFT Baffle

Often ultrasound transducers do not vibrate as a piston over the aperture. This can be due to the clamping of the active surface at its edges, or intentionally to reduce side lobes in the field. Applying for example a Gaussian apodization will significantly lower side lobes and generate a field with a more uniform point spread function as a function of depth. Apodization has previously been found and is introduced in Eq. (2) by writing<sup>13</sup>

$$h(\mathbf{r}_1, t) = \int_{\Theta_1}^{\Theta_2} \int_{d_1}^{d_2} a_p(r, \Theta) \frac{\delta\left(t - \frac{R}{c}\right)}{2\pi R} r dr d\Theta \quad (7)$$

in which  $a_p(r, \Theta)$  is the apodization over the aperture. Using the same substitutions as before yields

$$h(\mathbf{r}_1, t) = \frac{c}{2\pi} \int_{\Theta_1}^{\Theta_2} \int_{t_1}^{t_2} a_{p1}(t', \Theta) \delta(t - t') dt' d\Theta, \quad (8)$$

where  $a_{p1}(t', \Theta) = a_p(\sqrt{(ct')^2 - z_p^2}, \Theta)$ . The inner integral is a convolution of the apodization function with a  $\delta$ -function and readily yields

$$h(\mathbf{r}_1, t) = \frac{c}{2\pi} \int_{\Theta_1}^{\Theta_2} a_{p1}(t, \Theta) d\Theta \quad (9)$$

as noted by several authors.<sup>13-15</sup> The response for a given time point can, thus, be found by integrating the apodization function along the fixed arc with a radius of  $r = \sqrt{(ct)^2 - z_p^2}$  for the angles for the active aperture. Any apodization function can therefore be incorporated into the calculation by employing numerical integration.

Often the assumption of an infinite rigid baffle for the transducer mounting is not appropriate and another form of the Rayleigh integral must be used. For a soft baffle, in which the pressure on the baffle surface is zero, the Rayleigh-Sommerfeld integral is used. This is (Ref. 16, pp. 46-50)

$$h_s(\mathbf{r}_1, t) = \int_S \frac{\delta\left(t - \frac{|\mathbf{r}_1 - \mathbf{r}_2|}{c}\right)}{2\pi|\mathbf{r}_1 - \mathbf{r}_2|} \cos \varphi dS, \quad (10)$$

assuming that  $|\mathbf{r}_1 - \mathbf{r}_2| \gg \lambda$ . Here  $\cos \varphi$  is the angle between the line through the field point orthogonal to the aperture plane and the radius of the spherical wave as shown in Fig. 5. The angle  $\varphi$  is fixed for a given radius of the projected spherical wave and thus for a given time. It is given by

$$\cos \varphi = \frac{z_p}{R} = \frac{z_p}{ct}. \quad (11)$$

Using the substitutions from Sec. I the Rayleigh-Sommerfeld integral can then be rewritten as

$$h_s(\mathbf{r}_1, t) = \frac{z_p}{2\pi} c (\Theta_2 - \Theta_1) \int_{t_1}^{t_2} \frac{\delta(t - t')}{ct'} dt'. \quad (12)$$

Using the property of the  $\delta$ -function that

$$\int_{-\infty}^{+\infty} g(t') \delta(t - t') dt' = g(t) \quad (13)$$

then gives

$$h_s(\mathbf{r}_1, t) = \frac{z_p}{ct} \frac{\Theta_2 - \Theta_1}{2\pi} c = \frac{z_p}{ct} h(\mathbf{r}_1, t). \quad (14)$$

The spatial impulse response can, thus, be found from the spatial impulse response for the rigid baffle case by multiplying with  $z_p/(ct)$ .

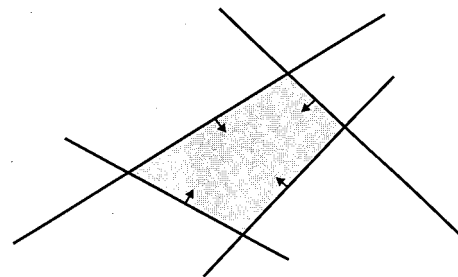


FIG. 6. Definition of bounding lines for polygon transducer. The arrows indicates the half-planes for the active aperture.

#### IV. SOLUTION FOR POLYGONS

The boundary of any polygon can be defined by a set of bounding lines as shown in Fig. 6. The active aperture is then defined as lying on one side of the line as indicated by the arrows, and a point on the aperture must be placed correctly in relation to all lines. The test whether a point is on the aperture is thus to go through all lines and test whether the point lies in the active half space for the line, and stop if it is not. The point is inside the aperture, if it passes the test for all the lines.

The intersections are found from the individual intersections between the projected circle and the lines. They are determined from the equations for the projected spherical wave and the line:

$$\begin{aligned} r^2 &= (x-x_0)^2 + (y-y_0)^2, \\ y &= \alpha x + y_1, \\ r^2 &= (ct)^2 - z_p^2. \end{aligned} \quad (15)$$

Here  $(x_0, y_0)$  is the center of the circle,  $\alpha$  the slope of the line, and  $y_1$  its intersect with the  $y$ -axis. The intersections are given from the solutions to:

$$\begin{aligned} 0 &= (1 + \alpha^2)x^2 + (2\alpha y_1 - 2x_0 - 2y_0\alpha)x \\ &\quad + (y_0^2 + y_1^2 + x_0^2 - 2y_0y_1 - r^2) \\ &= Ax^2 + Bx + C, \\ D &= B^2 - 4AC. \end{aligned} \quad (16)$$

The angles are

$$\Theta = \arctan\left(\frac{y-y_0}{x-x_0}\right). \quad (17)$$

Intersections between the line and the circle are only found if  $D > 0$ . A determinant  $D < 0$  indicates that the circle did not intersect the line. If the line has infinite slope, the solution is found from the equation:

$$\begin{aligned} x &= x_1, \\ 0 &= y^2 - 2y_0y + y_0^2 + (x_1 - x_0)^2 - r^2 \\ &= A_\infty y^2 + B_\infty y + C_\infty, \end{aligned} \quad (18)$$

in which  $A_\infty$ ,  $B_\infty$ ,  $C_\infty$  replaces  $A$ ,  $B$ ,  $C$ , respectively, and the solutions are found for  $y$  rather than  $x$ . Here  $x_1$  is the line's intersection with the  $x$ -axis.

The times for discontinuities in the spatial impulse response are given by the intersections of the lines that define the aperture's edges and by the minimum distance from the projected field point to the lines. The minimum distance is found from a line passing through the field point that is orthogonal to the bounding line. The intersection between the orthogonal line and the bounding line is:

$$\begin{aligned} x &= \frac{\alpha y_p + x_p - \alpha y_1}{\alpha^2 + 1}, \\ y &= \alpha x + y_1, \end{aligned} \quad (19)$$

where  $(x_p, y_p, z_p)$  is the position of the field point. For an infinite slope line the solution is  $x = x_1$  and  $y = y_p$ . The corresponding time is:

$$t_i = \frac{\sqrt{(x-x_p)^2 + (y-y_p)^2 + z_p^2}}{c}. \quad (20)$$

The intersections of the lines are also found, and the corresponding times are calculated by Eq. (20) and sorted in ascending order. They indicate the start and end time for the response and the time points for discontinuities in the response.

#### V. SOLUTION FOR CIRCULAR SURFACES

The other basic shape for a transducer apart from rectangular shapes is the flat, round surface used for single element piston transducers and annular arrays. For these the intersections are determined by two circles as depicted in Fig. 7. Here  $O_1$  is the center of the aperture with radius  $r_a$

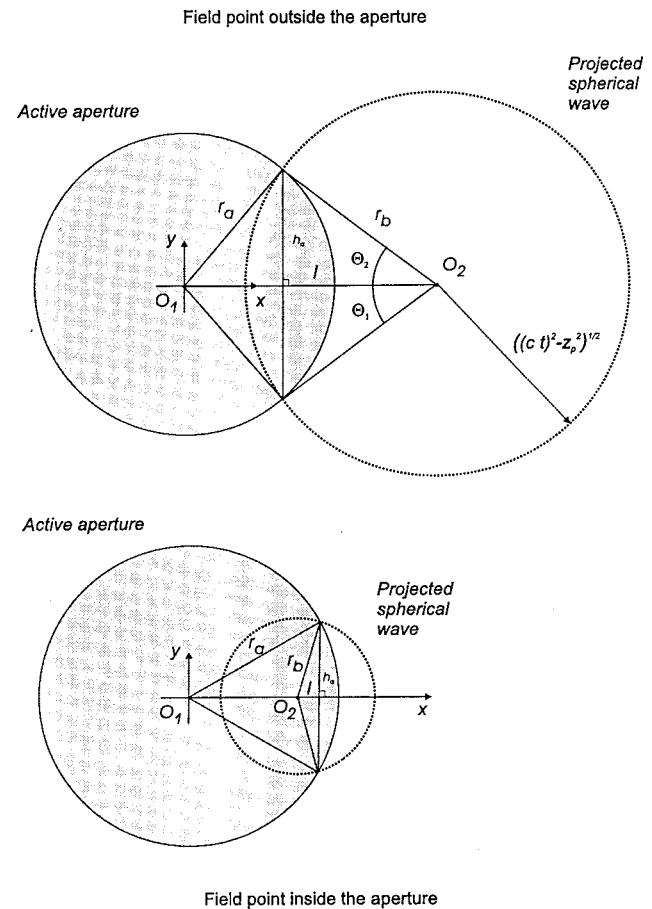


FIG. 7. Geometry for determining intersections between circles. The top graph shows the geometry when the field point denoted by  $O_2$  is outside the aperture, and the bottom graph when it is inside.

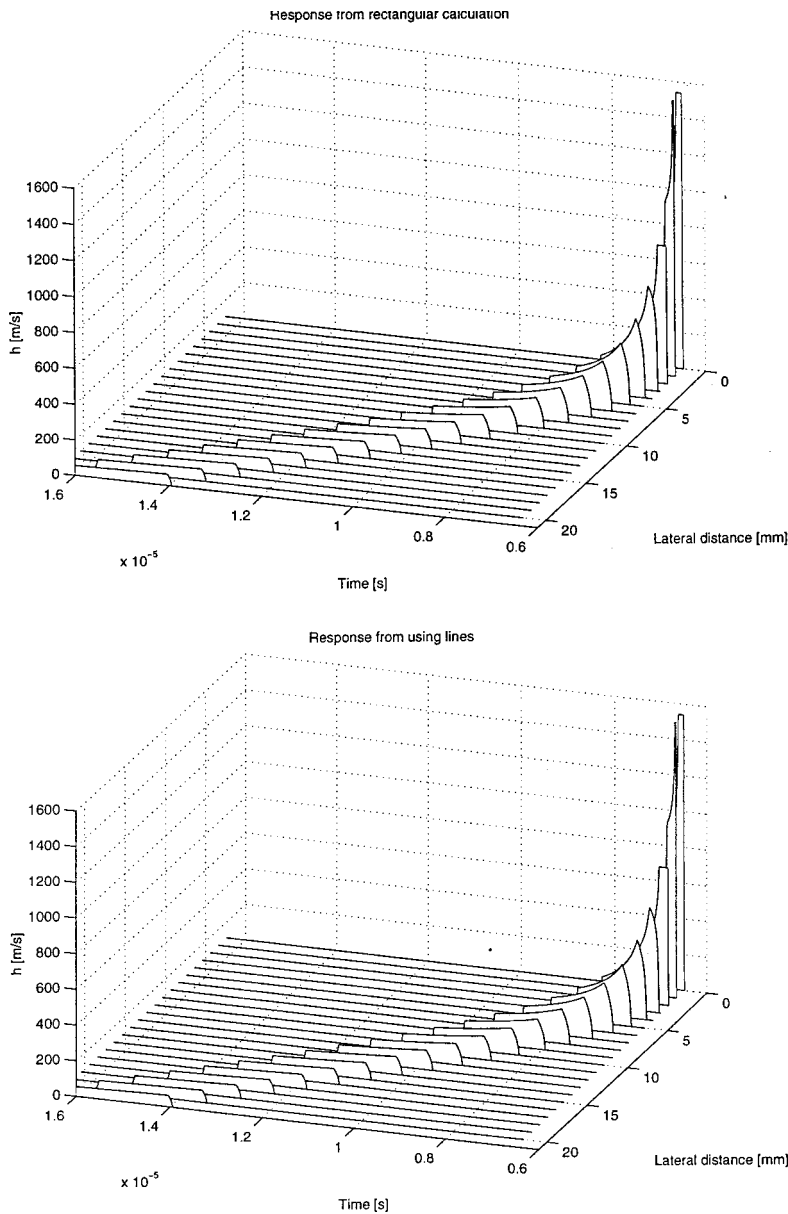


FIG. 8. Spatial impulse response calculated from a rectangular transducer of  $4 \times 5$  mm. The top graph shows the result from using traditional evaluation and the bottom graph is when using the new method with four bounding lines. The axial distance to the field point is 10 mm and the response is calculated for lateral distances from 0 to 21 mm off-axis in steps of 1 mm.

and the projected spherical wave is centered at  $O_2$  with radius  $r_b(t) = \sqrt{(ct)^2 - z_p^2}$ . The length  $h_a(t)$  is given by (Ref. 17, p. 66)

$$h_a(t) = \frac{2\sqrt{p(t)(p(t)-a)(p(t)-r_a)(p(t)-r_b(t))}}{a},$$

$$a = \|O_1 - O_2\|, \quad (21)$$

$$p(t) = \frac{a + r_a + r_b(t)}{2}.$$

In a coordinate system centered at  $O_1$  and an  $x$ -axis in the  $O_1 - O_2$  direction, the intersections are at

$$y = h_a(t),$$

$$l = \pm \sqrt{r_b^2(t) - h_a^2(t)}. \quad (22)$$

The sign for  $l$  depends on the position of the intersections. A negative sign is used if the intersections are for negative values of  $x$ , and positive sign is used for positive  $x$  positions.

When the field point is outside the active aperture the spatial impulse response is

$$h(\mathbf{r}_1, t) = \frac{|\Theta_2 - \Theta_1|}{2\pi} c = \frac{c}{\pi} \arctan\left(\frac{h_a(t)}{l}\right),$$

$$\Theta_2 = \arctan\left(\frac{h_a(t)}{l}\right) = -\Theta_1. \quad (23)$$

It must be noted that a proper four-quadrant arctan should be used to give the correct response. An alternative formula is (Ref. 18, p. 19)

$$h(\mathbf{r}_1, t) = \frac{c}{2\pi} \arcsin$$

$$\times \left( \frac{2\sqrt{p(t)(p(t)-a)(p(t)-r_a)(p(t)-r_b(t))}}{r_b^2(t)} \right),$$

$$= \frac{c}{2\pi} \arcsin\left(\frac{ah_a(t)}{r_b^2(t)}\right). \quad (24)$$

The start time  $t_s$  for the response is found from

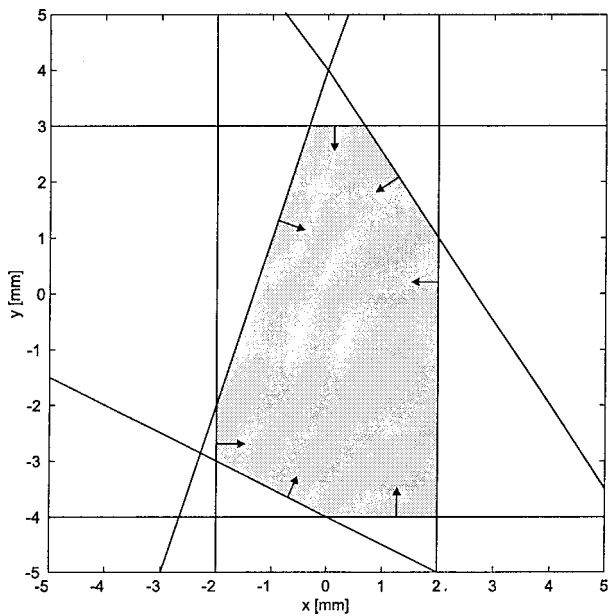


FIG. 9. Bounding lines defining a complex aperture. The arrows indicates the half-plane for the active aperture.

$$r_a + r_b(t) = \|O_1 - O_2\|, \quad (25)$$

$$t_s = \frac{\sqrt{r_b^2(t) + z_p^2}}{c} = \frac{\sqrt{(\|O_1 - O_2\| - r_a)^2 + z_p^2}}{c},$$

and the response ends at the time  $t_e$  when

$$r_b(t) = r_a + \|O_1 - O_2\|, \quad (26)$$

$$t_e = \frac{\sqrt{r_b^2(t) + z_p^2}}{c} = \frac{\sqrt{(\|O_1 - O_2\| + r_a)^2 + z_p^2}}{c}.$$

When the field point is inside the aperture, the response is

$$h(\mathbf{r}_1, t) = c \quad \text{for} \quad \frac{z_p}{c} \leq t \leq \frac{\sqrt{(r_a - \|O_1 - O_2\|)^2 + z_p^2}}{c}; \quad (27)$$

thereafter the arc lying outside the aperture should be subtracted, so that

$$h(\mathbf{r}_1, t) = \frac{2\pi - |\Theta_2 - \Theta_1|}{2\pi} c. \quad (28)$$

The response ends when

$$r_b(t) = r_a + \|O_1 - O_2\|, \quad (29)$$

$$t_e = \frac{\sqrt{(\|O_1 - O_2\| + r_a)^2 + z_p^2}}{c}.$$

The determination of which part of the arc that subtracts or adds to the response is determined by what the active aperture is. One ring in an annular array can be defined as consisting of an active aperture outside a circle combined with an active aperture inside a circle for defining the inner and outer rim of the aperture. A circular aperture can also be combined with a line for defining the active area of a split aperture used for continuous wave probing.

## VI. SOLUTION FOR PARAMETRIC SURFACES

For ellipses or other higher-order parametric surfaces it is in general not easy to find analytic solutions for the spatial impulse response. The procedure described can, however, devise a simple solution to the problem, since the intersections between the projected spherical wave and the edge of the aperture uniquely determine the spatial impulse response. It is therefore possible to use root finding for a set of (non-linear) equations for finding these intersections. The problem is to find when both the spherical wave and the aperture have crossing contours in the plane of the aperture, i.e., when

$$(ct)^2 - z_p^2 - (x - x_p)^2 - (y - y_p)^2 = 0, \quad (30)$$

$$S(x, y) = 0,$$

in which  $S(x, y) = 0$  defines the boundary of the aperture. The problem of numerically finding these roots is in general not easy, if a good initial guess on the position of the inter-

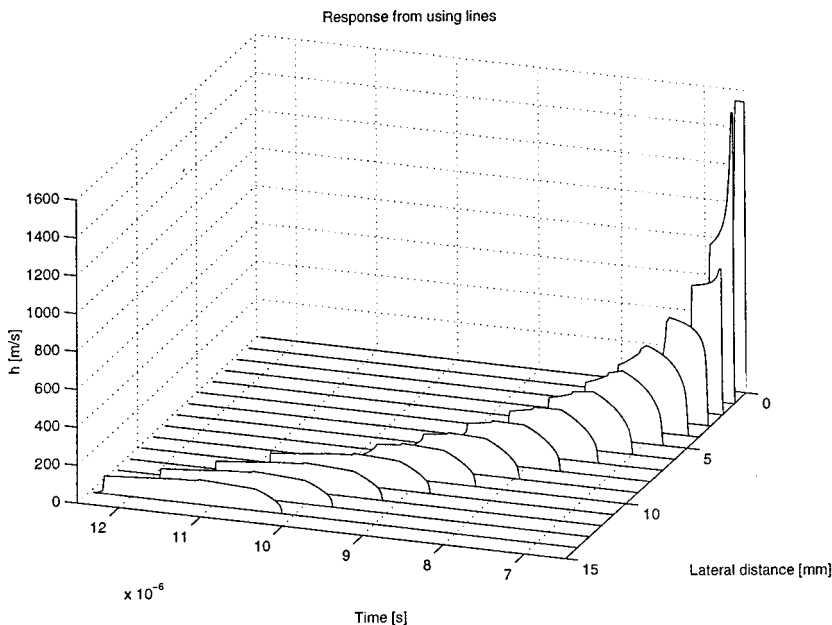


FIG. 10. Spatial impulse response from the complex aperture defined in Fig. 9. The axial distance to the field point is 10 mm and the response is calculated for lateral distances from 0 to 21 mm off-axis in the  $x$ -direction.

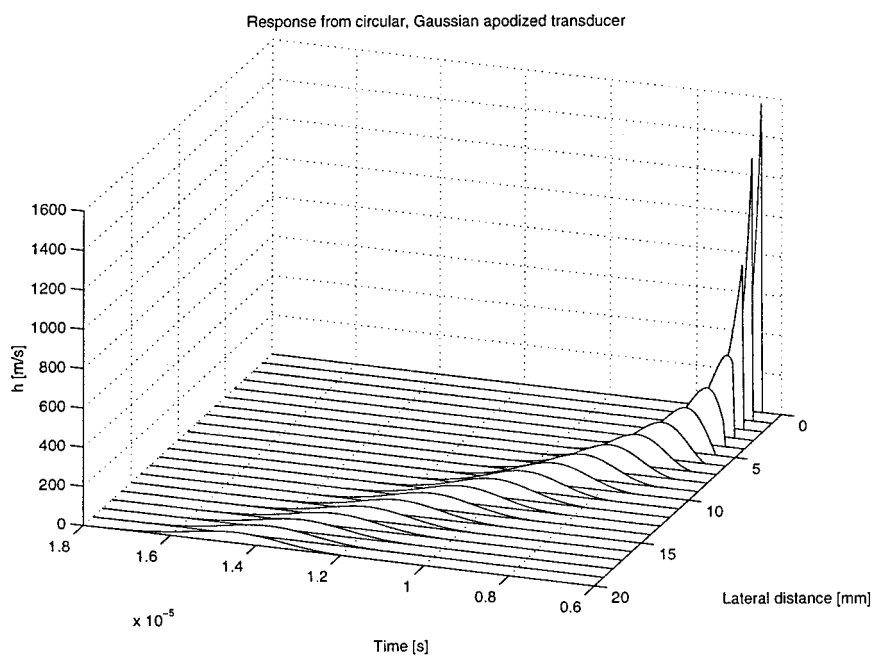
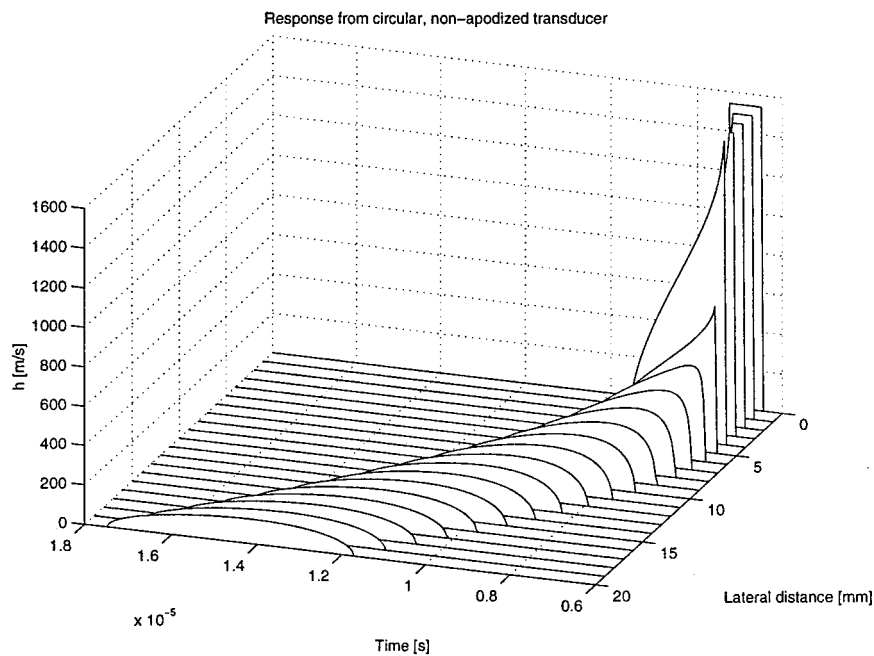


FIG. 11. Spatial impulse response from a circular aperture calculated with the new method. Graphs are shown without apodization of the aperture (top) and with a Gaussian apodization function (bottom). The radius of the aperture is 5 mm and the field is calculated 10 mm from the transducer surface.

sections is not found (Ref. 19, pp. 286–289). Good initial values are, however, found here, since the intersections must lie on the projected circle and the intersections only move slightly from time point to time point. An efficient Newton–Raphson algorithm can therefore be devised for finding the intersections, and the procedure detailed here can be made to find the spatial impulse response for any flat transducer geometry with an arbitrary apodization and both hard and soft baffle mounting.

## VII. EXAMPLES

The first example shows a comparison between the traditional method for calculating spatial impulse responses and the new method. The response from a  $4 \times 5$  mm rectangle is found for different spatial positions 10 mm from the front face of the transducer. The responses are found from the

center of the rectangle and out in steps of 1 mm in the  $x$ -direction to 21 mm away from the center of the rectangle. The results are shown in Fig. 8. It is seen that the two methods give identical results.

The second example is for a more complicated aperture, where its bounding lines are shown in Fig. 9. The calculated spatial impulse response is shown in Fig. 10. Responses have been calculated from the center position for  $x=0$  mm,  $y=0$  mm to the position  $x=14$  mm,  $y=0$  mm in increments of 1 mm. The distance to the transducer surface was always 10 mm ( $=z$ ). A complicated response with a number of discontinuities is seen due to the many edges of the aperture.

The last example shows the response from a circular, flat transducer calculated with the new method. Two different cases are shown in Fig. 11. The top graph shows the tradi-



tional spatial impulse response when no apodization is used, so that the aperture vibrates as a piston. The field is calculated 10 mm from the front face of the transducer starting at the center axis of the aperture. Twenty-one responses for lateral distance of 0–20 mm off-axis are then shown. The same calculation is repeated in the bottom graph, when a Gaussian apodization has been imposed on the aperture. The vibration amplitude is a factor of  $1/\exp(4)$  less at the edges of the aperture than at the center. It is seen how the apodization reduces some of the sharp discontinuities in the spatial impulse response.

## VIII. SUMMARY

The general theory for the calculation of spatial impulse responses for flat transducer apertures has been reviewed. It was shown that the response can be found from the intersections of the projected spherical wave with the edges of the aperture. This made it possible to derive a general procedure for calculating these responses, which can be used for any flat geometry. It was also shown that the calculation easily can include hard and soft baffle mountings and the incorporation of any apodization function over the aperture. The approach makes it possible to make simulation programs that can handle arbitrary, flat, apodized aperture geometries without making approximations in the evaluation of the spatial impulse response for the hard baffle case.

<sup>1</sup>G. E. Tupholme, "Generation of acoustic pulses by baffled plane pistons," *Mathematika* **16**, 209–224 (1969).

<sup>2</sup>P. R. Stepanishen, "Transient radiation from pistons in an infinite planar baffle," *J. Acoust. Soc. Am.* **49**, 1629–1638 (1971).

<sup>3</sup>P. R. Stepanishen, "The time-dependent force and radiation impedance on a piston in a rigid infinite planar baffle," *J. Acoust. Soc. Am.* **49**, 841–849 (1971).

<sup>4</sup>A. Penttinen and M. Luukkala, "The impulse response and nearfield of a curved ultrasonic radiator," *J. Phys. D* **9**, 1547–1557 (1976).

<sup>5</sup>M. Arditi, F. S. Forster, and J. Hunt, "Transient fields of concave annular arrays," *Ultrason. Imaging* **3**, 37–61 (1981).

<sup>6</sup>J. C. Lockwood and J. G. Willette, "High-speed method for computing the exact solution for the pressure variations in the nearfield of a baffled piston," *J. Acoust. Soc. Am.* **53**, 735–741 (1973).

<sup>7</sup>J. L. S. Emeterio and L. G. Ullate, "Diffraction impulse response of rectangular transducers," *J. Acoust. Soc. Am.* **92**, 651–662 (1992).

<sup>8</sup>J. A. Jensen, "Ultrasound fields from triangular apertures," *J. Acoust. Soc. Am.* **100**, 2049–2056 (1996).

<sup>9</sup>J. A. Jensen and N. B. Svendsen, "Calculation of pressure fields from arbitrarily shaped, apodized, and excited ultrasound transducers," *IEEE Trans. Ultrason. Ferroelectr. Freq. Control* **39**, 262–267 (1992).

<sup>10</sup>J. A. Jensen, "Simulating arbitrary geometry ultrasound transducers using triangles," in *Proc. IEEE Ultrason. Symp.*, Vol. 2 (IEEE, New York, 1996), pp. 885–888.

<sup>11</sup>G. R. Harris, "Review of transient field theory for a baffled planar piston," *J. Acoust. Soc. Am.* **70**, 10–20 (1981).

<sup>12</sup>P. R. Stepanishen, "Wide bandwidth near and far field transients from baffled pistons," in *Proc. IEEE Ultrason. Symp.* (IEEE, New York, 1977), pp. 113–118.

<sup>13</sup>G. R. Harris, "Transient field of a baffled planar piston having an arbitrary vibration amplitude distribution," *J. Acoust. Soc. Am.* **70**, 186–204 (1981).

<sup>14</sup>P. R. Stepanishen, "Acoustic transients from planar axisymmetric vibrators using the impulse response approach," *J. Acoust. Soc. Am.* **70**, 1176–1181 (1981).

<sup>15</sup>J. Naze Tjøtta and S. Tjøtta, "Near field and far field of pulsed acoustic radiators," *J. Acoust. Soc. Am.* **71**, 824–834 (1982).

<sup>16</sup>J. W. Goodman, *Introduction to Fourier Optics*, 2nd ed. (McGraw-Hill, New York, 1996).

<sup>17</sup>L. Raade and B. Westergreen, *β Mathematics Handbook* (Charwell-Bratt, Ltd., Kent, England, 1990).

<sup>18</sup>M. R. Spiegel, *Mathematical Handbook of Formulas and Tables* (McGraw-Hill, New York, 1968).

<sup>19</sup>W. H. Press, B. P. Flannery, S. A. Teukolsky, and W. T. Vetterling, *Numerical Recipes in C. The Art of Scientific Computing* (Cambridge University Press, Cambridge, 1988).

# Three-dimensional beam directivity of phase-steered ultrasound

Shi-Chang Wooh<sup>a)</sup> and Yijun Shi

*Department of Civil and Environmental Engineering, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139*

(Received 16 July 1998; accepted for publication 22 February 1999)

A model is developed to compute the three-dimensional far-field acoustic pressure distribution of the waves radiated from an ultrasonic linear phased array with finite elemental dimensions. Based on the model, the three-dimensional beam directivity function is introduced and analyzed to investigate the effects of geometric parameters on the steering performance. It is confirmed that the elevation dimension does not influence the beam directivity in the azimuthal plane. This is not true, however, in nonazimuthal planes. In general, increasing the elevation dimension reduces the energy leaking into nonazimuthal directions. The influences of other transducer parameters, including the number of elements, interelement spacing, and aperture length, are also investigated. It is found that they have only trivial effects on the sideleaking, although they are the important factors that influence the directivity in the azimuthal plane. © 1999 Acoustical Society of America.

[S0001-4966(99)00506-8]

PACS numbers: 43.38.Hz, 43.35.Yb [SLE]

## INTRODUCTION

The utilization of ultrasonic phased arrays in the field of nondestructive evaluation (NDE) has been receiving greater attention in recent years. Ultrasonic phased arrays are composed of multiple piezoelectric elements which are excited by predetermined time-delayed pulses to sweep the beam throughout the area of interest. Beam steering and dynamic focusing properties of phased arrays give rise to their advantages over conventional ultrasonic transducers, including high inspection speed, flexible control and signal processing, and fast imaging capabilities.

Among the various types of phased arrays available, linear arrays may be best suited for NDE purposes due to their simplicity, despite the fact that they only provide two-dimensional beam steering and focusing, i.e., their use is limited to a single steering (or azimuthal) plane. An effective utilization in NDE is closely dependent on the beam steering performance of the transducer, which requires analysis and optimal design details.

The steering and focusing characteristics of an array is determined by its geometric parameters including the interelement spacing  $d$  and number of elements  $N$  as well as the other dimensions such as element width  $a$  and elevation dimension  $L$ , which is also a function of loading materials and transducer frequency  $f$ . Many investigators studied the behavior of wave motion contributed by these factors both analytically and numerically.<sup>1-9</sup> In analyzing the characteristics, phased arrays are generally considered as an array of simple point sources arranged in the azimuthal plane. This approach is mathematically easier to deal with, yet it gives the same directivity in the azimuthal plane. However, an actual array made of such points may not perform its intended function of steering and it is required to understand the behavior of arrays made of finite dimensions. A lack of comprehensive

study on the effects of array parameters is a common drawback of these studies.

In our previous study, the far-field acoustic pressure field of a wave radiated from a linear array and the beam directivity were analyzed to find optimum beam steering parameters<sup>10</sup> and to study the effect of element widths.<sup>11</sup> In the analysis, an array was considered to have elements of finite width  $a$  and infinitely large elevation dimension. This model can be established by first finding a solution for a line source and then applying the Huygen's principle to obtain the solutions for each element and phased arrays. The solution of this model takes the form of a Hankel function of the second kind. This approach makes sense particularly when the array element size  $a$  is significantly smaller than the elevation dimension  $L$  or the aperture length  $D$ , which is a reasonable assumption. However, it is not always safe to discard the effects of  $a$  and  $L$ . In particular, the steering characteristics in the directions other than the azimuthal plane may be quite different if the elevation dimension is taken into account.

If each element has a certain finite elevation dimension  $L$ , it is quite predictable that a certain fraction of energy will propagate in the nonazimuthal directions. It is our interest to study the influence of the elevation dimension on the steering performance. From a design point of view, it is important to suppress the energy leaking into nonazimuthal directions as compared to the one in the azimuthal direction. In this paper, we focus on analyzing three-dimensional directivity of a linear phased array consisting of elements of finite dimensions  $a$  and  $L$ . It is also our interest to study the effect of other transducer parameters including  $N$ ,  $d$ , and  $D$ .

## I. THREE-DIMENSIONAL DIRECTIVITY ANALYSIS

As discussed before, beam steering is one of the attractive features of phased arrays, which requires good directivity of the beam propagating in the steering plane. The best steering effect can be achieved by analyzing the directivity

<sup>a)</sup>Electronic mail: scwooh@mit.edu

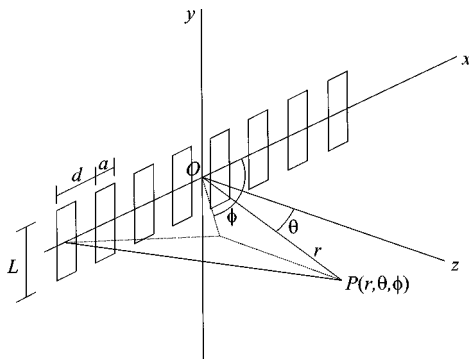


FIG. 1. Geometry and coordinate systems used to study acoustic waves radiated and phase-steered by a linear array with finite dimension ( $d$  = interelement spacing,  $a$  = element width, and  $L$  = elevation dimension).

and seeking a condition that provides maximum pressure in the direction of the desired main beam propagation (or in the steering direction), while suppressing the pressure in other directions.<sup>10</sup>

Before studying the effects of geometrical transducer parameters, we will express the directivities and analyze them for both single element and arrays. In general, the directivity of a transducer can be obtained by first evaluating the pressure at an arbitrary point in the propagating medium and then expressing the variation of pressure as a function of angle in the plane of interest. In our analysis, a phased array is modeled as an assemblage of rectangular elements of finite dimensions (width and elevation dimensions). The elements are activated sequentially at the times delayed by a predetermined linear relationship, resulting in steering of the beam in its azimuthal angle. Huygen's principle is applied first to compute the pressure field of a representative transducer element and an array consisting of discrete point sources. The pressure and directivity of an array of rectangular elements can be derived by multiplying these solutions.

### A. Single element

For a single rectangular element, the pressure field of a wave traveling in an acoustic medium can be derived by assuming that the element is composed of an infinitely large number of simple sources pulsating at the same amplitudes and phases. Figure 1 schematically shows the model used to study the directivity functions of a linear array consisting of  $N$  number of elements, where  $a$  is the element size (width),  $L$  is the elevation dimension, and  $d$  is the interelement spacing, respectively. The distance from a point  $P$  to the center of the array is denoted by  $r$ , and  $(\theta, \phi)$  are the corresponding azimuthal and rotational angles in the spherical coordinates with an origin at the center of the element. In this case, by assuming that  $r \gg a, L, \lambda$ , the pressure at the point  $P$  can be expressed in the form<sup>12</sup>

$$p(r, \theta, \phi) = \frac{jk\rho c U_0}{2\pi r} aL \operatorname{sinc}\left(\frac{\pi a \sin \theta \cos \phi}{\lambda}\right) \times \operatorname{sinc}\left(\frac{\pi L \sin \theta \sin \phi}{\lambda}\right) e^{jkr}, \quad (1)$$

where  $k$  is the wave number,  $\rho$  is the density of the material,

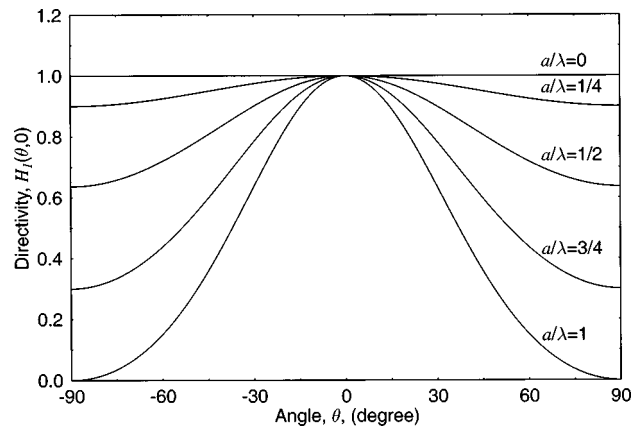


FIG. 2. Directivity of a single element as a function of angle  $\theta$  in the azimuthal plane for various element width  $a$ .

$U_0$  is the velocity amplitude of the element surface,  $\lambda$  is wavelength,  $c$  is the wavespeed in the medium, and  $j$  is the unit imaginary number, respectively.

The *beam directivity* or the *directivity function* is defined as a normalized pressure amplitude  $p(r, \theta, \phi)$  at any arbitrary direction  $(\theta, \phi)$ . Directivity is a good measure that indicates the direction of beam propagation, which is an important quantity to be studied in order to understand the beam propagation characteristics. From the definition, the directivity function for a single rectangular element  $H_1(\theta, \phi)$  can be written as a function of the element dimensions  $a$  and  $L$  as

$$H_1(\theta, \phi) = \left| \operatorname{sinc}\left(\frac{\pi a \sin \theta \cos \phi}{\lambda}\right) \operatorname{sinc}\left(\frac{\pi L \sin \theta \sin \phi}{\lambda}\right) \right|. \quad (2)$$

It can be also shown that the directivity in the azimuthal plane (steering plane,  $\phi=0$  or  $\phi=\pi$ ) can be rewritten as a function of only  $a$ ,

$$H_1(\theta, 0) = \left| \operatorname{sinc}\left(\frac{\pi a \sin \theta}{\lambda}\right) \right|, \quad (3)$$

which is exactly the same expression as the one obtained from a one-dimensional treatment of array elements.<sup>10,11</sup> Figure 2 shows the influence of element widths on the directivity in the azimuthal plane. When the azimuthal angle  $\theta=0$ , the directivity becomes  $H_1(0, \phi) = 1$ , meaning that the pressure reaches its maximum in this direction. In fact, the direction  $(0, \phi)$  is understood as the same as the direction  $(0, 0)$ .

### B. Phased array

As stated above, directivities of a phased array can be derived from the solutions for a single element and an array made of discrete point sources. In analyzing the arrays, it is assumed that the wave originated from each source has the same amplitude but different phases resulting from the given time delays  $\Delta\tau$  between any two neighboring elements,

$$\Delta\tau = \frac{d \sin \theta_s}{c}, \quad (4)$$

to produce a wavefront propagating at the steering angle  $\theta_s$  in the azimuthal plane. By applying Huygen's principle, the

far-field directivity function of the discrete point sources can be easily derived as

$$H_2(\theta, \phi) = \left| \frac{\sin[(\pi d/\lambda)(\sin \theta_s - \sin \theta \cos \phi)N]}{N \sin[(\pi d/\lambda)(\sin \theta_s - \sin \theta \cos \phi)]} \right|, \quad (5)$$

where  $(\theta, \phi)$  represent the angles of a vector in spherical coordinates whose origin is positioned at the center of the array. Note here that the function  $H_2$  is no longer restricted in the azimuthal plane. The resulting directivity for phased arrays is then derived by multiplying the two directivities,  $H(\theta, \phi) = H_1(\theta, \phi) \cdot H_2(\theta, \phi)$ , as

$$H(\theta, \phi) = \left| \operatorname{sinc} \left( \frac{\pi a \sin \theta \cos \phi}{\lambda} \right) \operatorname{sinc} \left( \frac{\pi L \sin \theta \sin \phi}{\lambda} \right) \right| \times \left| \frac{\sin[(\pi d/\lambda)(\sin \theta_s - \sin \theta \cos \phi)N]}{\sin[(\pi d/\lambda)(\sin \theta_s - \sin \theta \cos \phi)]} \right|. \quad (6)$$

Note that the directivity in Eq. (6) is a three-dimensional function that allows us to analyze the beam propagation pattern in a nonazimuthal plane. If the rotational angle  $\phi=0$  or only the azimuthal plane is considered,  $H(\theta, \phi)$  becomes a two-dimensional directivity  $H(\theta, 0)$  that is also consistent with the results from the two-dimensional analysis.

## II. GEOMETRIC PARAMETERS AND SIDELEAKING ENERGY

As shown in the previous section, the three-dimensional directivity  $H(\theta, \phi)$  is equivalent to the two-dimensional counterpart  $H(\theta)$  if  $\phi=0$ . Since the primary use of linear phased arrays is to steer the beam in the azimuthal plane where  $\phi=0$  (or  $\phi=\pi$ ), the two-dimensional directivity  $H(\theta)$  is sufficient to serve the purpose. However, it is anticipated that some energy may leak into nonazimuthal directions of the array, which is also one of the important phenomena that determines the transducer performance. This section studies the effects of the transducer parameters ( $L$ ,  $N$ ,  $d$ , and  $D$ ) on the sideleaking behavior.

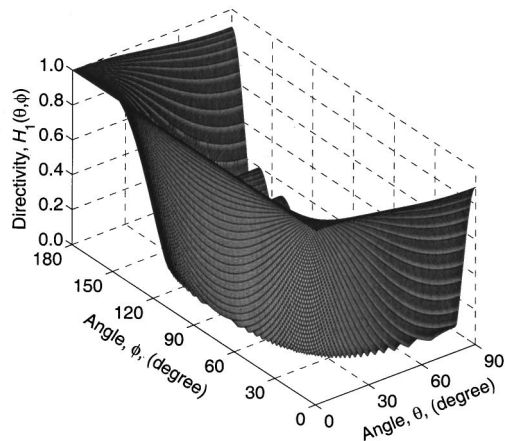
### A. Presentation of directivity functions

The directivities can be visualized by displaying three-dimensional plots in the forms of mesh plots, contour plots, and gray-scale image plots. As we will see, these presentations have their individual distinct merits in interpreting the physical meanings of directivities.

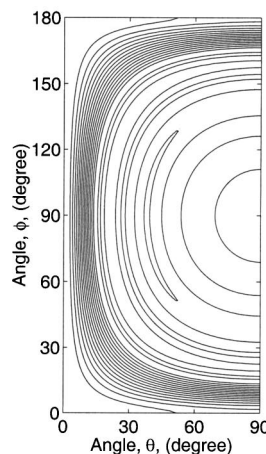
The first form of presentation is a three-dimensional mesh plot. For example, Fig. 3(a) shows the directivity as a function of azimuthal and rotational angles  $\theta$  and  $\phi$ . The respective parameters used to construct the plot are  $L=4\lambda$  and  $a=\lambda/4$  or  $L=16a$ . Note the general symmetry condition of the directivity

$$H(\theta, \phi) = H(-\theta, \phi - \pi). \quad (7)$$

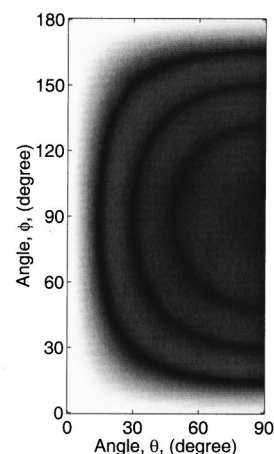
This means that the peaks in the range  $90^\circ < \phi < 180^\circ$  represent the directivity in the corresponding negative  $\theta$  direction, covering the entire  $180^\circ$  azimuthal and rotational angles in the plot. Since the directivity function exhibits these symmetry and antisymmetry in different octants of the coordinate system, it is sufficient to display only the first octant.



(a) Mesh plot with hidden line removal



(b) Contour plot

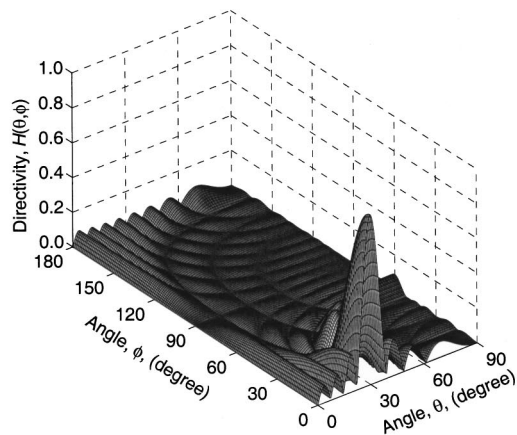


(c) Gray scale image plot

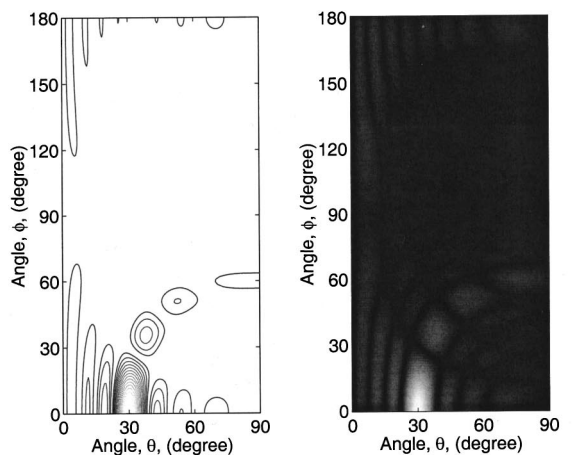
FIG. 3. Presentation schemes for visualizing three-dimensional directivity for a single element ( $L=4\lambda$ ,  $a=\lambda/4$ ).

This plot clearly shows the condition that  $H_1(0, \phi) = 1 = \text{const}$  for the plane represented by  $\theta=0$ . One can observe that the mesh plot cannot sufficiently display the variation of directivity along the azimuthal and rotational angles from this particular view angle. As a supplement, the function is presented by a contour plot shown in Fig. 3(b), where the lines represent the contour of the directivities of the same magnitude. Note that the values are not marked down on the plot because it will be otherwise too crowded. Finally, the same directivity can be displayed by an image plot shown in Fig. 3(c), where the directivity values are mapped into a gray-scale image.

Figure 4(a), (b), and (c) respectively represents the directivity mesh, contour, and image plots of the beam steered at  $\theta_s=30^\circ$  in the azimuthal direction, constructed using the transducer parameters of  $N=16$ ,  $d=\lambda/2$ ,  $L=4\lambda$ , and  $a=\lambda/4$ . Figure 5 shows the corresponding directivity  $H(\theta, 0)$  in the azimuthal plane, which is indeed a cross-sectional view of the three-dimensional directivity in the plane of  $\phi=0$  in Fig. 4(a). The main lobe appears at  $\theta=30^\circ$  and  $\phi=0^\circ$ , meaning that the beam is correctly steered at the angle and the beam steerability is good with these parameters. In addition, the main lobe (in the azimuthal plane) is reasonably narrow, which means that it has a good directivity as well.



(a) Mesh plot with hidden line removal



(b) Contour plot

(c) Gray scale image plot

FIG. 4. Presentation schemes for visualizing three-dimensional directivity for a phased array ( $N=16$ ,  $d=\lambda/2$ ,  $L=4\lambda$ ,  $a=\lambda/4$ , and  $\theta_s=30^\circ$ ).

The merit of the three-dimensional directivity analysis is that the beam propagation behavior in the directions other than the azimuthal plane is quantitatively predictable. For example, the directivity plot in Fig. 4 indicates that the beam leaks its energy into nonazimuthal planes ( $0 < \phi < \pi$ ) although the major beam propagates in the steering direction ( $\theta=30^\circ, \phi=0^\circ$ ). In this particular case, the main lobe has a

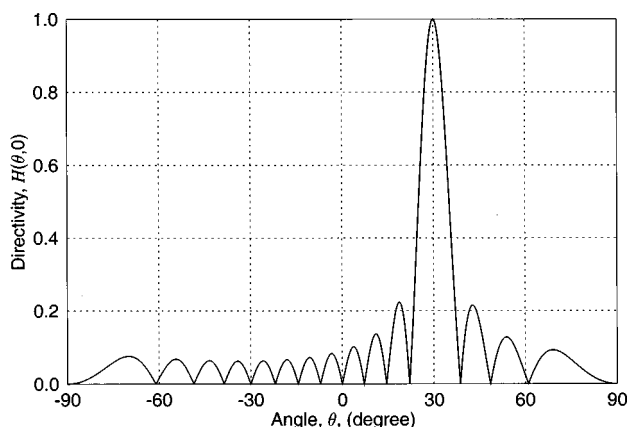
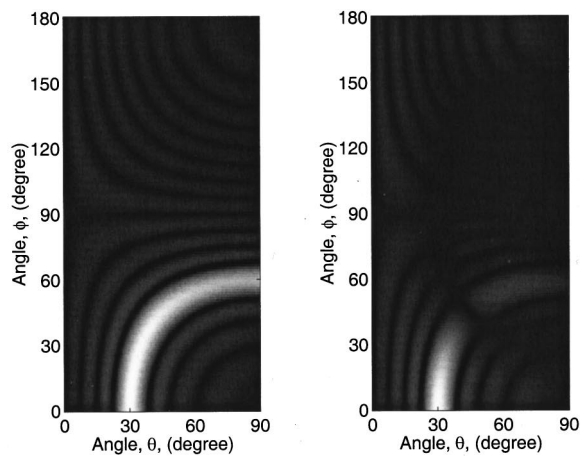
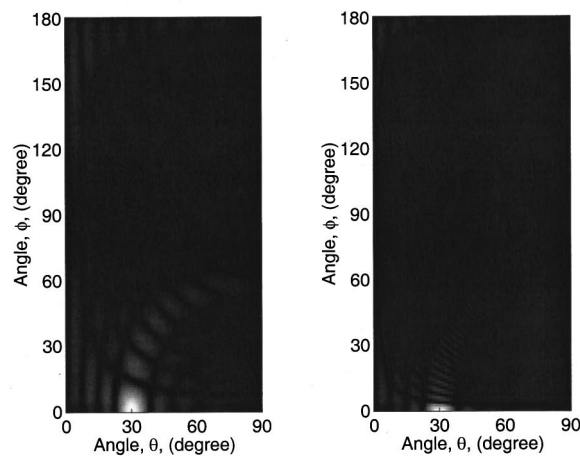


FIG. 5. Two-dimensional directivity of a phased array as a function of angle  $\theta$  in the azimuthal plane ( $N=16$ ,  $d=\lambda/2$ ,  $a=\lambda/4$ , and  $\theta_s=30^\circ$ ).



(a)  $L=2a$

(b)  $L=8a$



(c)  $L=32a$

(d)  $L=128a$

FIG. 6. Three-dimensional directivities showing the effect of elevation dimension  $L$  on the energy leaking for linear arrays ( $N=16$ ,  $d=\lambda/2$ ,  $a=\lambda/4$ , and  $\theta_s=30^\circ$ ).

certain width in the  $\phi$ -plane (up to approximately  $30^\circ$ ). What this implies is that the main lobe fans out within a sectorial angle determined by the width in the  $\phi$ -plane. This effect can also be visualized by the contour and image plots shown in Fig. 4(b) and (c). In the contour plots, densely packed lines mean higher sideleaking. In the image plot, it means that the larger the area of bright pixels, the more the energy leaks into the nonazimuthal directions. The sideleaking effect can thus be preferably reduced by suppressing the directivity in the range  $0 < \phi < \pi$ .

It should be noted here that each directivity plot is normalized by the maximum pressure in the plot, i.e., the pressure induced in the steering direction. Therefore, the directivity plots should not be directly compared with each other to calculate the energy field. These plots, however, can be used to show information about the pressure distribution and sideleaking patterns. It is desirable to compare the sideleaking energies for different parameters in a fashion as discussed below.

## B. Elevation dimension ( $L$ )

Figure 6 shows the directivity image plots for  $L=2a$ ,  $8a$ ,  $32a$ , and  $128a$ , ( $N=16$ ,  $d=\lambda/2$ ,  $a=\lambda/4$ , and  $\theta_s$

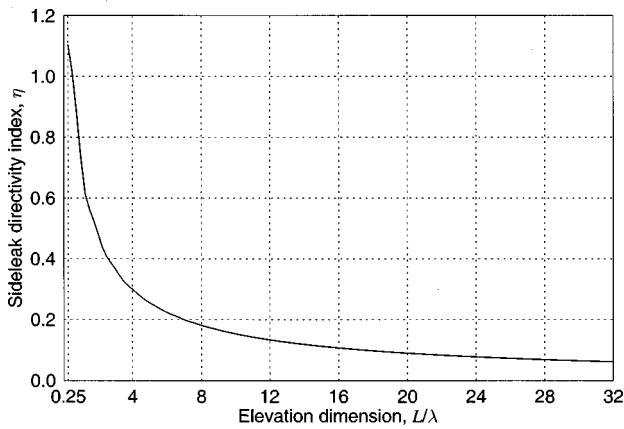


FIG. 7. Sideleak index as a function of elevation dimension ( $N=16$ ,  $d = \lambda/2$ ,  $a = \lambda/4$ , and  $\theta_s = 30^\circ$ ).

$=30^\circ$ ), from which the sideleaking effect can be observed. As the ratio  $L/a$  increases, the area of the bright zones decreases. For example, the white area of the image is the biggest for  $L=2a$  and it is the smallest for  $L=128a$ . This observation gives us a clue that the sideleaking energy can be reduced by enlarging the elevation dimension.

In order to check this assumption, a quantity called *sideleak directivity index* is introduced here, which is defined as the ratio

$$\eta = \frac{(1/2\pi) \int_0^{2\pi} \int_{-\pi/2}^{\pi/2} H(\theta, \phi) d\theta d\phi}{\int_{-\pi/2}^{\pi/2} H(\theta, 0) d\theta}. \quad (8)$$

The denominator represents the transmitting signal power in the azimuthal plane and the numerator is the average power over all directions. This index therefore represents the extent of energy propagating in the nonazimuthal directions relative to the one contained in the azimuthal direction. Note that a higher value of  $\eta$  means higher fraction of energy leaking into the sides of the transducer.

Figure 7 shows the influence of elevation dimension on the sideleaking behavior. The index  $\eta$  was obtained numerically and plotted as a function of elevation dimension  $L$ , for  $N=16$ ,  $d = \lambda/2$ ,  $a = \lambda/4$ , and  $\theta_s = 30^\circ$ . Note that denominator of Eq. (8) is not a function of  $L$  so that the numerator controls the behavior. This means that the directivity in the azimuthal plane is independent of  $L$ , so that the image plots not only show the distribution but also directly indicate sideleaking energy, in this case. As  $L$  approaches zero, the array can be considered as an ensemble of point sources, where each point acts as a source of spherical waves. Consequently, a significant portion of the energy propagates into unwanted directions and the array cannot perform its intended function of phase steering. Note that  $\eta > 1$  for very small  $L$  values, which means that more energy propagates in the nonazimuthal direction.

As  $L/\lambda$  reaches above 0.25, representing a square element ( $L=a$ ),  $\eta$  drops dramatically. This trend continues until  $L$  reaches four to eight times the wavelength and then slows down. As  $L$  approaches infinity, each element acts as an infinitely long line source so that the sideleak directivity index asymptotically approaches zero. It is thus desirable to increase the elevation dimension. However, there are some

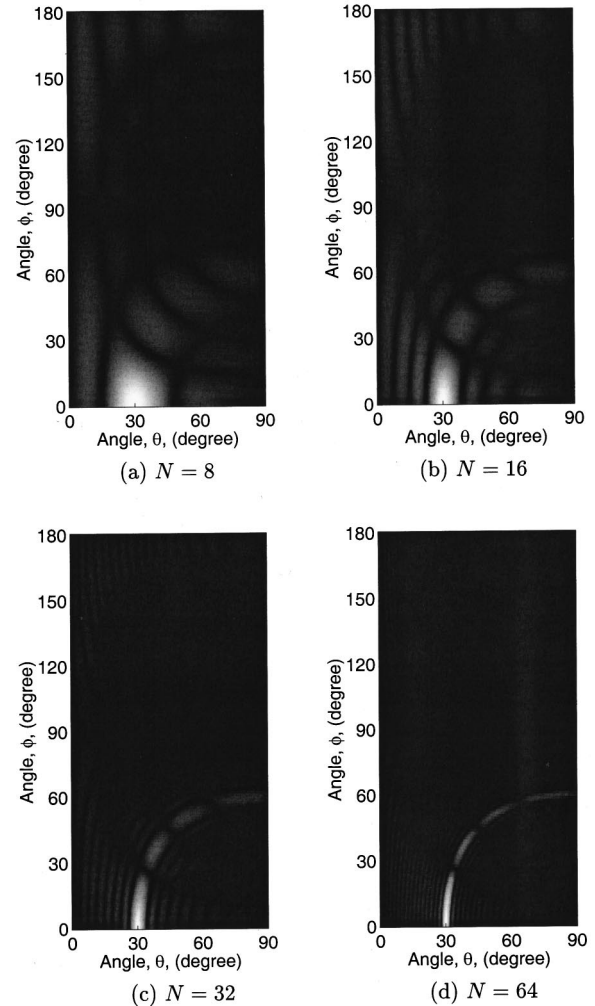


FIG. 8. Three-dimensional directivities showing the effect of number of elements  $N$  on the energy leaking for linear arrays ( $L=4\lambda$ ,  $d = \lambda/2$ ,  $a = \lambda/4$ , and  $\theta_s = 30^\circ$ ).

difficulties and disadvantages of making a transducer with large elevation dimension. Increased elevation dimension makes the transducer bulky and costly. Hence, from a practical transducer design stand point, it may be argued that a transducer whose elevation dimension  $L \geq 4\lambda = 16a$  (for  $\eta \approx 0.3$ ) or  $L \geq 8\lambda = 32a$  (for  $\eta \approx 0.19$ ) will perform reasonably well in terms of sideleaking.

### C. Number of elements

The number of elements  $N$  plays an important role in the transducer performance as discussed in our previous studies.<sup>10,11</sup> It was concluded that the main lobe width can be reduced by increasing the number of elements, resulting in improved and better directed beams propagating in the steered direction. It was also shown that the number of elements helps suppressing side lobes but has virtually no influence on grating lobes at all.

Here, we extend our interest to study the effect of  $N$  on the sideleaking energy. Figure 8 shows the respective three-dimensional directivities for  $N=8$ , 16, 32, and 64 ( $L=4\lambda$ ,  $d = \lambda/2$ , and  $a = \lambda/4$ ). Steered at  $30^\circ$ , maximum directivity occurs at  $\theta=30^\circ$  and  $\phi=0^\circ$ . As  $N$  increases from 8 to 64, the

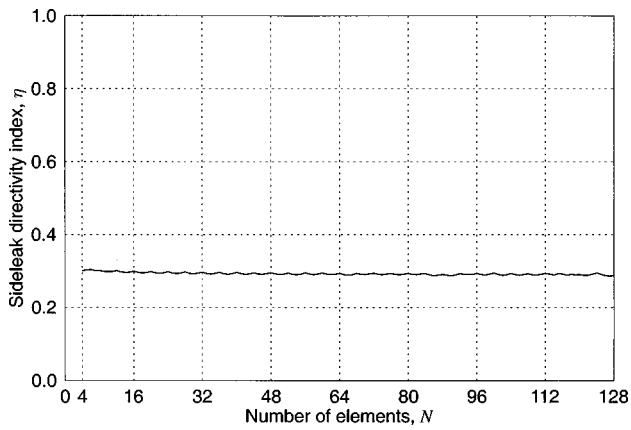


FIG. 9. Sideleak index as a function of number of elements  $N$  ( $L=4\lambda$ ,  $d=\lambda/2$ ,  $a=\lambda/4$ , and  $\theta_s=30^\circ$ ).

streak patterns (lobes) become much thinner. Increasing the number of elements reduces the lobe widths, resulting in a better directed beam, which is a preferable way. However, observe that the overall patterns remain similar for all cases, except that they get thinner. In all cases, the high intensity area in the neighborhood of the steering angle is extended in the  $\phi$  direction up to  $\phi=30^\circ$ , even for  $N=64$ . This means that some energy leaks out of the the azimuthal plane. This particular case can be pictured as a situation that the majority of the propagation energy is directed toward the steering direction but it fans out sharply within the approximate angle of  $\pm 30^\circ$ .

Figure 9 shows the relationship between  $\eta$  and  $N$  for the same parameter settings as given before. Interestingly, the index plot indicates that the sideleaking is almost independent of the number of elements. This plot, together with Fig. 8, implies that the average sideleaking energy decreases at almost the same rate as the one in the azimuthal plane, as  $N$  increases. Since the streak patterns shown in the three-dimensional directivity images are thinned without drastically changing their shapes, the average directivity changes at the same rate as the one in the azimuthal plane. The net result is the close index values for different  $N$  values. It should be reemphasized here that the image plots in Fig. 8 do not represent pressure distribution and should not be used to compare sideleaking energies between the cases. The total sideleaking energy in Fig. 8(d) is, in fact, bigger than that of Fig. 8(a) since higher pressure is induced by the contributions of more elements. It is thus concluded that higher number of elements yields narrow lobe widths and gives all positive influences on the steering behavior in the azimuthal plane without throwing in any negative influence on the sideleaking.

#### D. Interelement spacing

The interelement spacing  $d$  is also an important factor that critically influences the steering performance. It is known that increasing  $d$  improves the directivity in the azimuthal plane. However, if it reaches above a certain critical value, then grating lobes are introduced. These grating lobes

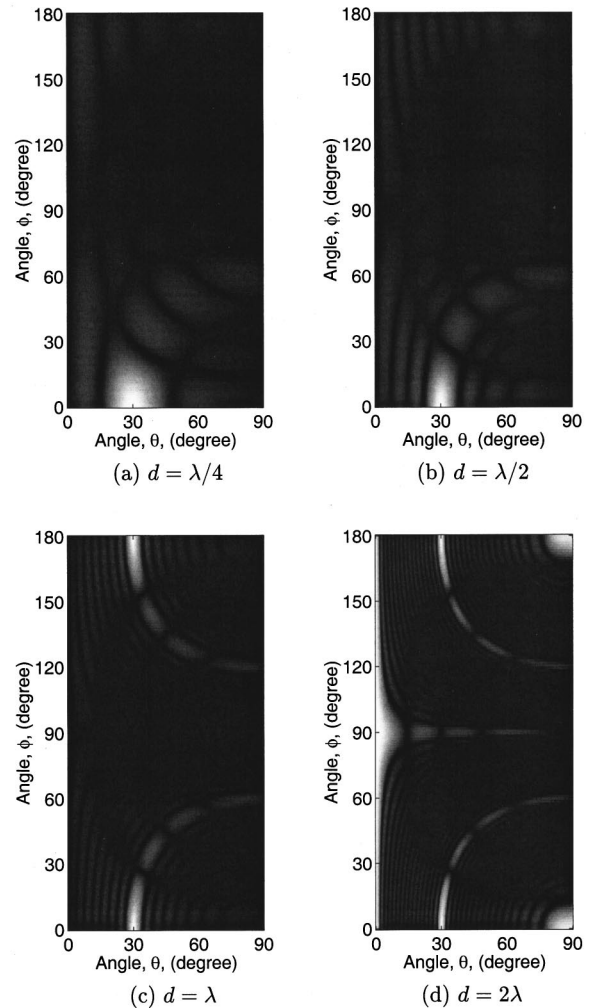


FIG. 10. Three-dimensional directivities showing the effect of interelement spacing  $d$  on the energy leaking for linear arrays ( $N=16$ ,  $L=4\lambda$ ,  $a=\lambda/4$ , and  $\theta_s=30^\circ$ ).

have the same magnitude as the main lobe and act detrimentally on the transducer performance. The critical value was found to be<sup>10</sup>

$$d_{\max} = \frac{\lambda}{1 + \sin \theta_s} \frac{N-1}{N}, \quad (9)$$

which guarantees the condition that completely eliminates the real-valued grating lobes out of the directivity plot.

Figure 10 shows the effect of interelement spacing on the directivity patterns. Three-dimensional image directivities are plotted for  $d=\lambda/4$ ,  $\lambda/2$ ,  $\lambda$ , and  $2\lambda$  ( $N=16$ ,  $L=4\lambda$ ,  $a=\lambda/4$ , and  $\theta_s=30^\circ$ ). By comparing the plot for  $d=\lambda/4$  and the one for  $d=\lambda/2$ , one can see that larger interelement spacing reduces the lobe widths.

As  $d$  keeps increasing above  $d_{\max}$ , which is the case for Fig. 10(c) and (d), the lobes become sharper, resulting in improved directivities. But, unfortunately, additional streak patterns appear in these images, which can be interpreted as grating lobes. For example, a grating lobe exists in Fig. 10(c) at  $(\theta, \phi)=(30^\circ, 0^\circ)$  or  $(-30^\circ, 0^\circ)$ , while the main lobe occurs at  $(30^\circ, 0^\circ)$ . In Fig. 10(d), four grating lobes exist in the azimuthal plane, standing at  $(0^\circ, 0^\circ)$ ,  $(30^\circ, 180^\circ)=(-30^\circ, 0^\circ)$ ,  $(90^\circ, 180^\circ)=(-90^\circ, 0^\circ)$ , and  $(90^\circ, 0^\circ)$ , respectively. The main

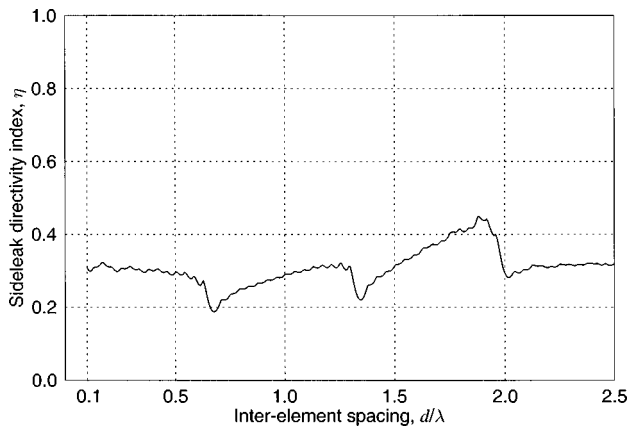


FIG. 11. Sideleak index as a function of interelement spacing  $d$  ( $N=16$ ,  $L=4\lambda$ ,  $a=\lambda/4$ , and  $\theta_s=30^\circ$ ).

lobe is still standing at the same position at  $(30^\circ, 0^\circ)$ . The line symmetry of the directivity plot about the  $\phi=90^\circ$  axis is introduced in Fig. 10(c) and (d) because integer multiples of the wavelength were chosen for  $d$ . In general, it is not symmetric.

Figure 11 shows the sideleak index as a function of interelement spacing. In the first region where  $0.1\lambda \leq d \leq d_{\max} = 0.625\lambda$ ,  $\eta$  stays almost constant. Since this is the region with no grating lobes, the streak patterns just get thinner for increasing  $d$ . As  $d$  passes the point  $d \approx d_{\max}$ , there is a sudden jump in magnitude. This phenomenon should be explained by the first grating lobe that starts appearing in the directivity space right at  $d_{\max}$ , whose zero-crossing point is located at  $\theta = -90^\circ$  in an image plot (not shown). As  $d$  increases, the grating lobe moves rapidly into the directivity plot. Consequently, the lobes continue to become sharper and  $\eta$  increases slowly until the next grating lobe appears. The same phenomenon repeats as  $d$  increases further. These grating lobes are deleterious and should be squelched by avoiding any spacing above  $d_{\max}$ . Within this range, the influence of  $d$  on the sideleaking is negligible.

### E. Aperture length

It has been shown that increasing  $N$  improves the overall steering properties. It was also observed that a larger interelement spacing, up to a certain value, provides good properties as well. It is thus quite natural to predict, from a designer's standpoint, that a transducer with an interelement spacing of  $d_{\max}$  and huge  $N$  will yield a well-performing beam. However, there are some drawbacks to this approach. The first reason why we should avoid too many elements in the array is that the transducer becomes bulky and costly. More importantly, increasing  $N$  (while fixing  $d$ ) also increases the overall aperture length  $D = (N-1)d + a$ , resulting in longer near-field length, and thus enlarges the dead zone of the transducer. Therefore, a compromise should be made.

One of our studies reveals an interesting fact that the main lobe width can be maintained by fixing the product  $Nd$ . In other words, the same aperture length ( $D \approx Nd$  for large  $N$  or  $d \approx a$ ) produces the same main lobe width. For example, doubling  $N$  and reducing  $d$  to half at the same time, and vice

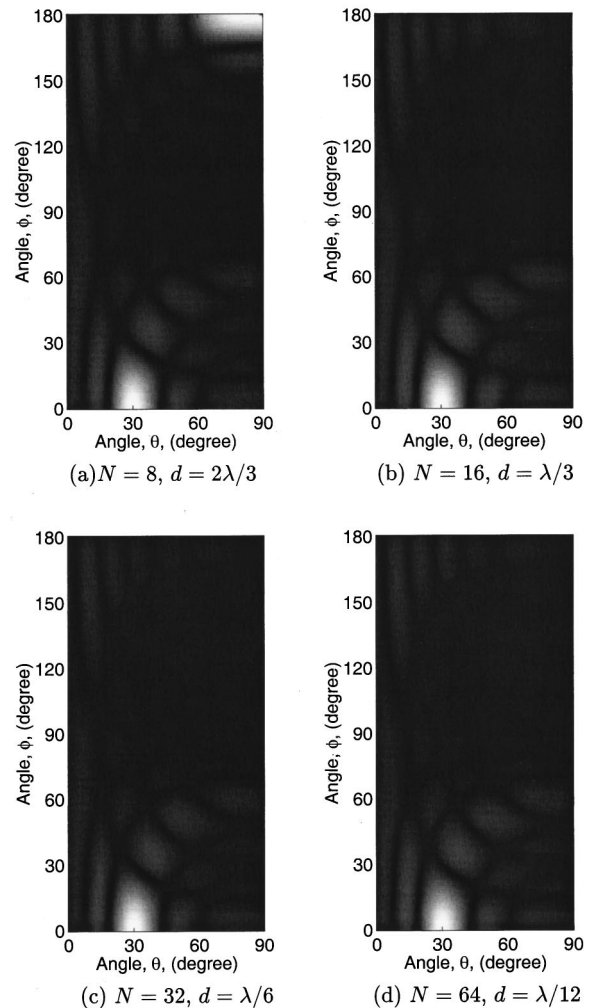


FIG. 12. Three-dimensional directivities showing the effect of aperture length  $D$  on the energy leaking for linear arrays ( $L=4\lambda$ ,  $a=\lambda/4$ , and  $\theta_s=30^\circ$ ).

versa, does not change the main lobe width. Of course, it should be done within the limiting condition  $d < d_{\max}$  to prevent the grating lobes. As a matter of fact, it was concluded from the study that the optimum design strategy is to use the largest possible number of elements which can be practically placed in the given aperture. We extend the analysis for the nonazimuthal directions in this paper.

Since both  $N$  and  $d$  negligibly affect  $\eta$ , it is quite obvious that the effect of the product  $Nd$  is also negligible. Figure 12 shows three-dimensional directivities for constant  $Nd$  combinations  $(N, d) = (8, 2\lambda/3)$ ,  $(16, \lambda/3)$ ,  $(32, \lambda/6)$ , and  $(64, \lambda/12)$ , i.e.,  $Nd = 16\lambda/3$ . It can be observed that the sideleaking energy of the first set is the largest of all. This can be again explained by the existence of a grating lobe since  $d$  is larger than its maximum allowable value  $d_{\max} = 0.625\lambda$ . If  $d$  is decreased below the critical value with an accordingly increasing  $N$ , the grating lobe disappears and apparently sideleaking energy drops.

It is interesting to observe that the directivities of the three cases in Fig. 12(b), (c), and (d), which represents regions where  $d$  is below  $d_{\max}$ , are exactly the same. This can be explained from Eq. (6) by noting that the numerator in the last term remains constant for fixed  $Nd$  while the denomina-



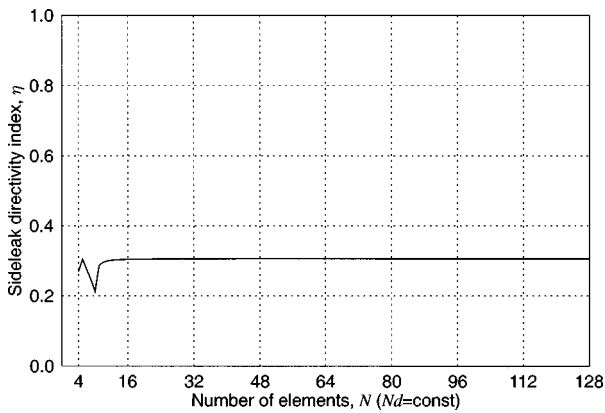


FIG. 13. Sideleak index as a function of number of elements  $N$  ( $L=4\lambda$ ,  $a=\lambda/4$ ,  $\theta_s=30^\circ$ , and fixed  $Nd=16\lambda/3$ ).

tor approaches unity for small  $d$ . Consequently, directivities are the same for constant  $Nd$  values. This phenomenon is further evidenced by Fig. 13 which relates the sideleaking with the number of elements for a fixed aperture length  $D \approx Nd=16\lambda/3$ . The index remains constant everywhere except some anomalies at low  $N$  values, which is again explained by the existence of grating lobes.

Therefore, we can make the same conclusion as before: it is desirable to use as many elements as possible for fixed aperture length, to provide a good directivity in the azimuthal plane. For  $d < d_{\max}$ ,  $N$  and  $d$  make little influences and  $Nd$  makes absolutely no influence on the leakage. In practical design, it is suggested to make arrays with  $N$  at least more than 16.

### III. SUMMARY AND CONCLUSIONS

Three-dimensional expressions of directivity function were obtained for single rectangular element sources and linear phased arrays, based on Huygen's principle which models the motion of waves radiated from multiple sources. The directivities were quantitatively analyzed to interpret the wave propagation behavior in the directions other than the azimuthal plane. The distribution was presented in various forms and the sideleaking directivity index was expressed in terms of geometrical transducer parameters including elevation dimension  $L$ , number of elements  $N$ , interelement spacing  $d$ , and the aperture length  $Nd$ .

It was shown that the directivity patterns of a single element and phased arrays in the azimuthal plane are equivalent to the corresponding two-dimensional analysis. In particular, it was verified that the beam directivity in the azimuthal plane is not affected by changing the elevation dimension. On the other hand, the directivities in the other directions are very much influenced by the elevation dimension. Increasing the elevation dimension can effectively reduce the sideleaking energy, and hence the beam is better

directed and confined in the azimuthal plane. For a finite dimension  $L$ , increasing the number of elements or interelement spacing improves the beam directivity in the steering angle but they have negligible influence on the sideleaking in an average sense. The interelement spacing should be chosen below the critical value and  $N$  should be also limited.

The effect of aperture was also studied and a design strategy was established. The following general design procedure can be suggested. First and foremost, determine the aperture length depending on the application requirements such as transducer frequency, inspection area, near-field length, etc. Then, the next step would be to determine the desired elevation dimension depending on the required suppression amount of sideleaking energy. Use as many elements as possibly fit into the given aperture length. Calculate the sideleaking energy and calibrate the transducer by checking the directivity pattern. Iterate for an optimal design set of transducer parameters.

### ACKNOWLEDGMENTS

This study was supported by the Korea Highway Corporation. We are grateful to Dr. Chang-Guen Lee, Program Manager, and Keon-Chang Cho, Director of Highway Research Institute of the KHC, for their encouragement and support.

- <sup>1</sup>B. D. Steinberg, *Principles of Aperture and Array System Design* (Wiley, New York, 1976).
- <sup>2</sup>D. K. Lemon and G. J. Posakony, "Linear array technology in NDE applications," *Mater. Eval.* **38**, 34–37 (1980).
- <sup>3</sup>O. T. von Ramm and S. T. Smith, "Beam steering with linear arrays," *IEEE Trans. Biomed. Eng.* **BME-30**, 438–452 (1983).
- <sup>4</sup>S. T. Smith, H. G. Pavy, Jr., and O. T. von Ramm, "High-speed ultrasound volumetric imaging system—Part I: Transducer design and beam steering," *IEEE Trans. on Ultrason. Ferroelectr. Freq. Control* **38**, 100–108 (1991).
- <sup>5</sup>S. T. Smith, H. G. Pavy, Jr., and O. T. Von Ramm, "High-speed ultrasound volumetric imaging system—Part II: Parallel processing and image display," *IEEE Trans. Ultrason. Ferroelectr. Freq. Control* **38**, 109–115 (1991).
- <sup>6</sup>D. H. Turnbull and S. F. Foster, "Beam steering with pulsed two-dimensional transducer arrays," *IEEE Trans. Ultrason. Ferroelectr. Freq. Control* **38**, 320–333 (1991).
- <sup>7</sup>G. R. Lockwood, P.-C. Li, M. O'Donnell, and F. S. Foster, "Optimizing the radiation pattern of sparse periodic linear arrays," *IEEE Trans. Ultrason. Ferroelectr. Freq. Control* **43**, 7–14 (1996).
- <sup>8</sup>G. R. Lockwood and F. S. Foster, "Optimizing the radiation pattern of sparse periodic two-dimensional arrays," *IEEE Trans. Ultrason. Ferroelectr. Freq. Control* **43**, 15–19 (1996).
- <sup>9</sup>J. Assaad and C. Bruneel, "Radiation from finite phased and focused linear array including interaction," *J. Acoust. Soc. Am.* **101**, 1859–1867 (1997).
- <sup>10</sup>S. C. Wooh and Y. Shi, "Optimum beam steering of linear phased arrays," *Wave Motion* **29**, 245–265 (1999).
- <sup>11</sup>S. C. Wooh and Y. Shi, "Influence of phased array element size on beam steering behavior," *Ultrasonics* **36**, 737–749 (1998).
- <sup>12</sup>P. M. Morse and K. U. Ingard, *Theoretical Acoustics* (McGraw-Hill, New York, 1968).

# Robust control of a sensorless bass-enhanced moving-coil loudspeaker system

Mingsian R. Bai and Hsinping Wu

*Department of Mechanical Engineering, National Chiao-Tung University, 1001 Ta-Hseuh Road, Hsin-Chu, Taiwan, Republic of China*

(Received 13 November 1998; accepted for publication 5 March 1999)

Moving-coil loudspeakers generally exhibit poor response in the low-frequency range because the diaphragms are unable to produce sufficient volume velocity. To alleviate the problem, this study focuses on enhancing the low-frequency performance of loudspeakers by means of modern control techniques. A self-sensing velocity observer is utilized for producing the cone velocity signal required by the controller. Feedback  $H_\infty$  robust control and feedforward  $H_2$  model matching control are employed to simultaneously achieve robust stabilization and tracking performance. The proposed controller is implemented using a combined digital signal processor and operational amplifier circuitry. © 1999 Acoustical Society of America. [S0001-4966(99)02806-4]

PACS numbers: 43.38.Ja, 43.38.Ar [SLE]

## INTRODUCTION

In general, moving-coil loudspeakers exhibit poor response in the low-frequency range because the speaker diaphragms are unable to produce sufficient volume velocity below the mechanical resonance frequency (Borwick, 1994). Insufficient bass content significantly affects the listening quality of audio systems. How to maintain a uniform acoustic output from loudspeakers at very low frequency is a difficult problem. One method of improving the low-frequency response is to increase the radius of the speaker. However, the increase in efficiency is not as great as might be expected, because the mass of the speaker also increases with radius. The low-frequency response can also be enhanced by reducing the stiffness of the suspension, thereby lowering the mechanical resonance frequency. However, if the stiffness of the mechanical system is excessively reduced, its displacement at low frequency becomes very large, which may lead to harmonic distortions resulting from displacement of the voice coil into nonuniform regions of the magnetic field. Although efficiency can also be improved by increasing the magnetic flux density in the air gap (Kinsler *et al.*, 1982), this would result in an undesirable decrease of low-frequency sensitivity. Another conventional approach is electronic compensation, where audio systems are equipped with equalizers to boost the bass output. In doing so, only the magnitude of the low-frequency response is increased, while the phase is distorted even further [unless a linear phase FIR (finite impulse response) digital filter is used]. In contrast to conventional equalizers, this study adopted a different approach of electronic compensation that seeks to increase the bass level without disturbing the phase response so that the waveform distortion is minimized. A very good collection of references on loudspeaker development in last 30 years can be found in Borwick (1994).

The majority of loudspeaker designers to date have focused on the mechanical aspects of loudspeakers (Colloms, 1991). However, an increasing number of researchers are recognizing the potential of using modern control methods to enhance the performance of loudspeakers. Radcliffe and

Gogate (1992) employed classical proportional control and velocity feedback to improve low-frequency performance of loudspeakers. Kuriyama and Furukawa (1989) used a feedforward least-mean-square (lms) algorithm to achieve the same purpose. Along the same line, this study attempts to enhance the low-frequency performance by means of modern control techniques, under the electromechanical constraints of the original system. It seeks to optimally increase the magnitude without altering the phase (within a pure delay) of the low-frequency response. To this end, a hybrid structure is employed in the control design. In particular, robust stabilization is achieved using an  $H_\infty$  robust feedback controller, whereas tracking performance is maintained via an  $H_2$  model matching feedforward controller. The feedback controller is implemented by using analog operational amplifier (OP) circuitry to avoid excessive delay that might destabilize the closed-loop system. The feedforward controller is implemented on a floating-point digital signal processor (DSP).

The proposed system is different from the other electronically compensated systems (Colloms, 1991) in that this system does not require sensors such as accelerometers or microphones. A self-sensing velocity observer circuit based on the idea of Okada *et al.* (1995) is used for producing the cone velocity signal.

Experimental results show that the proposed system has an enhanced frequency response in the low-frequency region and better tracking performance of time-domain signals than the uncompensated system.

## I. MOVING-COIL LOUDSPEAKER

### A. Modeling

In this section, a brief review of the model of moving-coil loudspeakers which is also similar to the Thiele-Small's model (Small, 1972) is given. The following definitions are used (Beranek, 1954):

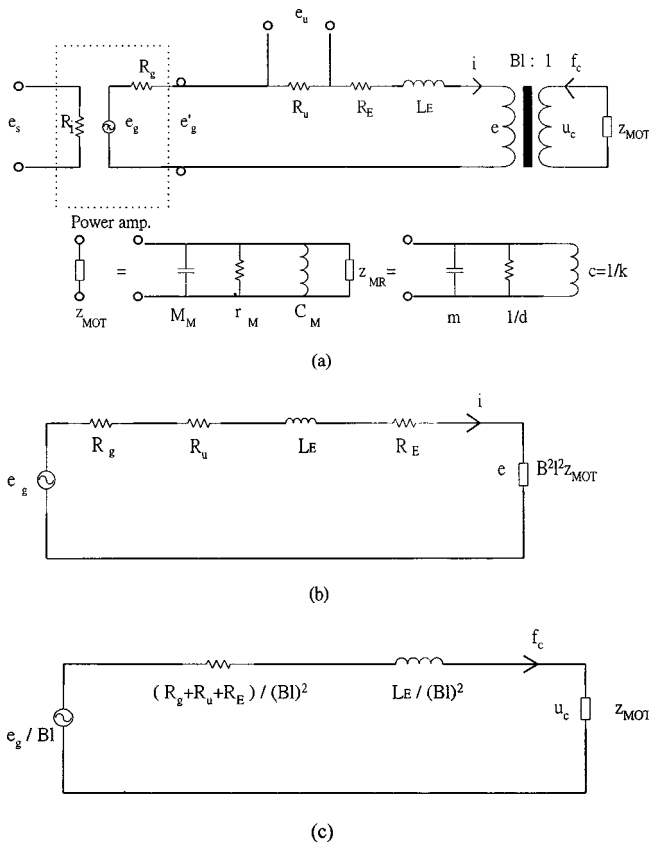


FIG. 1. Electromechanical analogous circuits. (a) Analogous circuit of the mobility type; (b) electrical circuit showing motional impedance; (c) analogous circuit of the mobility type referred to the mechanical side.

$e_s, R_i$	input voltage and input resistance of the power amplifier
$e_g, R_g$	open-circuit voltage and internal resistance of the generator
$e'_g$	output voltage of the power amplifier
$R_u$	resistance for the velocity observer
$e_u$	voltage drop across the resistor, $R_u$
$L_E, R_E$	inductance and resistance of the coil measured with the voice coil blocked ( $u_c = 0$ )
$Bl$	the electromagnetic coupling factor (magnetic flux density $\times$ coil length)
$e, i$	back electromotive force (EMF) and the current of the coil
$u_c, f_c$	coil velocity and Lorentz force
$M_M, C_M, r_M$	equivalent mass, compliance, and responsiveness of the mechanical system
$Z_{MR}$	radiation mobility, $Z_{MR} = (Z_{MR})^{-1}$ , $Z_{MR}$ being radiation impedance
$Z_{MOT}$	motional mobility of mechanical and acoustical systems
$m, d, k$	effective mass, damping, stiffness of the mechanical and acoustical systems

An equivalent circuit can be drawn by using the mobility analog, as shown in Fig. 1(a). Note that the resistance  $R_u$  is inserted in the circuit for measuring current of the coil. The motional mobility  $Z_{MOT}$  can be expressed in terms of equivalent mass ( $m$ ), damping ( $d$ ), and stiffness ( $k$ ):

$$Z_{MOT} = \frac{1}{j\omega m + d + k/j\omega}. \quad (1)$$

If the circuit of Fig. 1(a) is referred to the electrical side, as shown in Fig. 1(b), the current  $i$  can be calculated as

$$i = \frac{e_g}{R_t + j\omega L_E + (Bl)^2 Z_{MOT}}, \quad (2)$$

where the total resistance is defined as  $R_t \triangleq R_g + R_u + R_E$ . Assume that ( $|j\omega L_E| \ll R_t$ ) is in the low-frequency range and substitute Eq. (1) into (2):

$$\frac{i}{e_g} \cong \frac{(j\omega)^2 m + j\omega d + k}{(j\omega)^2 R_t m + j\omega(R_t d + B^2 l^2) + R_t k}. \quad (3)$$

Equation (3) is a second-order system with a resonance frequency  $\sqrt{k/m}$  and damping modified by the factor  $Bl$ . Note that a pair of lightly damped zeros are located at the same frequency  $\sqrt{k/m}$ . Conversely, if the circuit is referred to the mechanical side, as shown Fig. 1(c), the coil velocity  $u_c$  can be solved as

$$\begin{aligned} u_c &= \frac{e_g}{Bl} \left[ Z_{MOT} \left/ \left( \frac{R_g + R_u + R_E}{B^2 l^2} + j\omega \frac{L_E}{B^2 l^2} + Z_{MOT} \right) \right. \right] \\ &= e_g Bl \left/ \left[ R_t \left( \frac{1}{Z_{M1}} \right) + j\omega L_E \left( \frac{1}{Z_{M1}} \right) + B^2 l^2 \right] \right. \\ &= \frac{e_g Bl}{(R_t + j\omega L_E)(j\omega m + d + k/j\omega) + B^2 l^2}. \end{aligned} \quad (4)$$

The low-frequency approximation of coil velocity  $u_c$  can be obtained by neglecting the inductance  $L_E$

$$\frac{u_c}{e_g} \cong \frac{j\omega Bl}{(j\omega)^2 R_t m + j\omega(R_t d + B^2 l^2) + R_t k}. \quad (5)$$

Equation (5) represents a second-order system with a dc blocking zero and resonance frequency  $\sqrt{k/m}$ .

## B. Implementation of the velocity observer

In the far field, the sound pressure of a direct-radiator loudspeaker is related to the diaphragm velocity (Beranek, 1954). Hence, cone velocity is selected as the controlled variable in our design. However, direct access of cone velocity requires sensors such as accelerometers that may result in adverse effects of mass loading. A simpler solution is the *self-sensing velocity observer* (Okada *et al.*, 1995). From the electrical side of Fig. 1(a),

$$e_g = (R_t + j\omega L_E)i + Bl u_c. \quad (6)$$

Knowing that  $i = e_u / R_u$ , we can thus express the coil velocity  $u_c$  as

$$u_c = \frac{1}{Bl} \left[ e_g - \left( \frac{R_t + j\omega L_E}{R_u} \right) e_u \right]. \quad (7)$$

Hence, a velocity observer can be constructed based on the idea of Eq. (7), provided parameters  $R_u$ ,  $R_t$ ,  $L_E$ , and  $Bl$  have been measured. However, common calibration procedures (Beranek, 1954) that treat these parameters as ideal constants appeared insufficient for our purpose. We use a different approach to accommodate the frequency variation of the parameters. Rewrite Eq. (7) in terms of the output voltage of the power amplifier ( $e'_g$ )

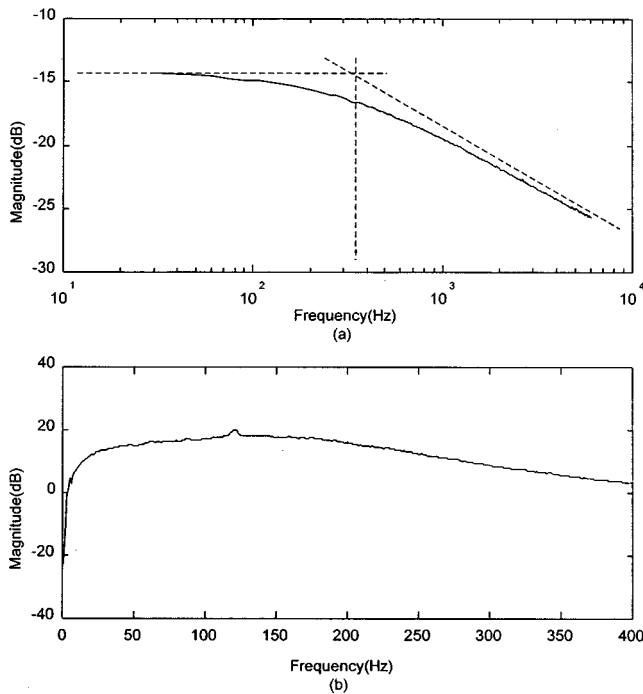


FIG. 2. Frequency responses of electrical parameters. (a) Bode plot of  $H(j\omega) = R_u / (R'_t + j\omega L_E)$ ; (b)  $1/Bl$ .

$$u_c = \frac{1}{Bl} \left[ e'_g - \left( \frac{R'_t + j\omega L_E}{R_u} \right) e_u \right], \quad (8)$$

where  $R'_t \triangleq R_u + R_E$ . Now, setting  $u_c = 0$  in Eq. (8) leads to

$$\frac{e_u}{e'_g} = \frac{R_u}{R'_t + j\omega L_E} = \frac{(R_u/R'_t)}{1 + j\omega/(R'_t/L_E)} \triangleq H(j\omega), \quad (9)$$

where  $H(j\omega)$  is the *blocked* frequency response (with the speaker diaphragm held still, i.e.,  $u_c = 0$ ) between the voltage drop across the current sensing resistor and the output voltage of the power amplifier. The function  $H(j\omega)$  is a first-order low-pass function with the corner frequency  $\omega_c = R'_t/L_E$  and dc level  $H_0 = R_u/R'_t$ . To satisfy the blocking condition, we simply place a hard rubber stopper firmly against the cone. With  $R_u$  pre-specified,  $R'_t$  and  $L_E$  can be identified from the Bode plot of  $H(j\omega)$ , i.e.,

$$R'_t = R_u/H_0 \quad \text{and} \quad L_E = R'_t/\omega_c. \quad (10)$$

The Bode plot of  $H(j\omega)$  obtained for our system is shown in Fig. 2(a). The remaining work is to identify the factor  $(1/Bl)$ . Dividing both sides of Eq. (8) by  $e_s$  gives

$$\frac{u_c}{e_s} = \frac{1}{Bl} \left[ \frac{e'_g}{e_s} - \left( \frac{R'_t + j\omega L_E}{R_u} \right) \frac{e_u}{e_s} \right]. \quad (11)$$

We then measure the coil velocity  $u_c$  using a Polytec OFV 2500 laser velocity sensor. From the frequency responses  $u_c/e_s$ ,  $e'_g/e_s$ ,  $e_u/e_s$ , and the identified value of  $(R'_t + j\omega L)/R_u$ , the factor  $(1/Bl)$  is determined, as shown in Fig. 2(b). This factor appears to be frequency dependent. This frequency dependence entails the need to treat the factor  $(1/Bl)$  in Eq. (8) as a filter so that the magnitude and phase of coil velocity can be accurately described. The resulting

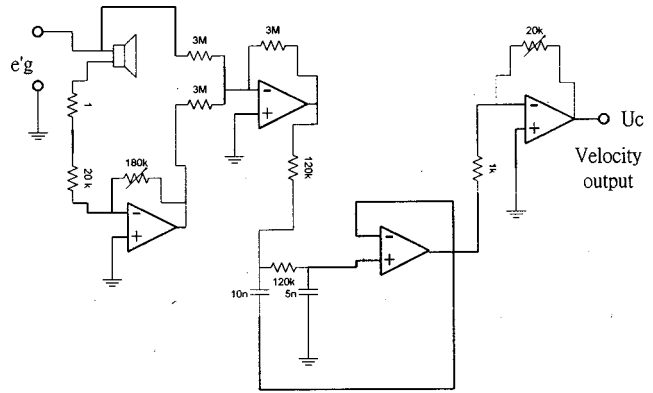


FIG. 3. OP circuit of the self-sensing velocity observer.

velocity observer in Eq. (11) is implemented by an OP circuit (Schaumann *et al.*, 1989), as shown in Fig. 3.

## II. ROBUST CONTROL DESIGN

The hybrid structure (Aström, 1990) composed of a feedforward controller and a feedback controller is adopted in the control design. The design strategy is first to find an  $H_\infty$  feedback controller that stabilizes the open-loop plant, where “plant” means “the controlled system.” The reason for using a feedback module is to increase robustness against plant uncertainties and perturbations so that the cone velocity is nearly constant (Morari and Zafirov, 1989). Next, a feedforward controller is introduced to achieve tracking performance without degrading the stability of the feedback-compensated system. It is noted that an optimally matched feedforward control is a step beyond merely using a linear phase FIR digital filter that does not take into account the phase response of the plant.

### A. $H_\infty$ robust feedback controller

The feedback structure of Fig. 4 is considered. To find an  $H_\infty$  controller, we weight the sensitivity function  $\tilde{S}(z)$  by  $W_1(z)$ , the control input  $u(k)$  by  $W_2(z)$ , and the complementary sensitivity function  $\tilde{T}(z)$  with  $W_3(z)$ , where

$$\tilde{S}(z) = \frac{1}{1 + P(z)C(z)} \quad (12)$$

and

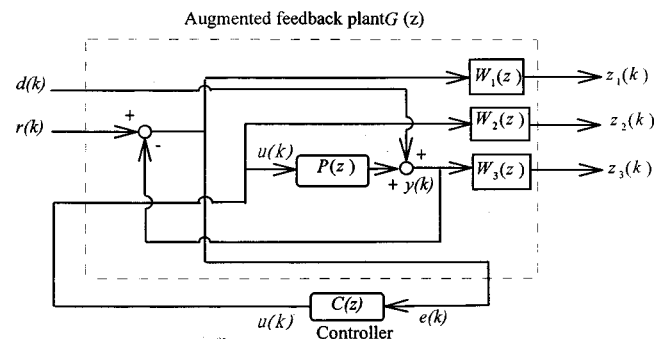


FIG. 4. System diagram of feedback control.

$$\tilde{T}(z) = \frac{P(z)C(z)}{1 + P(z)C(z)}. \quad (13)$$

For good disturbance rejection and tracking performance, the nominal performance condition must be satisfied (Doyle *et al.*, 1992)

$$\|\tilde{S}(z)W_1(z)\|_\infty < 1. \quad (14)$$

On the other hand, for system stability against plant perturbations, the robustness condition must be satisfied

$$\|\tilde{T}(z)W_3(z)\|_\infty < 1. \quad (15)$$

The tradeoff between  $\tilde{S}(z)$  and  $\tilde{T}(z)$  dictates the performance and stability robustness of the feedback design:

$$\| |\tilde{S}(z)W_1(z)| + |\tilde{T}(z)W_3(z)| \|_\infty < 1 \quad (16)$$

The input–output relation of the augmented plant of the feedback structure is

$$\begin{bmatrix} Z_1(z) \\ Z_2(z) \\ Z_3(z) \\ E(z) \end{bmatrix} = \begin{bmatrix} W_1(z) & -W_1(z)P(z) \\ 0 & W_2(z) \\ 0 & W_3(z)P(z) \\ 1 & -P(z) \end{bmatrix} \begin{bmatrix} D(z) \\ U(z) \end{bmatrix}, \quad (17)$$

where  $Z_1(z)$ ,  $Z_2(z)$ , and  $Z_3(z)$  are controlled variables, and  $P(z)$  is the open-loop plant. The extraneous input  $D(z)$  consists of the reference  $r(k)$  and disturbance  $d(k)$ . The signals  $U(z)$  and  $E(z)$  are the control input to the plant and the measured output from the plant, respectively. In  $H_\infty$  theory, the suboptimal condition of the feedback controller reads (Doyle *et al.*, 1992).

$$\left\| \begin{bmatrix} W_1(z)\tilde{S}(z) \\ W_2(z)\tilde{R}(z) \\ W_3(z)\tilde{T}(z) \end{bmatrix} \right\|_\infty < 1, \quad (18)$$

where

$$\tilde{R}(z) = \frac{C(z)}{1 + P(z)C(z)}. \quad (19)$$

In  $H_\infty$  control synthesis, the optimal tradeoff between performance and stability robustness is achieved by tuning the weighting functions  $W_1(z)$ ,  $W_2(z)$ , and  $W_3(z)$ . In general,  $W_1(z)$  and  $W_3(z)$  are chosen as a low-pass function and a high-pass function, respectively. The details of how to select the weighting functions can be found in Bai and Lin (1997).

### B. $H_2$ feedforward model matching controller

Having stabilized the plant  $G(z)$  by the feedback controller, the design effort can then be focused on finding a feedforward controller  $C(z)$  which makes the plant output track the desired output of a reference model  $M(z)$ . In our study,  $M(z)$  is chosen as the following function:

$$M(z) = \frac{z^{-10}(1.4695 - 0.0609z^{-1})}{1 - 0.5305z^{-1}}. \quad (20)$$

Note that the above function contains a pure delay term  $z^{-10}$  and a first-order low-pass function. The low-pass function is

to attenuate the excessive gain outside the control bandwidth. The pure delay, which will not introduce waveform distortion, is essential in calculating the controller using the model matching principle, and is detailed as follows.

The matching procedure is based on the  $H_2$  criterion. Let the squared error be

$$\epsilon_2^2 = \|M(z) - C(z)G(z)\|_2^2, \quad (21)$$

where the 2-norm of a transfer function is defined as

$$\|G(z)\|_2 \triangleq \left( \frac{1}{2\pi} \int_{-\pi}^{\pi} |G(e^{j\theta})|^2 d\theta \right)^{1/2}. \quad (22)$$

The feedforward model matching problem reduces to finding a stable transfer function  $C(z)$  to minimize  $\epsilon_2^2$ . We perform an *inner–outer factorization* (Doyle *et al.*, 1992) on  $G(z)$ :

$$G(z) = V_a(z)V_m(z), \quad (23)$$

with  $V_a(z)$  being an all-pass function and  $V_m(z)$  being a minimum phase function. Substitute Eq. (23) into Eq. (21) and omitting  $(z)$  for simplicity, we have

$$\begin{aligned} \epsilon_2^2 &= \|M - V_a V_m C\|_2^2 = \|V_a V_a^{-1} M - V_a V_m C\|_2^2 \\ &= \|V_a (V_a^{-1} M - V_m C)\|_2^2 = \|V_a^{-1} M - V_m C\|_2^2. \end{aligned} \quad (24)$$

In the last step, the fact that  $V_a$  has a bounded constant magnitude on the unit circle is used. Now, decompose  $V_a^{-1}M$  as follows:

$$V_a^{-1}M = (V_a^{-1}M)_+ + (V_a^{-1}M)_-, \quad (25)$$

where  $(V_a^{-1}M)_+$ , and  $(V_a^{-1}M)_-$  correspond to the unstable part and the stable part, respectively. Then Eq. (24) can be rewritten into

$$\begin{aligned} \epsilon_2^2 &= \|(V_a^{-1}M)_+ + (V_a^{-1}M)_- - V_m C\|_2^2 \\ &= \|(V_a^{-1}M)_+\|_2^2 + \|(V_a^{-1}M)_- - V_m C\|_2^2. \end{aligned} \quad (26)$$

The Pythagoras theorem in Hilbert space is used in the last equality. Thus the optimal  $C_{\text{opt}}(z)$  becomes

$$C_{\text{opt}}(z) = V_m^{-1}(V_a^{-1}M)_-. \quad (27)$$

### III. EXPERIMENTAL INVESTIGATIONS

The experimental setup is composed of a one-way closed-box woofer with an 8-in. moving-coil direct radiator speaker driven by a 30-W power amplifier, a mid-range speaker, and a tweeter driven by a 25-W power amplifier.

Prior to controller design, the model of the plant  $G(z)$  needs to be determined. One way of constructing the plant model is to identify all parameters in Eq. (5). Alternatively, one may use a system identification procedure to construct the model, based on the input data  $e_g$  and output data  $u_c$ . This paper chose the latter approach which may capture more dynamics overlooked by the low-frequency analytical model in Eq. (5). The comparison in Fig. 5 shows good agreement between the measured frequency response and the regenerated frequency response of the identified plant.

On the basis of the identified plant, the aforementioned  $H_\infty$  procedure and  $H_2$  procedure are applied to obtain the optimal feedback controller and feedforward controller, re-

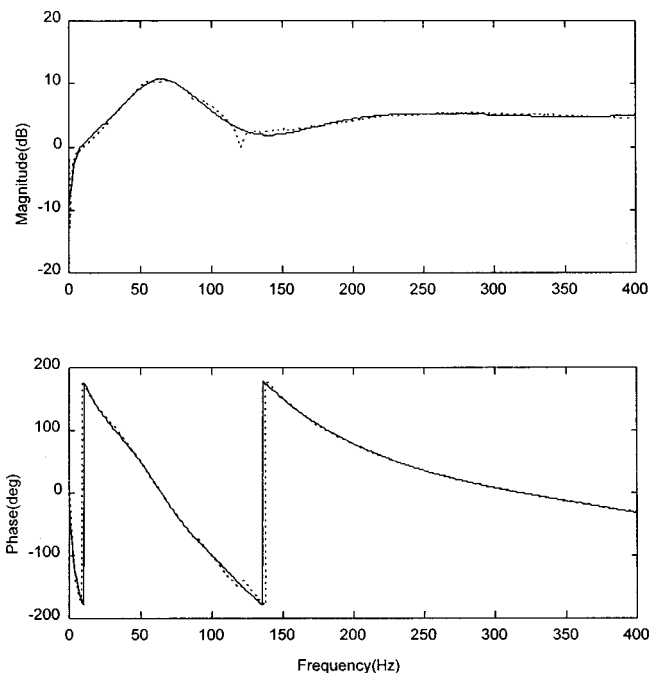


FIG. 5. Comparison between the measured and the regenerated frequency responses of the open-loop plant. Measured data —; regenerated data - -.

spectively. The feedback controller is implemented by an OP circuit to avoid excessive delay, whereas the feedforward controller is implemented by a digital filter on a 32-bit floating-point DSP (TMS320C31) with a sampling rate 2 kHz that is nearly five times of the control bandwidth, which meets the Nyquist sampling criterion. A two-channel 16-bit input/output daughter module is used. The board is equipped with anti-aliasing filters and smoothing filters. The dynamic range of quantization is 90 dB. The total system diagram is shown in Fig. 6. The frequency response of the feedback controller and the associated electric circuit of the analog controller are shown in Figs. 7 and 8, respectively. Prior to the feedforward controller design, the model of the feedback-compensated plant is determined via a frequency-domain

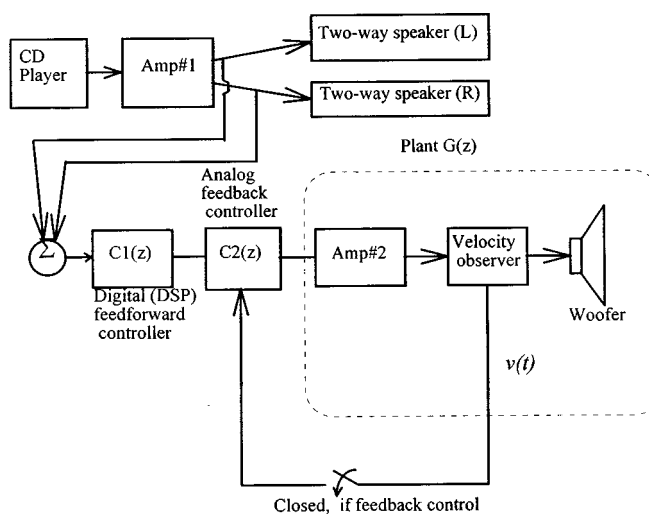


FIG. 6. Total system diagram of the bass-enhanced audio system.

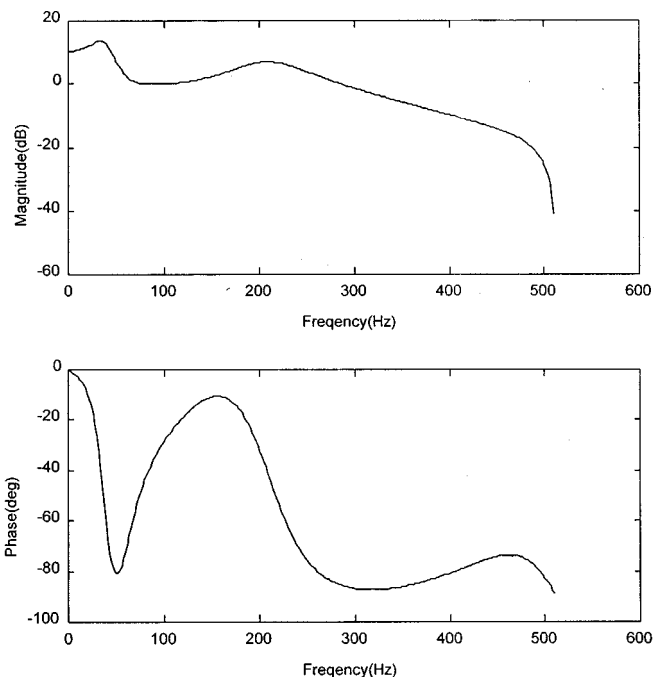


FIG. 7. Frequency response of the  $H_\infty$  feedback controller.

identification procedure (Juang, 1994). The measured frequency response and the regenerated frequency response of the compensated plant are compared in Fig. 9. By  $H_2$  optimization, the optimal feedforward controller is found (Fig. 10). As expected, the controller exhibits high gain in the low-frequency range.

The feedback controller and the feedforward controller are hence combined to enhance the bass quality of the audio system. Although coil velocity is the controlled variable in our design, a monitoring microphone is also set in the far field (20 cm away) to evaluate the acoustical performance. The frequency responses for the uncompensated system and the system compensated by the hybrid control are compared in term of the coil velocity and the far-field sound pressure (Figs. 11 and 12). The experimental result shows significant improvement of bass quality by driving the low-frequency limit of the loudspeaker from approximately 60 Hz down to 20 Hz. The magnitude enhancement ranges from 3 dB to 20 dB throughout the 400-Hz control bandwidth. In addition, the phase linearity is not degraded (with a pure delay) because of the electronic compensation. Note that there is a hump around 10 dB. This peculiar result is done on purpose

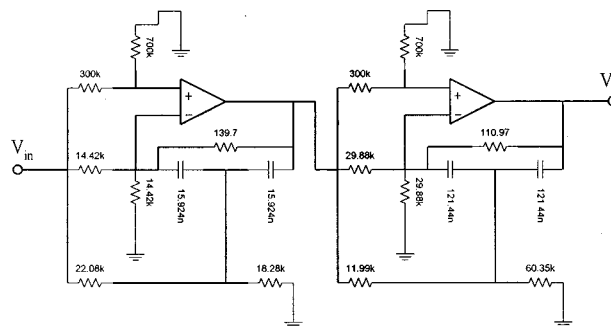


FIG. 8. Circuit diagram of the  $H_\infty$  feedback controller.

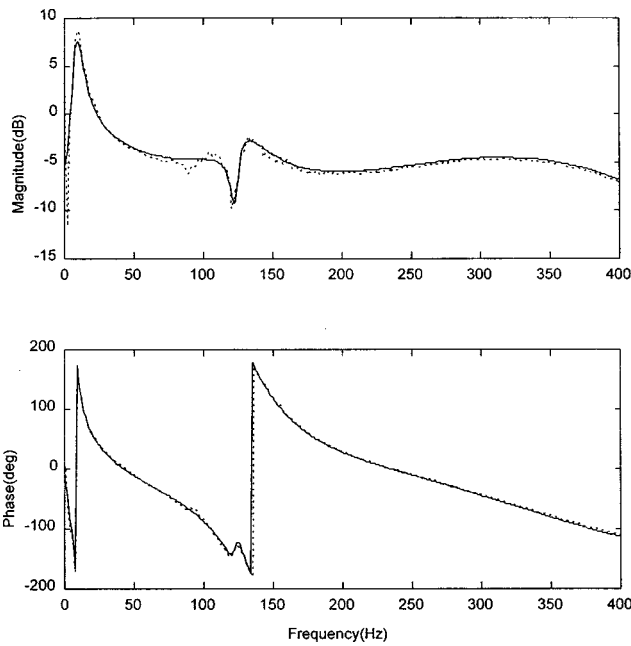


FIG. 9. Comparison between the measured and the regenerated frequency responses of the feedback-compensated plant. Measured data — — —; regenerated data —.

by giving more weight to the low-frequency response of the feedforward controller. After a subjective evaluation, we felt a boosted bass would give a more impressive sense of “super bass” than just a flat response. To test the practicality of the proposed system, a pop music sample is used as the command signal to evaluate the tracking performance of the hybrid controller. The time-domain responses for the uncompensated system and the compensated system are compared in Fig. 13. It is evident from the comparison that the proposed system indeed produces more satisfactory tracking performance.

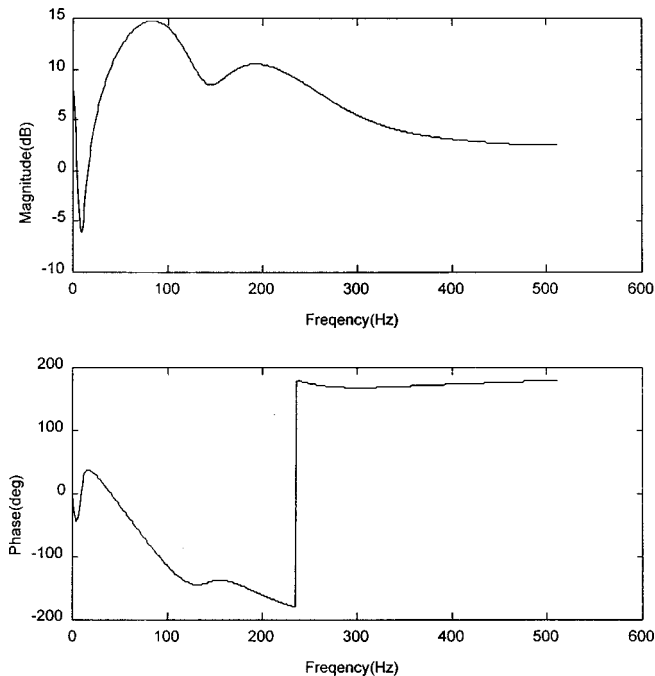


FIG. 10. Frequency response of the  $H_2$  feedforward controller.

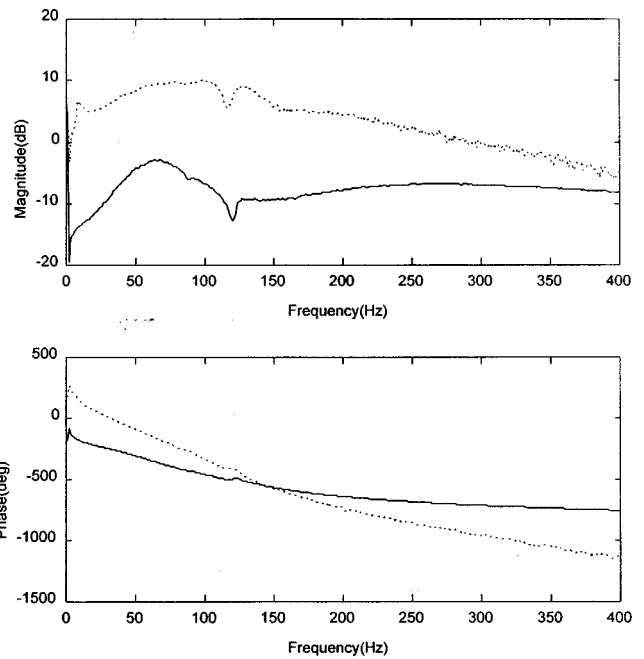


FIG. 11. Frequency responses between a white noise input and the cone velocity output for the uncompensated system and the compensated system. Control off —; control on — — —.

#### IV. CONCLUSIONS

Modern control techniques are exploited to enhance the low-frequency performance of moving-coil loudspeakers, under the electromagnetic properties and acoustical constraints. A self-sensing velocity observer is developed for cone velocity estimation without additional motional or acoustical sensors.  $H_\infty$  feedback control is employed for robust stabilization, while the  $H_2$  feedforward model matches

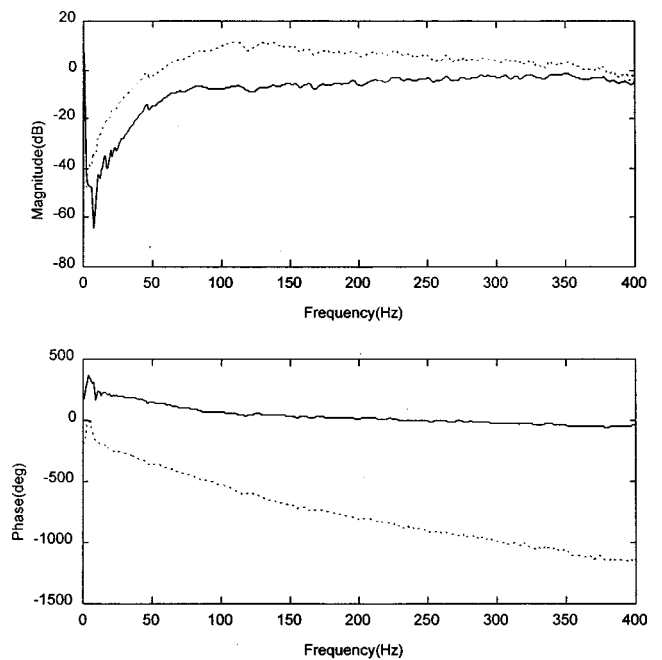


FIG. 12. Frequency responses between a white noise input and the far-field sound pressure output for the uncompensated system and the compensated system. Control off —; control on — — —.

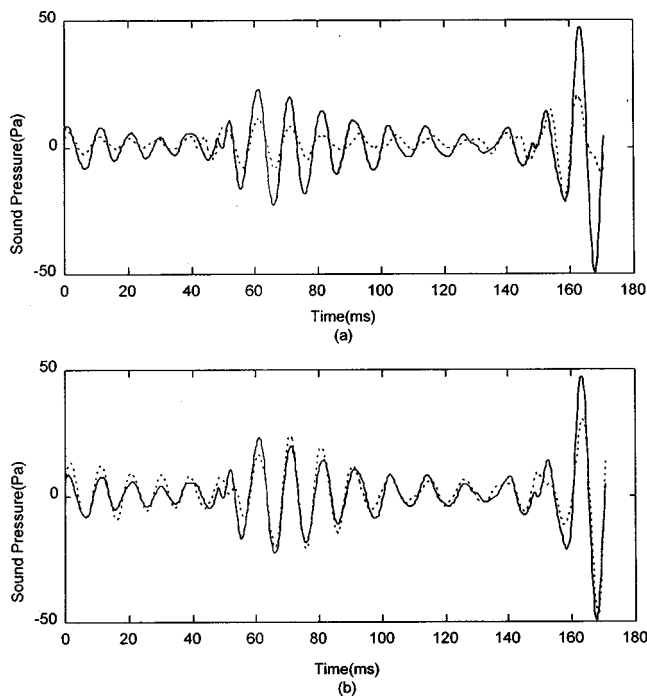


FIG. 13. Time-domain tracking performance of the bass-enhanced system for the pop music signal. (a) Command signal versus sound pressure output signal before bass compensation; (b) Command signal versus sound pressure output signal after bass compensation. Command signal — output signal

control for tracking performance. The results obtained in experiments indicate that the proposed system yields improved performance over the uncompensated one. However, as a limitation of the proposed methodology, the success of this method relies on adequate design of the original mechanical system and acoustical system (such as a sufficiently large diameter of the speaker). That is, one can never adequately control a poorly designed mechanical system. Overemphasis on the proposed electronic compensation will likely result in undesirable nonlinearity in the system.

Although this paper mainly focuses on audio loudspeakers, the same rationale can be extended to the other applica-

tions, e.g., control speakers for active noise cancellation, linear electromagnetic actuators for active vibration control and isolation, where efficient low-frequency response is crucial. Since the research was originally targeted at the subwoofer, only one driver was tested. However, the feasibility of the proposed technique applied to the systems of multiple drivers should be examined. Future research is planned in these areas.

## ACKNOWLEDGMENTS

We wish to thank Dr. William Thompson, Jr. of Applied Research Laboratory, Penn State University for providing class notes on acoustic transducers. The work was supported by the National Science Council in Taiwan, Republic of China, under the Project No. NSC 86-2212-E-009-003.

- Aström, J., and Wittenmark, B. (1990). *Computer-Controlled Systems* (Prentice-Hall, Englewood Cliffs, NJ).
- Bai, M. R., and Lin, H. H. (1997). "Comparison of active noise control structures in the presence of acoustical feedback by using the  $H_\infty$  synthesis technique," *J. Sound Vib.* **206**, 453–471.
- Beranek, L. L. (1954). *Acoustics* (McGraw-Hill, New York).
- Borwick, J. (1994). *Loudspeaker and Headphone Handbook* (Butterworth-Heinemann, Oxford).
- Colloms, M. (1991). *High Performance Loudspeakers* (Wiley, New York).
- Doyle, J. C., Francis, B. A., and Tannenbaum, A. R. (1992). *Feedback Control Theory* (Macmillan, New York).
- Juang, J. N. (1994). *Applied System Identification* (Prentice-Hall, Englewood Cliffs, NJ).
- Kinsler, L. E., Frey, A. R., Coppens, A. B., and Sanders, J. V. (1982). *Fundamentals of Acoustics* (Wiley, New York).
- Kuriyama, J., and Furukawa, Y. (1989). "Adaptive loudspeaker system," *J. Audio Eng. Soc.* **37**, 919–926.
- Okada, Y., Matsuda, K., and Hashitani, H. (1995). "Self-sensing active vibration control using the moving-coil-type actuator," *ASME J. Vib. Acoust.* **117**, 411–415.
- Morari, M., and Zafiroiu, E. (1989). *Robust Process Control* (Prentice-Hall, Englewood Cliffs, NJ).
- Radcliffe, J., and Gogate, S. D. (1992). "Identification and modeling of speaker dynamics for acoustic control application," *ASME J. Vib. Acoust.* **38**, 295–300.
- Schaumann, R., Ghausi, M. S., and Laker, K. R. (1989). *Design of Analog Filters: Passive, Active RC, and Switched Capacitor* (Prentice-Hall, Englewood Cliffs, NJ).
- Small, R. H. (1972). "Closed-box Loudspeaker Systems, Part I: Analysis," *J. Audio Eng. Soc.* **20**, 798–808.



# Propagation characteristics of leaky Lamb waves in a liquid-loaded double-layered substrate consisting of a thin piezoelectric ceramic plate and a thin glass plate

Kohji Toda and Kenji Motegi

*Department of Electronic Engineering, National Defense Academy, Hashirimizu, Yokosuka 239-8686, Japan*

(Received 16 March 1998; revised 5 February 1999; accepted 22 February 1999)

Operational performance of a leaky wave interdigital transducer on a double-layered substrate composed of two thin plates, a piezoelectric ceramic plate and a glass plate, is described for the purpose of two kinds of transducer functions—one for the effective radiation of underwater ultrasound and the other for liquid sensing. Two parameters, electromechanical coupling coefficient ( $k^2$ ) and mode conversion ratio ( $C$ ), are evaluated on a two-dimensional  $p$ - $fd$  plane as the fundamental factors for the leaky wave transducer, where  $p$  is the thickness ratio of the ceramic part and the double-layered substrate, and  $fd$  is the product of the frequency ( $f$ ) and the total thickness of the substrate ( $d$ ). For an effective transducer for underwater ultrasound, both  $k^2$  and  $C$  should be large at a certain operation point of  $p$  and  $fd$ . The maximum transducer efficiency from the product of  $k^2$  and  $C$  in the second mode operation in the case of  $p=0.37$  is three times as large as that in the case of a single-layered piezoelectric ceramic plate. For constructing a concise liquid sensing device, on the other hand, a combination of large  $k^2$  and small  $C$  is required. A condition satisfying the above requirement is recognized in the third mode operation in the case of  $p=0.55$  and  $fd=3.4$  MHz·mm. The combination of a piezoelectric ceramic plate and a glass plate is promising for realizing highly functional ultrasound transducers for underwater ultrasound and for liquid sensing.  
© 1999 Acoustical Society of America. [S0001-4966(99)00406-3]

PACS numbers: 43.38.Rh, 43.20.Hq, 43.20.Jr [SLE]

## INTRODUCTION

Elastic wave propagation is essentially an important problem to be solved in constructing various ultrasonic devices. A high-performance transducer is required for various underwater ultrasound fields such as medical diagnosis, non-destructive testing, and physical acoustics.<sup>1-5</sup> A thickness-drive transducer is generally used for the excitation of the longitudinal wave into the liquid.<sup>6,7</sup> An interdigital transducer (IDT) is widely used as the most fundamental component in surface acoustic wave (SAW) technology.<sup>8,9</sup> The IDT is also able to operate effectively as a leaky wave transducer at a liquid–solid interface.<sup>10</sup> A leaky SAW along the liquid-loaded surface can be applied to transducers for underwater ultrasound.<sup>11-13</sup>

A leaky Lamb wave transducer using a thin piezoelectric plate has merits such as multiple-mode operation and frequency-dependent radiation direction of a longitudinal wave in a liquid.<sup>14</sup> Propagation characteristics of the leaky Lamb waves in a single-layered substrate have been studied both numerically and experimentally.<sup>15-21</sup> This type of transducer, on the other hand, is mechanically weak. Reference 22 dealt with the velocity dispersion of Lamb waves in an aluminum (nonpiezoelectric) plate coated with a polymer placed in air.

In this paper, we describe the propagation characteristics of leaky Lamb waves and transducer performance, in a layered substrate consisting of two thin plates, one piezoelectric ceramic thin plate and one glass thin plate. The following two parameters, electromechanical coupling coefficient ( $k^2$ ) of the IDT and mode conversion ratio ( $C$ ) from the leaky

Lamb wave to the longitudinal wave in the liquid, are fundamental factors for considering the device performance. Both  $k^2$  and  $C$  are calculated from the complex leaky Lamb wave velocities in the form of a function of  $fd$ , where  $f$  is the frequency and  $d$  is the total thickness of the double-layered substrate. For an efficient underwater ultrasound transducer, both  $k^2$  and  $C$  are desired to be large. From the viewpoint of a liquid sensor application, on the other hand, a combination of large  $k^2$  and small  $C$  is required, because of low propagation loss of the leaky Lamb wave under the existence of the loading liquid. Compared with the case of a device configuration using a single-layered substrate, the values of  $k^2$  and  $C$  in the double-layered substrate depend not only on  $fd$  but also on  $p$ , which could lead to the increase of functional combination. Selection of the substrate material is essentially important for transducer performances although the use of a single-layered piezoelectric plate is, in general, for Lamb wave devices.

As far as we are aware, there has been no report on leaky Lamb waves propagating in layered-structured substrates with various thickness ratios in terms of the two parameters,  $k^2$  and  $C$ . When a double-layered substrate is used, the thickness ratio would be a key parameter for broadening the choice of the transducer performance. For a certain thickness ratio, both  $k^2$  and  $C$  take their respective peak values at almost the same value of  $fd$ , which enables us to construct an underwater ultrasound transducer with high efficiency. For another thickness ratio,  $C$  takes a small value at an  $fd$  giving the peak of  $k^2$ , which enables us to design a liquid sensing device operating under a liquid-loaded condition. In this study, a layered substrate consisting of a thin piezoelectric

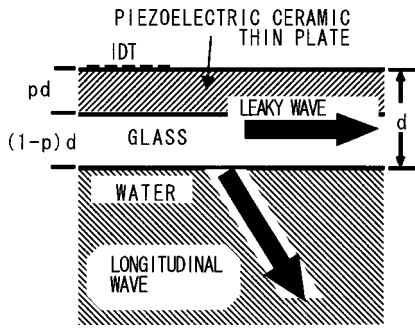


FIG. 1. Schematic diagram of propagation directions of a leaky Lamb wave in a double-layered substrate and a longitudinal wave radiated into a liquid.

ceramic plate and a thin glass plate, a typical combination of well-known materials, is introduced and the thickness ratio of the ceramic to the total thickness is varied from 0 to 1. Substitution of the glass with other materials could broaden the choice of the transducer configuration.

### I. CONSTRUCTION OF LAYERED SUBSTRATE

Figure 1 shows schematically the directions of two kinds of waves, a leaky Lamb wave propagating in a double-layered substrate and a longitudinal wave in a liquid. The substrate is composed of two different materials, a piezoelectric ceramic thin plate and a glass thin plate, which are adhered to each other. The ceramic is 101 A produced by TDK Co., which has a poling axis in the thickness direction. The glass plate is No. 7059 produced by Corning Co. The total thickness of the double-layered substrate is  $d$ . The thickness of the piezoelectric plate is  $p \cdot d$ , where  $p$  is a parameter with a value between 0 and 1. Hence it follows that the thickness of the glass plate is  $(1-p) \cdot d$ . The surface of the piezoelectric plate on the glass side is electrically shorted and the surface of the opposite side is electrically opened or shorted. The free surfaces of the glass plate and the piezoelectric plate are exposed to the liquid or air, respectively.

The efficiency of leaky Lamb wave excitation in the substrate, evaluated by  $k^2$ , can be obtained from the following relation:<sup>23</sup>

$$k^2 = 2 \frac{v_0 - v_s}{v_0}, \quad (1)$$

where  $v_0$  and  $v_s$  are the phase velocities of the leaky Lamb waves when the free surface of the piezoelectric plate is electrically opened and shorted, respectively. While propagating in the  $x$ -direction, the leaky Lamb waves are mode converted to the longitudinal wave in the liquid. The mode conversion efficiency  $C$  is evaluated by the following relation:

$$C = 1 - \exp\left(4\pi \frac{v_I}{v_R}\right), \quad (2)$$

where  $v_R$  and  $v_I$  are the real and imaginary parts of the complex leaky Lamb wave velocity, respectively.  $C$  is the ratio of mode-converted energy from the leaky Lamb wave to the longitudinal wave in water while the leaky Lamb wave propagates in the  $x$ -direction by one wavelength.

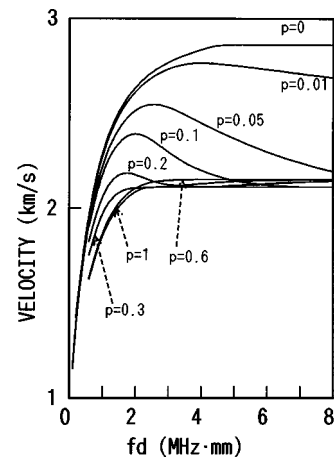


FIG. 2. Phase velocity dispersion curves of the zeroth mode leaky Lamb wave propagating in layered-structured substrates with various thickness ratios.  $p=0$  corresponds to the single layer of glass.

### II. NUMERICAL CALCULATION RESULTS

The complex velocity of the leaky Lamb wave was obtained by developing Farnell's method.<sup>24,25</sup> We focus on the shift of the dispersion curves for increasing  $p$ . Figure 2 shows the dispersion curves of the zeroth ( $A_0$ ) mode for the  $p$  values of 0, 0.01, 0.05, 0.1, 0.2, 0.3, 0.6, and 1. In these cases, both surfaces of the ceramic plate is electrically shorted. When  $p$  is small ( $p=0.01, 0.05, 0.1$ , and  $0.2$ ), the velocity curve initially goes up with  $fd$  and takes the maximum value at a certain  $fd$  value and then approaches the Rayleigh wave velocity of the piezoelectric ceramic. These figures explain the behavior of  $p=0$  in terms of the transition from the thin piezoelectric layer to the nonpiezoelectric layer.

Figure 3 shows the curves of the first ( $S_0$ ) mode for  $p$  in the range of 0–1, in steps of 0.1. The velocity values of the first mode at  $fd=0$  change monotonically with  $p$ . These velocities correspond to the respective longitudinal wave velocities in the thin layered substrates with different  $p$  values, because the particle displacement in the substrate is confined in the propagation direction of the leaky Lamb wave, while the particle displacements for other  $fd$  values are on oval

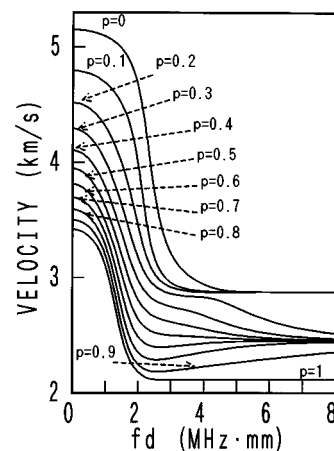


FIG. 3. Phase velocity dispersion curves of the first mode leaky Lamb wave propagating in layered-structured substrates with various thickness ratios.

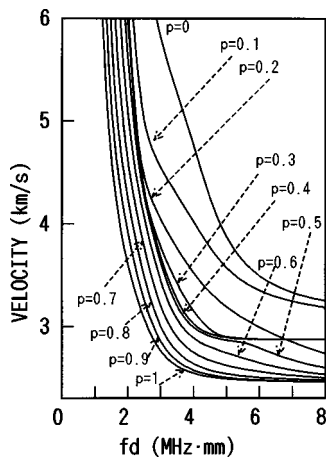


FIG. 4. Phase velocity dispersion curves of the second mode leaky Lamb wave propagating in layered-structured substrates with various thickness ratios.

tracks in the plane containing the propagation directions of both leaky Lamb wave and the longitudinal wave. The velocities of the first mode approach the transverse wave velocity of the piezoelectric ceramic for  $p$  in the range of 0.3–0.9 with the increase of  $fd$ , while they approach the Rayleigh wave velocity of the glass in the cases of  $p=0, 0.1$ , and  $0.2$ . The coupling between the first and second propagation modes causes an abrupt transition between the two curves for  $p=0.2$  and  $p=0.3$ .<sup>26</sup> In the case of  $p=1$ , the first mode velocity approaches the Rayleigh wave velocity of the piezoelectric plate.

Higher order modes have their respective cutoff frequencies corresponding to the  $p$  value and their velocities decrease as  $fd$  increases. Figures 4 and 5 show the velocity dispersion curves of the second and third modes, respectively, for various  $p$ . For a certain  $p$  value between 0 and 1, the double-layered substrate can be regarded as a single-layered substrate whose elastic properties are between those of the glass and piezoelectric ceramic plates, as far as the velocity dispersion curves are concerned. By changing  $p$ , we can broaden the choice of the material properties of the substrate.

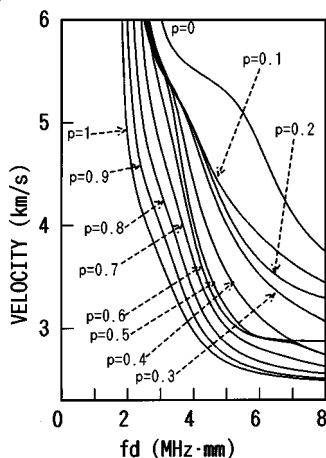


FIG. 5. Phase velocity dispersion curves of the third mode leaky Lamb wave propagating in layered-structured substrates with various thickness ratios.

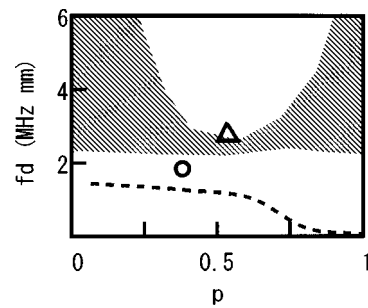


FIG. 6. Large  $C$ -valued region (shaded for  $C>0.2$ ) and trace of maximum  $k^2$  on a  $p$ - $fd$  plane, for the first mode operation. The circle and the triangle on the plane correspond to the operation points of two kinds of prepared transducer configurations,  $p=0.37$  and  $p=0.55$ , respectively.

Figures 6 through 8 show the calculated results of  $k^2$  and  $C$  on a two-dimensional plane ( $p$ - $fd$  plane) in the cases of the first, second, and third modes, respectively. The shaded domains in the figures represent the combinations of  $p$  and  $fd$  where the calculated  $C$  is over 0.2, a value that is considered large enough as a mode conversion ratio. The thick dashed curves show the  $fd$  values giving the peak values of  $k^2$  for respective  $p$ . In designing a transducer for underwater ultrasound, both  $k^2$  and  $C$  must be large. Figures 6 and 8 say that the first and third mode operations of the transducer are not promising for any combinations of  $p$  and  $fd$ . In Fig. 7, however, the dashed curve crosses one of the shaded domains, indicating that the layered substrates with the thickness ratios from 0.23 through 0.4 are promising for transducers operating in the second mode with high  $k^2$  and high  $C$ , if the interdigital periodicity of the IDTs are properly selected so that  $fd$  of the operation point gives the peak value of  $k^2$ . The maximum product of  $k^2$  and  $C$  is about three times as large as that in the case of a transducer using a single layer of the ceramic plate. As for an application to a liquid sensor having a combination of high  $k^2$  and low  $C$ , the three figures indicate that there are many combinations of  $p$  and  $fd$  which give favorable transducer configurations as liquid sensing devices.

### III. EXPERIMENTAL RESULTS AND DISCUSSIONS

Two devices were prepared using double-layered substrates with two different  $p$  values. One used a bilayer substrate with the total thickness of  $630 \mu\text{m}$  and the  $p$  value of

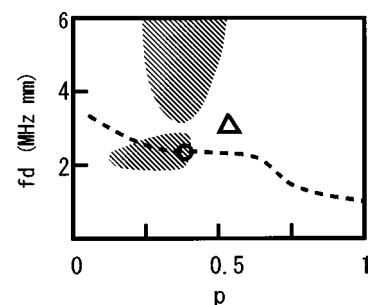


FIG. 7. Large  $C$ -valued region (shaded for  $C>0.2$ ) and trace of maximum  $k^2$  on a  $p$ - $fd$  plane, for the second mode operation. The circle and the triangle on the plane correspond to the operation points of two kinds of prepared transducer configurations,  $p=0.37$  and  $p=0.55$ , respectively.

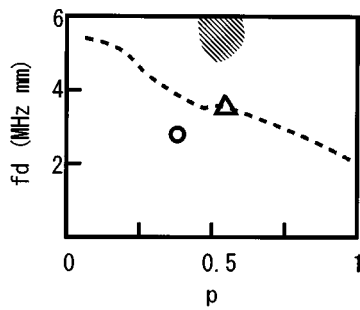


FIG. 8. Large  $C$ -valued region (shaded for  $C > 0.2$ ) and trace of maximum  $k^2$  on a  $p$ - $fd$  plane, for the third mode operation. The circle and the triangle on the plane correspond to the operation points of two kinds of prepared transducer configurations,  $p = 0.37$  and  $p = 0.55$ , respectively.

0.37. Two arch-shaped IDTs with the periodicity of  $1200 \mu\text{m}$ , one for the input and the other for the output, were deposited on the free surface of the ceramic. The IDT, composed of three pairs of electrode fingers, had an aperture angle of  $45^\circ$ . The operation points of this transducer in the first, second, and third modes are shown by solid circles in Figs. 6–8, respectively. This transducer was designed so that it can most effectively operate in the second mode. The phase velocities at the operation points of the transducer are plotted by open squares in Fig. 9, which shows the velocity dispersion curves of the leaky Lamb waves in the bilayer substrate with the  $p$  value of 0.37.

When an rf electrical pulse is applied to one of two arch-shaped IDTs on the double-layered substrate surface, a leaky Lamb wave is excited at the part of the input IDT and is immediately mode converted to a longitudinal wave in a liquid (water). After being reflected by a placed brass plate via the water layer, the wave arrives at another transducer to be detected as a delayed electrical signal. Figure 10 shows the observed input and output signals for the second mode operation at the carrier frequency of 3.7 MHz, where an output amplitude of 170 mV is observed for an input of 5 V. It should be noted that the transducer effectively radiates the leaky wave and detects the delayed signal after the propagation in the form of a longitudinal wave in water even for the electrode finger pairs of 3. The center frequencies of the transducer are 2.7, 3.7, 4.4, and 5.4 MHz, corresponding to the first, second, third, and fourth modes, respectively. The

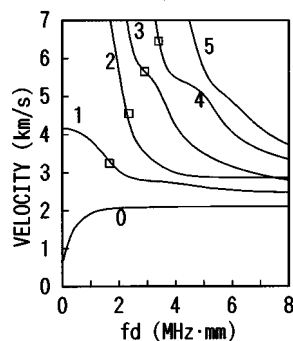


FIG. 9. Phase velocity dispersion curves of the six lowest order modes of the leaky Lamb waves propagating in the layered substrate with the  $p$  value of 0.37. Open squares show the operation points of the transducer configuration with the interdigital periodicity of  $1200 \mu\text{m}$ .

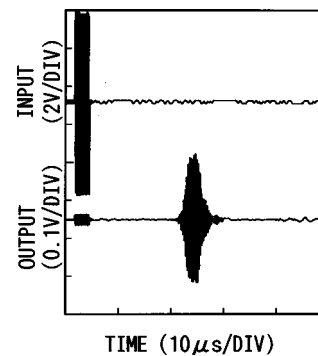


FIG. 10. Observed waveforms of input and output signals for the second mode at the carrier frequency of 3.7 MHz, in the case of  $p = 0.37$  and  $d = 630 \mu\text{m}$ .

experimental results say that the insertion losses of each mode are 41.1, 29.4, and 30.0 dB for the first, second, and fourth mode, respectively. No output was observed for the third mode. The above results are well consistent with the calculated product of  $k^2$  and  $C$ .

The other device configuration used a double layer with the total thickness of  $900 \mu\text{m}$  and the thickness ratio of 0.55. The interdigital periodicity of the arch-shaped IDTs with three electrode finger pairs were  $1050 \mu\text{m}$ . The operation points of this transducer are indicated by triangles in Figs. 6–8, which do not satisfy the condition required for underwater ultrasound transducers. In the third mode operation, however, the leaky Lamb wave can be detected effectively without much attenuation accompanied by the radiation into the water layer, as shown in Fig. 11. This agrees with the calculated results in Fig. 8, corresponding to the combination of large  $k^2$  and small  $C$ .

In this case, the leaky Lamb wave mode propagates in the layered substrate, of which the velocity is considered to be dependent on the material properties of the loading liquid. The calculated velocity change of the third mode leaky Lamb wave in the layered substrate with  $p = 0.55$  is represented in Fig. 12 as a function of the concentration of the salified water as a loading liquid. The velocity change could be observed in the form of the change of the delay time between the input and output transducers. This indicates the possibility of constructing a liquid sensing device.

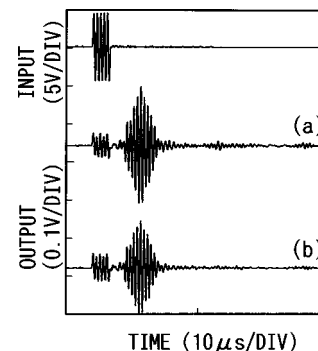


FIG. 11. Observed waveforms of input and output signals for the third mode at the carrier frequency of 3.8 MHz, in the case of  $p = 0.55$  and  $d = 900 \mu\text{m}$ , where the free surface of the glass plate is in contact with air (a) and water (b).

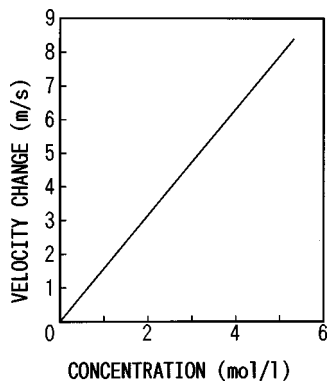


FIG. 12. Calculated relationship between the concentration of salified water and the change in the phase velocity of the third mode leaky Lamb wave in a double-layered substrate with  $p=0.55$ , under the condition of  $f=3.8$  MHz and  $d=900$   $\mu\text{m}$ .

#### IV. CONCLUSIONS

We have described propagation characteristics of leaky Lamb waves in a double-layered substrate composed of two kinds of thin plates, ceramic and glass, from the viewpoint of leaky wave transducers. Change of the thickness ratio between the two layers strongly influences the velocity dispersion curves, which makes it possible to realize wide combinations of  $k^2$  and  $C$  values. The favorable configuration in using the double-layered substrate is related to the selection of the thickness ratio of the two thin plates as well as the interdigital periodicity. The operational performances of this type of leaky wave transducer are quite different from the other type of transducer configurations without the use of a glass thin plate. The introduction of a glass thin plate would be useful for realizing a unique transducer for nondestructive testing and liquid sensing.

- <sup>1</sup>D. P. Morgan, "Surface acoustic wave devices and applications," *Ultrasonics* **11**, 121 (1973).
- <sup>2</sup>H. Wohltjen, "Surface acoustic wave microsensors," *Tech. Digest, 4th Int. Conf. Solid-State Sensors and Actuators (Transducer '87)*, Tokyo, Japan, June 2–5, pp. 471–477 (1987).
- <sup>3</sup>K. Toda and H. Urabe, "A hydrophone with a liquid delay line oscillator," *J. Appl. Phys.* **53**, 4072–4074 (1982).
- <sup>4</sup>K. Toda and A. Sawaguchi, "An ultrasonic observation of the heat stress in a glass plate by means of a leaky SAW transducer," *Proc. 3rd Int. Conf. Residual Stresses, Residual Stresses—III, Science and Technology*, Vol. 2, pp. 1115–1120 (1992).
- <sup>5</sup>T. Nomura, S. Shiokawa, and T. Moriizumi, "Measurement and Mapping of anisotropy of solid using a leaky SAW excited by an interdigital transducer," *IEEE Trans. Sonics Ultrason.* **32**, 235–240 (1985).

- <sup>6</sup>N. Murayama and H. Obara, "Piezoelectric polymers and their applications," *Jpn. J. Appl. Phys.* **22**, 3–6 (1982).
- <sup>7</sup>H. Takeuchi, S. Jyoumura, and C. Nakaya, "Highly anisotropic piezoelectric ceramics and their application in ultrasonic probes," *Proc. IEEE Ultrason. Symp.*, San Francisco, CA, USA, Oct. 16–18, pp. 605–608 (1985).
- <sup>8</sup>R. F. Milsom, M. Redwood, and N. H. C. Reilly, "The interdigital transducer," in *Surface Wave Filters*, edited by Herbert Matthews (Wiley, New York, 1977), Chap. 2.
- <sup>9</sup>R. M. White, "Surface acoustic wave sensors," *Proc. IEEE Ultrason. Symp.*, San Francisco, CA, USA, Oct. 16–18, pp. 490–494 (1985).
- <sup>10</sup>K. Toda and Y. Murata, "Acoustic focusing device with an interdigital transducer," *J. Acoust. Soc. Am.* **62**, 1033–1035 (1977).
- <sup>11</sup>G. W. Farnell and C. K. Jen, "Microscope lens using conversion of Rayleigh to compressional waves," *Proc. IEEE Ultrason. Symp.*, Boston, MA, Nov. 5–7, pp. 673–676 (1980).
- <sup>12</sup>K. Kobayashi, T. Moriizumi, and K. Toda, "Longitudinal acoustic wave radiated from an arched interdigital transducer," *J. Appl. Phys.* **52**, 5386–5388 (1981).
- <sup>13</sup>T. Nomura and T. Yasuda, "A new method of exciting focused surface acoustic waves using an interdigital transducer," *Jpn. J. Appl. Phys.* **26**, 109–111 (1986).
- <sup>14</sup>K. Toda, "An acoustical focusing device with an interdigital transducer on a thin piezoelectric plate," *Proc. IEEE* **67**, 1173–1174 (1979).
- <sup>15</sup>A. Sawaguchi and K. Toda, "Acoustical imaging using a leaky Lamb wave transducer," *Jpn. J. Appl. Phys.* **30**, 191–193 (1991).
- <sup>16</sup>H. Überall, B. Hosten, M. Descamps, and A. Gerard, "Repulsion of phase velocity dispersion curves and the nature of plate vibration," *J. Acoust. Soc. Am.* **96**, 908–916 (1994).
- <sup>17</sup>J. Dickey, G. Maidanik, and H. Überall, "The splitting of dispersion curves for the fluid-loaded plate," *J. Acoust. Soc. Am.* **98**, 2365–2367 (1995).
- <sup>18</sup>A. H. Nayfeh and P. B. Nagy, "Excess attenuation of leaky Lamb waves due to viscous fluid loading," *J. Acoust. Soc. Am.* **101**, 2649–2658 (1997).
- <sup>19</sup>W. Hassan and P. B. Nagy, "Circumferential creeping waves around a fluid-filled cylindrical cavity in an elastic medium," *J. Acoust. Soc. Am.* **101**, 2496–2503 (1997).
- <sup>20</sup>O. Poncelet and M. Deschamps, "Lamb waves generated by complex harmonic inhomogeneous plane waves," *J. Acoust. Soc. Am.* **102**, 292–300 (1997).
- <sup>21</sup>W. Hassan and P. B. Nagy, "Why fluid loading has an opposite effect on the velocity of dilatational waves in thin plates and rods," *J. Acoust. Soc. Am.* **102**, 3478–3483 (1997).
- <sup>22</sup>J. Laperre and W. Thys, "Experimental and theoretical study of Lamb waves dispersion in aluminum/polymer bilayers," *J. Acoust. Soc. Am.* **94**, 268–278 (1993).
- <sup>23</sup>S. G. Joshi and Y. Jin, "Electromechanical coupling coefficients of ultrasonic Lamb waves," *J. Acoust. Soc. Am.* **94**, 261–267 (1993).
- <sup>24</sup>G. W. Farnell, *Elastic Surface Waves, Surface Wave Filters*, edited by Herbert Matthews (Wiley, New York, 1977), Chap. 1, p. 26.
- <sup>25</sup>K. Toda and K. Ibuki, "Analysis of performance of leaky Lamb wave transducer operating at liquid-solid interface," *Ferroelectrics* **73**, 419–430 (1987).
- <sup>26</sup>K. Motegi and K. Toda, "Interaction between two leaky Lamb wave modes propagating in a water-loaded bilayer consisting of a piezoelectric ceramic plate and a glass plate," *Ultrasonics* **37**, 141–147 (1999).

# Modal characteristics of in-plane vibration of rectangular plates

N. H. Farag and J. Pan

*Department of Mechanical and Materials Engineering, The University of Western Australia, Nedlands, Western Australia 6907, Australia*

(Received 17 February 1998; accepted for publication 12 February 1999)

The equations of in-plane vibration of thin plates are solved for rectangular panels with two parallel edges clamped, while the other two edges are either both clamped, both free, or one clamped and one free. The propagation characteristics of in-plane waves are investigated and the frequency bands of attenuation and propagation of the waves are identified. Simple expressions for the cutoff frequencies of different wave components are derived. The nature of the coupling between in-plane longitudinal and in-plane shear waves is mathematically and physically illustrated. Although all the modes in a plate panel are coupled through Poisson's effect, it is shown that the coupling is very weak in plate panels with all edges clamped as compared to the cases where one or two parallel edges are free. As a first order approximation, the resonant modes of in-plane waves are classified into uncoupled and coupled mode pairs. By using the coupled mode pairs approximation, it is possible to predict the dynamic response of the plate panel with a reasonable accuracy. Simple expressions are given for approximate estimation of the resonance frequencies of coupled and uncoupled modes. Mode shapes are given, for each case of edge conditions, which satisfy both the displacement and force conditions at the plate edges. A simple procedure is given for the determination of resonance frequencies and mode shapes without excessive computations. The predicted resonance frequencies and mode shapes are compared to the finite element results and good agreement is found. The mode shapes of the in-plane vibration are depicted for the first eight resonant modes for each case of edge conditions. © 1999 Acoustical Society of America. [S0001-4966(99)03305-6]

PACS numbers: 43.40.Dx [CBB]

## INTRODUCTION

The prediction of in-plane dynamic response and resonance frequencies of flat plate panels is of practical importance. This is because in most applications the forces transmitted to platelike structures, even if they are designed to be perpendicular to the middle plane of the plate panel, are practically inclined, and have in-plane components due to imperfections in the manufacturing, assembly, or alignment of the supporting mounts. In some applications the main force components are designed to act parallel to or in the middle plane of the plate.

The resonance frequencies and mode shapes for flexural vibration are well documented.<sup>1,2</sup> Much effort has also been devoted to investigate the forced response of flexural vibrations in platelike structures (see, for example, Refs. 3–5). Similar information for the in-plane vibration, to the knowledge of the authors, are lacking.

In recent years, the in-plane vibrational response of single and coupled plates has attracted the attention of researchers. The forced response of in-plane vibration in infinite plates can be found in many references (see, for example, Refs. 3, 4), much less work is reported on the in-plane response of finite plates. References 6–10 report some of the present efforts to include the in-plane response into predictions and measurements of the response to simultaneous flexural and in-plane excitations. They emphasize the importance of in-plane response at high frequencies and in

large coupled platelike structures. Only recently, Farag and Pan presented a method to compute the in-plane response and to extract the modal characteristics of thin rectangular plates.<sup>11</sup> The application of receptance concepts in the investigation of vibrational characteristics of plates can be found in Refs. 5 and 12. Ljunggren<sup>13</sup> has discussed an interesting coincidence coupling between a pressure wave over an area of a plate and the quasi-longitudinal waves in the plate. It is expected that in-plane waves in this case will have a predominant influence on energy flow from external pressure and through the coupling boundaries.

The resonant characteristics of in-plane vibrations of rectangular plates is the concern of this paper. The boundary conditions considered here represent the practical case of flat plate panels fixed to frames from two parallel edges. The other two edges may be either connected to stiffening beams or free. Linear displacements along the connected boundaries can be considered negligible while rotational displacements are permissible. These correspond to the simply supported conditions for flexural vibration and the clamped conditions for in-plane vibration.

The modal receptance solution developed in this paper is different from the solution presented in the previous work<sup>11,14</sup> in that:

- (1) A modal receptance solution is developed for plate panels with one or two parallel edges free while all other edges are clamped. These two cases of edge conditions

are important when two or more plates are coupled together where the coupling forces are applied on free edges, as detailed in Ref. 14.

- (2) In the present solutions the response is obtained as a receptance function which is mathematically formulated as a modal series summation.
- (3) The present solution provides more insight into the modal behavior and characteristics of rectangular panels. For example, the cutoff frequencies are defined mathematically and the relation between the uncoupled and coupled modes are discussed in detail. The free vibrational characteristics can also be obtained directly without first solving for the forced response.
- (4) Detailed examination of the modal behavior led to the development of approximate and simple procedures for quick estimation of the in-plane resonance frequencies and mode shapes.
- (5) Reference 14 presents an approximate receptance function for prediction of the in-plane response of finite plates excited by in-plane forces perpendicular to one free edge when all other edges are clamped. The in-plane response parallel to the free edge is neglected in Ref. 14.

In this paper, the two coupled equations governing the two in-plane response components are presented. These equations are based on the assumptions of the thin plate theory. The first section discusses the free vibrational characteristics of plates with two parallel edges clamped. In this section, the propagation and attenuation frequency bands and cutoff frequencies are examined, and expressions are derived for the resonance frequencies of coupled and uncoupled modes. In the second section, a solution for the forced response to in-plane force excitations is developed, taking into consideration the coupling between the in-plane longitudinal and in-plane shear waves. The response is obtained in the form of modal receptance functions. Subsequent sections discuss the mode shapes, resonance frequencies, and forced response of three different boundary conditions of rectangular plates with two parallel edges clamped and the other edges are (1) both clamped, (2) both free, or (3) one clamped and one free. The mode shapes satisfy both displacement and force conditions at the plate edges. Examples are given where the resonance frequencies and mode shapes of in-plane vibration are obtained from the present method and compared to the finite element predictions for plate panels with the abovementioned boundary conditions. Mode shapes corresponding to the first eight resonance frequencies are depicted.

### I. GENERAL CHARACTERISTICS OF IN-PLANE VIBRATION OF PLATES WITH TWO PARALLEL EDGES CLAMPED

The equations presented in the present and next sections are general for plate panels with two clamped edges at  $x=0$  and  $x=a$ . They are valid for any combination of edge conditions at  $y=0$  and  $y=b$ . The subsequent section will consider the three specific cases shown in Fig. 1, namely;

- (1) Clamped edges at  $y=0$  and  $y=b$ ;
- (2) Clamped edge at  $y=0$  and free edge at  $y=b$ ;

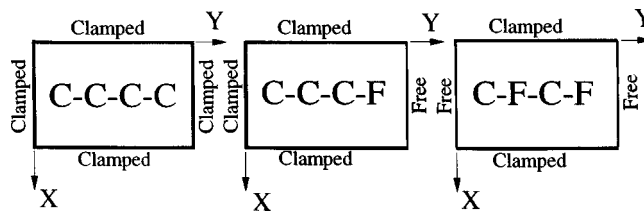


FIG. 1. Three cases of boundary conditions considered for in-plane vibration of rectangular plates; C-C-C-C, C-C-C-F, and C-F-C-F.

- (3) Free edges at  $y=0$  and  $y=b$ .

The first case is the simplest. A solution to this first case is presented in Ref. 11 following a completely different approach. The second and third cases are important for analysis of coupled plates.

#### A. The general equations for forced in-plane vibration

The equations governing the in-plane vibration of flat plates can be found in Refs. 10, 11 and 15. They may be written in the form:

$$C_L^2 \frac{\partial^2 u}{\partial x^2} + C_T^2 \frac{\partial^2 u}{\partial y^2} + (v C_L^2 + C_T^2) \frac{\partial^2 v}{\partial x \partial y} - \ddot{u} = - \frac{X_e}{\rho h} \delta(x - x_e) \delta(y - y_e), \quad (1a)$$

$$C_L^2 \frac{\partial^2 v}{\partial y^2} + C_T^2 \frac{\partial^2 v}{\partial x^2} + (v C_L^2 + C_T^2) \frac{\partial^2 u}{\partial x \partial y} - \ddot{v} = - \frac{Y_e}{\rho h} \delta(x - x_e) \delta(y - y_e). \quad (1b)$$

A list of symbols is given in the Appendix.

When two edges parallel to the  $Y$  axis are clamped, modal decomposition in the  $X$  direction can be adopted and a solution to Eqs. (1) may be assumed in the form:

$$u(x, y, t) = \sum_{m=1}^{\infty} \sin \frac{m \pi x}{a} U_m(y) e^{-j \omega t}, \quad (2a)$$

$$v(x, y, t) = \sum_{n=1}^{\infty} \sin \frac{n \pi x}{a} V_n(y) e^{-j \omega t}. \quad (2b)$$

The mode shapes  $U_m(y)$  and  $V_n(y)$  may be chosen to satisfy any combination of edge conditions at the edges  $y=0$  and  $y=b$ . Substitution of Eq. (2) into Eq. (1) and using the orthogonal properties of the sine functions in the  $X$  direction and the selective property of Dirac delta functions lead to the following equations:

$$\begin{aligned} & [C_T^2 U_m'' + \left[ \omega^2 - C_L^2 \left( \frac{m \pi}{a} \right)^2 \right] U_m \\ & + \sum_{n=1}^{\infty} \left[ (v C_L^2 + C_T^2) \left( \frac{2}{a} \right) I_{m,n} \right] V_n' \\ & = - \frac{2 X_e}{\rho h a} \sin \frac{m \pi x_e}{a} \delta(y - y_e), \end{aligned} \quad (3a)$$

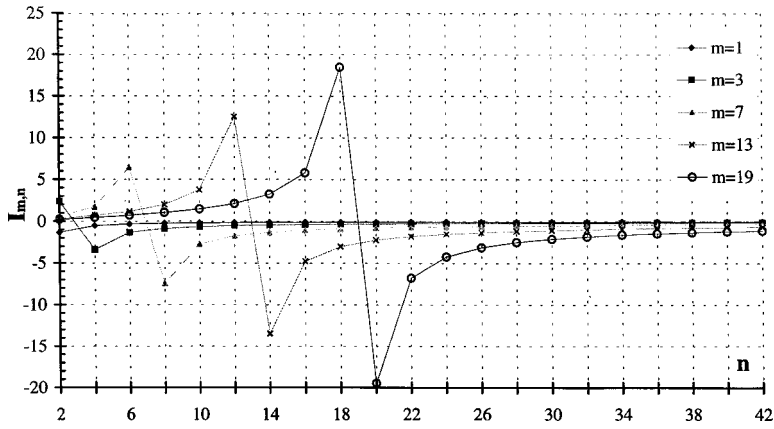


FIG. 2. Modal coupling factors of in-plane modes [Eq. (5)].

$$\begin{aligned}
 [C_L^2]V_n'' + \left[ \omega^2 - C_T^2 \left( \frac{n\pi}{a} \right)^2 \right] V_n \\
 + \sum_{m=1}^{\infty} \left[ (vC_L^2 + C_T^2) \left( \frac{2}{a} \right) I_{n,m} \right] U_m' \\
 = - \frac{2Y_e}{\rho h a} \sin \frac{n\pi x_e}{a} \delta(y - y_e), \quad (3b)
 \end{aligned}$$

where the coupling factors between the  $m$  and  $n$  modes are:

$$\begin{aligned}
 I_{m,n} &= \int_{x=0}^a \left( \frac{n\pi}{a} \right) \sin \frac{m\pi x}{a} \cos \frac{n\pi x}{a} dx, \\
 I_{n,m} &= \int_{x=0}^a \left( \frac{m\pi}{a} \right) \sin \frac{n\pi x}{a} \cos \frac{m\pi x}{a} dx. \quad (4)
 \end{aligned}$$

Evaluation of the above integrals gives:

$$I_{m,n} = \begin{cases} \frac{2mn}{(m^2 - n^2)} = -I_{n,m}, & \text{for odd values of } (m+n); \\ 0, & \text{otherwise.} \end{cases} \quad (5)$$

This indicates that the mode  $U_m(y)$  is directly coupled to the mode  $V_n(y)$  only when  $(m+n)$  is an odd integer. They are uncoupled when  $(m+n)$  is even. Equations (3) may be written in a more compact form as follows:

$$U_m'' + A_m U_m + \sum_{n=1}^{\infty} B_{m,n} V_n' = -X_m \delta(y - y_e), \quad (6a)$$

$$V_n'' + G_n V_n + \sum_{m=1}^{\infty} D_{n,m} U_m' = -Y_n \delta(y - y_e), \quad (6b)$$

where

$$\begin{aligned}
 A_m &= \frac{C_L^2}{C_T^2} \left[ \left( \frac{\omega^2}{C_L^2} \right) - \left( \frac{m\pi}{a} \right)^2 \right], \\
 G_n &= \frac{C_T^2}{C_L^2} \left[ \left( \frac{\omega^2}{C_T^2} \right) - \left( \frac{n\pi}{a} \right)^2 \right], \\
 B_{m,n} &= \left( \frac{2}{a} \right) \left( v \frac{C_L^2}{C_T^2} + 1 \right) (I_{m,n}),
 \end{aligned}$$

$$D_{n,m} = \left( \frac{2}{a} \right) \left( v + \frac{C_T^2}{C_L^2} \right) (I_{n,m}), \quad (7)$$

$$X_m = \frac{2X_e}{\rho h a C_T^2} \sin \frac{m\pi x_e}{a}, \quad Y_n = \frac{2Y_e}{\rho h a C_L^2} \sin \frac{n\pi x_e}{a}. \quad (8)$$

Before entering into the consideration of full coupling of all modes, it is necessary to obtain some insight into the modal and coupling characteristics by investigating the coupling between one pair of modes.

## B. Modal coupling characteristics

Equations (3) and (5) clearly show that the two components of in-plane response are coupled only when  $(m+n)$  is an odd integer. Equation (5) is plotted in Fig. 2 for values of  $m=1, 3, 7, 13,$  and  $19$  and  $n=2$  to  $42$ . The following characteristics of the coupling between the two components of in-plane response in finite plates can be observed from Eq. (5) and Fig. 2:

- (1) The maximum coupling always exists between adjacent modes, i.e., mode pair  $(m, n=m-1)$  and mode pair  $(m, n=m+1)$ .
- (2) The maximum value of the coupling factor increases with mode number.
- (3) Equation (5) can be used to investigate the relation between the values of the maximum coupling factors for two adjacent modes as follows:

$$\frac{I_{m,m+1}}{I_{m,m-1}} = - \left( \frac{1+1/m}{1-1/m} \right) \left( \frac{2-1/m}{2+1/m} \right). \quad (9)$$

From the above relation, the ratio  $I_{m,m+1}/I_{m,m-1}$  is  $-1.23, -1.11, -1.05,$  and  $-1.0$  for  $m=5, 10, 20,$  and  $\infty,$  respectively.

- (4) The nature of the coupling term in each of Eqs. (6) can now further be investigated for higher order modes. The coupling term is a product of the coupling factor and the first derivative of the coupled component of response. Away from the boundary the spatial distribution of a resonant response is generally sinusoidal. The products  $I_{m,n}V_n'$  and  $I_{n,m}U_m'$  when summed up over the equal number of modes at both sides of the two modes of maximum coupling factors will produce very small quantities (or practically cancel to zero). This is due to



the fact that the coupling factors are of nearly equal values and opposite signs for the modes of equal spacing from those of maximum coupling factors [e.g., the two modes  $(m, n = m - 3)$  and  $(m, n = m + 3)$ ]. This is also clear from examining Fig. 2 and from the fact that  $I_{m, m+1}/I_{m, m-1}$  approaches  $-1.0$  for large values of  $m$ . This nature of the coupling terms is also supported by the computational results presented in Sec. VI of this paper where it is shown that considering the coupling of more modes tends to reduce the value of the coupling terms.

### C. Solution of the equations of uncoupled modes

When  $m + n$  is an even integer,  $U_m(y)$  and  $V_n(y)$  are uncoupled to each other even though they are still coupled to other modes as indicated by Eqs. (6). Considering one mode pair  $(m$  and  $n)$  and neglecting the coupling to other modes, Eqs. (6) become:

$$U_m'' + A_m U_m = -X_m \delta(y - y_e), \quad (10a)$$

$$V_n'' + G_n V_n = -Y_n \delta(y - y_e). \quad (10b)$$

The complementary solution to Eqs. (10) is obtained in the form:

$$U_m(y) = U_m^+ e^{\mu_m y} + U_m^- e^{-\mu_m y}, \quad (11)$$

$$V_n(y) = V_n^+ e^{\mu_n y} + V_n^- e^{-\mu_n y},$$

where:

$$\mu_m^2 = -A_m \quad \text{and} \quad \mu_n^2 = -G_n.$$

When the effect of other modes on a specific  $(m, n)$  pair are ignored, the above solution represents the free in-plane vibration in a flat plate with two edges clamped (parallel to the  $Y$  axis) and with infinite extension in the  $Y$  direction. The frequencies at which  $A_m = 0$  and  $G_n = 0$  are the cutoff frequencies of the individual uncoupled modes given, respectively, by:

$$\omega_m = C_L \left( \frac{m\pi}{a} \right) \quad \text{and} \quad \omega_n = C_T \left( \frac{n\pi}{a} \right). \quad (12)$$

Below these cutoff frequencies,  $A_m$  and  $G_n$  are negative,  $\mu_m$  and  $\mu_n$  are real, and the response is mathematically expressed as hyperbolic functions (sinh and cosh) representing free disturbances which are attenuating exponentially in the  $Y$  direction. Above the cutoff frequencies,  $A_m$  and  $G_n$  are positive,  $\mu_m$  and  $\mu_n$  are imaginary, and the response is mathematically represented by trigonometric functions (sin and cos) representing free disturbances which will propagate or form a standing wave pattern in the  $Y$  direction. It will be shown in the following sections that for a plate with finite extension in the  $Y$  direction the cutoff frequencies are always higher than those given by Eqs. (12).

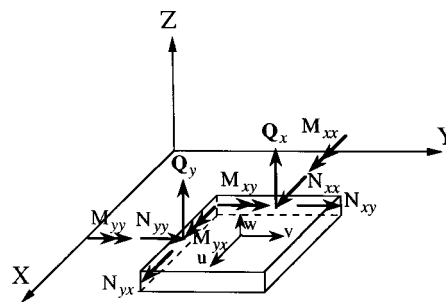


FIG. 3. Positive directions of forces and displacements for in-plane vibration.

For free uncoupled response,  $U_m(y)$  and  $V_n(y)$  must satisfy the displacement and force boundary conditions at the edges  $y = 0$  and  $y = b$ . The uncoupled response—Eqs. (11)—may be expressed in a general form as:

$$U_m(y) = \bar{U}_m f_u(y), \quad V_n(y) = \bar{V}_n f_v(y). \quad (13)$$

The in-plane longitudinal and shear forces (per unit width) at any point in the middle plane of the plate—see Fig. 3—are given as.<sup>5,14</sup>

$$N_{yy} = -\frac{Eh}{1-\nu^2} \left[ \frac{\partial v}{\partial y} + \nu \frac{\partial u}{\partial x} \right], \quad N_{xy} = -Gh \left[ \frac{\partial v}{\partial x} + \frac{\partial u}{\partial y} \right]. \quad (14)$$

Substituting the solution of Eqs. (2) and (13) into Eqs. (14) gives rise to the following relations for the longitudinal and shear forces at the edges  $y = 0$  and  $y = b$  for one mode pair  $(m$  and  $n)$ :

$$N_{yy} = \frac{-Eh}{1-\nu^2} \left[ \bar{V}_n \sin \frac{n\pi x}{a} f'_v(y) + \nu \bar{U}_m \left( \frac{m\pi}{a} \right) \cos \frac{m\pi x}{a} f_u(y) \right], \quad (15a)$$

$$N_{xy} = -Gh \left[ \bar{V}_n \left( \frac{n\pi}{a} \right) \cos \frac{n\pi x}{a} f_v(y) + \bar{U}_m \sin \frac{m\pi x}{a} f'_u(y) \right]. \quad (15b)$$

#### 1. For clamped edges at $y=0$ and $y=b$

Below the cutoff frequency, the response is expressed mathematically by a hyperbolic sine function of  $y$  which satisfies the condition of zero displacement at  $y = 0$ . But this function will not satisfy the condition of zero displacement at  $y = b$ . It follows that the uncoupled solution must not exist below the cutoff frequency. Above the cutoff frequency, the response, represented mathematically by a sine function, can satisfy both the boundary conditions of zero displacements and finite in-plane normal and in-plane shear forces at  $y = 0$  and  $y = b$ . It follows that the uncoupled response does exist above the cutoff frequency.

#### 2. For at least one free edge at $y=0$ and $y=b$

Considering Eqs. (15a) and (15b) at a free edge ( $y = y_f$ ), multiplying the first equation by  $\sin(n\pi x/a)$  and the

second equation by  $\sin(m\pi x/a)$ , integrating with respect to  $x$  from  $x=0$  to  $x=a$  and employing Eqs. (4), one obtains:

$$0 = \bar{V}_n \left( \frac{a}{2} \right) f'_v(y_f) + v \bar{U}_m I_{n,m} f'_u(y_f),$$

$$0 = \bar{V}_n I_{m,n} f'_v(y_f) + \bar{U}_m \left( \frac{a}{2} \right) f'_u(y_f).$$

For the special case of uncoupled modes, where  $I_{m,n} = I_{n,m} = 0$ , the above relations become:

$$\bar{V}_n f'_v(y_f) = 0, \quad \bar{U}_m f'_u(y_f) = 0.$$

Substituting the above relations into Eqs. (15) and noting that at the free edge  $f_u(y_f)$  and  $f_v(y_f)$  assume finite values, it becomes obvious that the condition for in-plane forces to vanish at a free edge is:  $\bar{U}_m = 0$  and  $\bar{V}_n = 0$ . This result means that the uncoupled modes do not exist for plate panels with one or two parallel edges free while the other edges are clamped.

#### D. Coupled mode pairs of in-plane vibration

Considering the coupling between one pair of modes ( $m$  and  $n$ ), Eqs. (6) may be written, for the case of free vibration, in the form:

$$U''_m + A_m U_m + B_{m,n} V'_n = 0, \quad V''_n + G_n V_n + D_{n,m} U'_m = 0. \quad (16)$$

The above equations, which govern the free vibration of a pair of coupled modes, when solved simultaneously for  $U_m(y)$  and  $V_n(y)$ , give rise to the following equations:

$$U_m^{iv} + E_{m,n} U_m'' + F_{m,n} U_m = 0, \quad (17)$$

$$V_n^{iv} + E_{m,n} V_n'' + F_{m,n} V_n = 0,$$

where

$$E_{m,n} = A_m + G_n - B_{m,n} D_{n,m} \quad \text{and} \quad F_{m,n} = A_m G_n. \quad (18)$$

Solution to Eqs. (17) may be written in the form of Eqs. (13), where:

$$f_u(y) = a_1 e^{K_1 y} + a_2 e^{-K_1 y} + a_3 e^{K_2 y} + a_4 e^{-K_2 y}, \quad (19)$$

$$K_1^2 = - \left( \frac{E}{2} \right) - \sqrt{\left( \frac{E}{2} \right)^2 - F} = -\alpha^2, \quad (20)$$

$$K_2^2 = - \left( \frac{E}{2} \right) + \sqrt{\left( \frac{E}{2} \right)^2 - F} = \beta^2.$$

$\alpha$  and  $\beta$  are related by the relation:

$$\beta^2 = \alpha^2 - E, \quad -\alpha^2 \beta^2 = F. \quad (21)$$

Although Eqs. (19)–(21) are written for one pair of modes, the subscripts  $m, n$  are removed for clarity.  $f_v(y)$  has the same expression as Eq. (19).

A nondimensional frequency parameter, defined as  $\Omega = \omega a / C_L$ , is substituted into Eqs. (7) and (18) to obtain:

$$E_{m,n} = \left( \frac{3-v}{1-v} \right) \left( \frac{1}{a^2} \right) \Omega^2 + \left( \frac{(1+v)^2}{1-v} \right) \left( \frac{2}{a^2} \right) I_{m,n}^2 - \left( \frac{2}{1-v} \frac{m^2 \pi^2}{a^2} + \frac{1-v}{2} \frac{n^2 \pi^2}{a^2} \right), \quad (22)$$

$$F_{m,n} = \left( \frac{m^2 \pi^2}{a^2} \right) \left( \frac{n^2 \pi^2}{a^2} \right) \left[ \left( \frac{2}{1-v} \frac{1}{m^2 \pi^2} \frac{1}{n^2 \pi^2} \right) \Omega^4 - \left( \frac{1}{m^2 \pi^2} + \frac{2}{1-v} \frac{1}{n^2 \pi^2} \right) \Omega^2 + 1 \right]. \quad (23)$$

$K_1^2$  and  $K_2^2$  are plotted in Fig. 4 against the nondimensional frequency parameter  $\Omega$  for the modes (1,2), (1,4), (2,1), (2,3), (3,2), and (3,4). The plate dimension across the clamped edges is  $a = 1.0$  m. The plate material is aluminum of properties  $E = 70 \times 10^9$  N/m<sup>2</sup>,  $\rho = 2700$  Kg/m<sup>3</sup>, and  $\nu = 0.33$ .

It is evident from Fig. 4 that each pair of coupled modes has two recognized frequencies at  $K_1^2 = 0$  and  $K_2^2 = 0$ . From the second of Eqs. (21) it is concluded that  $F_{m,n} = 0$  at each of these two recognized frequencies. The second of Eqs. (18) becomes  $A_m G_n = 0$  at these two recognized frequencies. Hence, these two frequencies are the two cutoff frequencies corresponding to  $A_m = 0$  and  $G_n = 0$ .

Three distinct frequency ranges can be identified:

- (1) Below the lower cutoff frequency,  $K_1^2$  and  $K_2^2$  are both positive and the two parts of the solution are exponentially decaying. This is a fully attenuated response.
- (2) Above the lower and below the upper cutoff frequencies,  $K_1^2$  is negative and  $K_2^2$  is positive. The first part of the solution is propagating and the second part is attenuating. The in-plane response can be recognized in this frequency range.
- (3) Above the second cutoff frequency,  $K_1^2$  and  $K_2^2$  are both negative. The two parts of the response are both propagating.

This result illustrates the meaning of the two resonance frequencies for each pair of coupled modes. For a coupled mode pair, the lower resonance frequency occurs in the frequency range bounded by the two cutoff frequencies while the upper resonance frequency is higher than the upper cutoff frequency. It has also to be noted that for some mode pairs [as shown in Fig. 4(a) and (d)] at low frequencies (small values of  $\Omega$ ),  $K_1^2$  and  $K_2^2$  are imaginary. In such cases the in-plane response, at least away from the source of excitation, is completely suppressed.

#### E. Resonance frequencies of coupled modes

In this section, an approximate solution is obtained for in-plane response to in-plane force excitations. The solution is expressed in the form of receptance functions for each pair

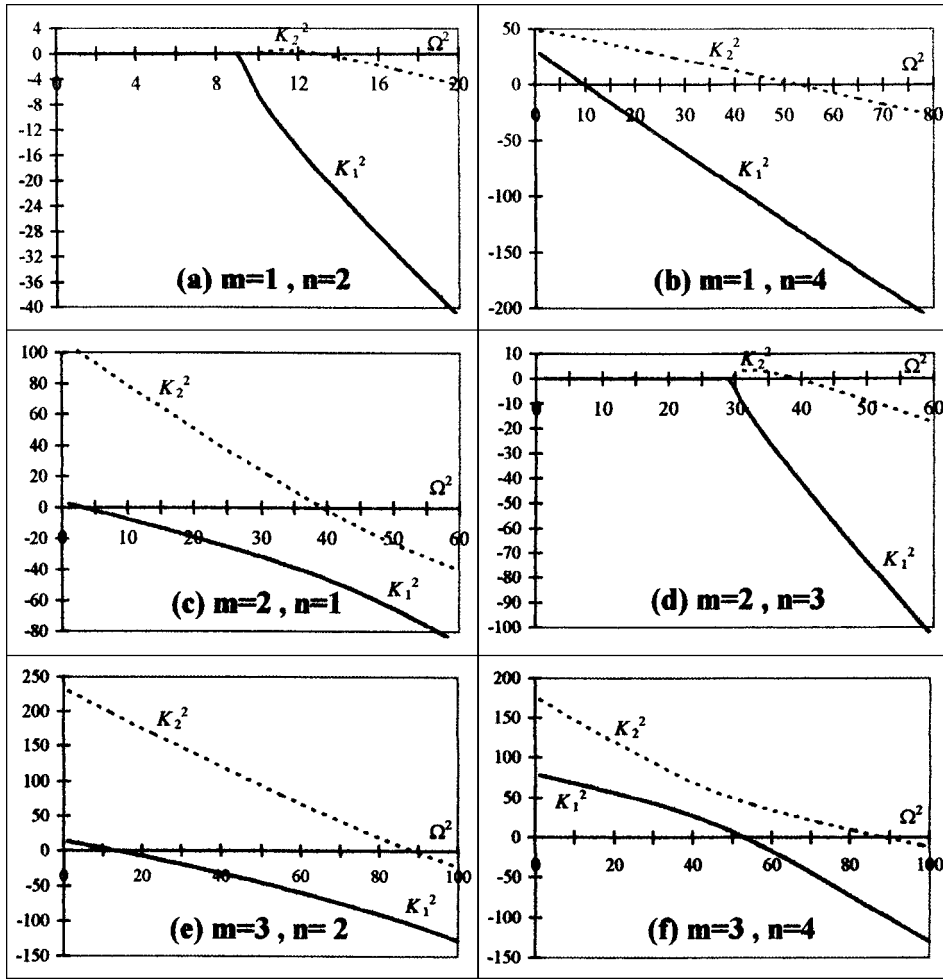


FIG. 4. Wave propagation characteristic frequencies for in-plane vibration.

of coupled modes. The total response is the summation of the receptance functions for all the participating modes.

Equations (13) are substituted into Eqs. (3), and the resultant equations are:

$$[f_p''(y) + A_m f_p(y)] \bar{U}_m + \sum_{n=1}^{\infty} B_{m,n} f_s'(y) \bar{V}_n = -X_m \delta(y - y_e), \quad (24a)$$

$$[f_s''(y) + G_n f_s(y)] \bar{V}_n + \sum_{m=1}^{\infty} D_{n,m} f_p'(y) \bar{U}_m = -Y_n \delta(y - y_e), \quad (24b)$$

where  $f_p(y)$  and  $f_s(y)$  are replacing  $f_u(y)$  and  $f_v(y)$ , respectively.

Multiplying the first of the above equations by  $f_p(y)$  and the second by  $f_s(y)$  and integrating from  $y=0$  to  $y=b$ , the following equations are obtained:

$$(-\alpha_p^2 + A_m) \bar{U}_{m,p} + \sum_n \frac{2B_{m,n}}{bS_p} \times \sum_s \int_{y=0}^b f_p f_s' dy \bar{V}_{n,s} = \frac{-2X_m f_p(y_e)}{bS_p}, \quad (25a)$$

$$(-\alpha_s^2 + G_n) \bar{V}_{n,s} + \sum_m \frac{2D_{n,m}}{bS_s} \times \sum_p \int_{y=0}^b f_s f_p' dy \bar{U}_{m,p} = \frac{-2Y_n f_s(y_e)}{bS_s}. \quad (25b)$$

An approximate solution for the modal participation factors  $\bar{U}_{m,p}$  and  $\bar{V}_{n,s}$  is obtained by considering the coupling between one mode pair  $(m,p)$  and  $(n,s)$  and neglecting the effect of other modes. Thus Eqs. (25) are approximated as:

$$\begin{bmatrix} (-\alpha_p^2 + A_m) & \frac{2B_{m,n}}{bS_p} \int_{y=0}^b f_p f_s' dy \\ \frac{2D_{n,m}}{bS_s} \int_{y=0}^b f_s f_p' dy & (-\alpha_s^2 + G_n) \end{bmatrix} \begin{bmatrix} \bar{U}_{m,p} \\ \bar{V}_{n,s} \end{bmatrix} = \begin{bmatrix} \frac{-2X_m f_p(y_e)}{bS_p} \\ \frac{-2Y_n f_s(y_e)}{bS_s} \end{bmatrix}, \quad (26)$$

where:

$$S_p = \frac{2}{b} \int_0^b f_p^2(y) dy, \quad S_s = \frac{2}{b} \int_0^b f_s^2(y) dy, \quad (27)$$

$$\int_0^b f_p(y) f_p''(y) dy = -\alpha_p^2 S_p, \quad (28)$$

$$\int_0^b f_s(y) f_s''(y) dy = -\alpha_s^2 S_s.$$

The subscripts  $p$  and  $s$  are introduced for identification of the mode shapes of  $U_m(y)$  and  $V_n(y)$ . Their mathematical and physical meaning will be illustrated in subsequent sections when considering specific boundary conditions at  $y=0$  and  $y=b$ .

Substitution of  $A_m$ ,  $G_n$ ,  $B_{m,n}$ , and  $D_{n,m}$  from Eqs. (7) into (26) gives rise to the following equations:

$$\begin{bmatrix} (\omega_{m,p}^2 - \omega^2) & -\frac{4}{abS_p}(vC_L^2 + C_T^2)I_{m,n}I_{p,s} \\ -\frac{4}{abS_s}(vC_L^2 + C_T^2)I_{n,m}I_{s,p} & (\omega_{n,s}^2 - \omega^2) \end{bmatrix} \times \begin{bmatrix} \bar{U}_{m,p} \\ \bar{V}_{n,s} \end{bmatrix} = \begin{bmatrix} \frac{2X_m C_T^2 f_p(y_e)}{bS_p} \\ \frac{2Y_n C_L^2 f_s(y_e)}{bS_s} \end{bmatrix}, \quad (29)$$

where:

$$I_{p,s} = \int_0^b f_p f_s' dy, \quad I_{s,p} = \int_0^b f_s f_p' dy, \quad (30)$$

$$\omega_{m,p}^2 = \left(\frac{m\pi}{a}\right)^2 C_L^2 + \alpha_p^2 C_T^2, \quad \omega_{n,s}^2 = \left(\frac{n\pi}{a}\right)^2 C_T^2 + \alpha_s^2 C_L^2.$$

The condition for in-plane vibration to take place at no external excitation is obtained by equating the determinant of the square matrix in the left hand side of Eq. (29) to zero. This gives the following characteristic equation:

$$\omega^4 - (\omega_{m,p}^2 + \omega_{n,s}^2)\omega^2 + (\omega_{m,p}^2 \omega_{n,s}^2 - R^2) = 0, \quad (31)$$

where:

$$R^2 = -\left(\frac{4}{ab}\right)^2 (vC_L^2 + C_T^2)^2 I_{m,n}^2 \frac{I_{p,s} I_{s,p}}{S_p S_s}. \quad (32)$$

The solution of Eq. (31) is:

$$\omega_{c1, c2}^2 = \left(\frac{\omega_{m,p}^2 + \omega_{n,s}^2}{2}\right) \mp \sqrt{\left(\frac{\omega_{m,p}^2 - \omega_{n,s}^2}{2}\right)^2 + R^2}. \quad (33)$$

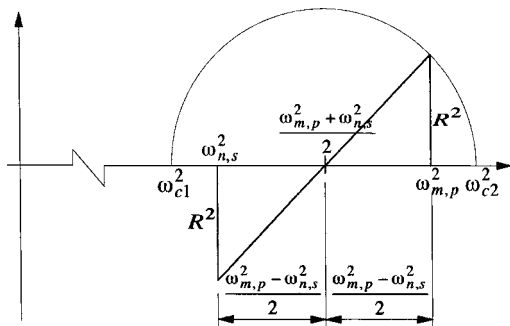


FIG. 5. Relation between coupled and uncoupled resonance frequencies.

$\omega_{c1}$  and  $\omega_{c2}$  are the two resonance frequencies for the coupled mode pair  $U_{m,p}$  and  $V_{n,s}$ .  $\omega_{m,p}$  and  $\omega_{n,s}$  are the resonance frequencies if the mode pair were not coupled. They will be referred as “the uncoupled resonance frequencies.” Figure 5 is a diagrammatic representation of Eq. (33). It shows the relation between the coupled and the uncoupled resonance frequencies. The following observations on the characteristics of the resonance frequencies  $\omega_{c1}$  and  $\omega_{c2}$  of the coupled mode pair  $U_{m,p}$  and  $V_{n,s}$  are worth recording:

- (1) The second term under the square root in Eq. (33) represents the effect of coupling on the resonance frequencies. It will be shown in subsequent sections that this term is positive for the three specific boundary conditions considered in this work.
- (2) It is clear from Fig. 5 that  $\omega_{c1}$  and  $\omega_{c2}$  are always outside the range  $[\omega_{m,p}, \omega_{n,s}]$ . This means that as a result of coupling  $\omega_{c1}$  is lower than  $\min[\omega_{m,p}, \omega_{n,s}]$  and  $\omega_{c2}$  is higher than  $\max[\omega_{m,p}, \omega_{n,s}]$ .
- (3)  $\omega_{c1}$  and  $\omega_{c2}$  are approximate values because the coupling with other modes is neglected. Consideration of the coupling to other modes will change slightly the values of  $\omega_{c1}$  and  $\omega_{c2}$  but the properties stated in (2) above will still hold. The effect of neglecting the coupling to other modes on the accuracy of predictions of the resonance frequencies will be investigated by computational examples in subsequent sections.
- (4) Each coupled mode pair has two resonance frequencies. This is due to the fact that the vibrational behavior of the coupled mode pair is mathematically represented—as a two degrees of freedom system—by the two equations [Eq. (29)]. Hence, the solution produces two resonance frequencies. This property of the coupled in-plane modes was observed in the previous work.<sup>11</sup> Its mathematical proof and graphical illustration are demonstrated here by Eq. (33) and Fig. 5.

## II. IN-PLANE FORCED RESPONSE

In this section an approximate solution will be derived for the forced response of each pair of coupled modes by neglecting the coupling to other modes. A more accurate solution can then be obtained by iteration employing the approximate solutions and considering the coupling between the different mode pairs.

The simultaneous solution of Eqs. (29) for  $\bar{U}_{m,p}$  and  $\bar{V}_{n,s}$  may be written in the form:

$$\begin{bmatrix} \bar{U}_{m,p} \\ \bar{V}_{n,s} \end{bmatrix} = \frac{1}{M} \frac{1}{(\omega_{m,p}^2 - \omega^2)(\omega_{n,s}^2 - \omega^2) - R^2} \times \begin{bmatrix} (\omega_{n,s}^2 - \omega^2) & C_Y \\ C_X & (\omega_{m,p}^2 - \omega^2) \end{bmatrix} \begin{bmatrix} X_{m,p} \\ Y_{n,s} \end{bmatrix}. \quad (34)$$

The denominator in the above equation may be factorized as presented in the previous section. Equation (34) is then written in the form:

$$\begin{bmatrix} \bar{U}_{m,p} \\ \bar{V}_{n,s} \end{bmatrix} = \frac{1}{M} \frac{1}{(\omega_{c1}^2 - \omega^2)(\omega_{c2}^2 - \omega^2)} \times \begin{bmatrix} (\omega_{n,s}^2 - \omega^2) & C_Y \\ C_X & (\omega_{m,p}^2 - \omega^2) \end{bmatrix} \begin{bmatrix} X_{m,p} \\ Y_{n,s} \end{bmatrix}, \quad (35)$$

where  $\omega_{c1}^2$  and  $\omega_{c2}^2$  are the squares of the approximate resonance frequencies of the mode pair  $(m,p)$  and  $(n,s)$ ,

$$M = \frac{\rho h a b}{4},$$

$$X_{m,p} = \frac{X_e}{S_p} \sin \frac{m \pi x_e}{a} f_p(y_e), \quad Y_{n,s} = \frac{Y_e}{S_s} \sin \frac{n \pi x_e}{a} f_s(y_e), \quad (36)$$

$$C_X = \frac{2}{a b S_s} C_L^2 (1 + \nu) I_{n,m} I_{s,p},$$

$$C_Y = \frac{2}{a b S_p} C_L^2 (1 + \nu) I_{m,n} I_{p,s}.$$

Equation (35) is used to obtain the approximate solutions for  $\bar{U}_{m,p}$  and  $\bar{V}_{n,s}$ . These approximate values are then substituted into the following equations, which are obtained directly from Eqs. (25), to get better approximation:

$$\bar{U}_{m,p} = \frac{1}{(\omega_{m,p}^2 - \omega^2)} \left[ \frac{X_{m,p}}{M} + \frac{2}{a b S_p} C_L^2 (1 + \nu) \times \sum_n I_{m,n} \sum_s I_{p,s} \bar{V}_{n,s} \right], \quad (37a)$$

$$\bar{V}_{n,s} = \frac{1}{(\omega_{n,s}^2 - \omega^2)} \left[ \frac{Y_{n,s}}{M} + \frac{2}{a b S_s} C_L^2 (1 + \nu) \times \sum_n I_{n,m} \sum_p I_{s,p} \bar{U}_{m,p} \right]. \quad (37b)$$

The characteristics of the coupling term as represented by Eqs. (5) and (9) and illustrated graphically in Fig. 2 can be employed to reduce the computational effort in Eqs. (37). It is sufficient to consider the coupling with only three neighboring modes in the summations of Eqs. (37).

The second term in the right hand sides of Eqs. (37) represents the coupling between the in-plane modes for  $u$  and  $v$  displacements. It is also clear from these equations that a pair of modes is uncoupled when  $I_{m,n} = 0$  and/or  $I_{p,s} = 0$ .

The approximate response of the uncoupled modes (if they exist) is obtained from Eqs. (37) by neglecting the coupling term. To obtain the response of the coupled modes, an iteration process is required. In the first step the coupling terms are neglected and an approximate solution is obtained. This approximate solution is then substituted into Eqs. (37) to obtain a second order approximation. The iteration process continues until the required accuracy is reached. The modal coupling characteristics presented in part ‘‘C’’ of Sec. I and illustrated by Fig. 2 have to be considered in the choice of the modes participating in the summations of Eqs. (37). For example, Fig. 2 shows that if the mode  $(m = 13, n = 12)$  is included in the summation, then the mode  $(13, 14)$  must be

included, and also the modes  $(13, 10)$ ,  $(13, 16)$ ,  $(13, 8)$ , and  $(13, 18)$  should be included.

### III. THE IN-PLANE FORCED RESPONSE OF RECTANGULAR PLATES WITH ALL EDGES CLAMPED

#### A. The uncoupled modes

Applying the displacement boundary conditions  $u = v = 0$  at  $y = 0$  to the general solution for uncoupled modes, Eqs. (11) become,

$$U_m(y) = \bar{U}_m [e^{\mu_m y} - e^{-\mu_m y}],$$

$$V_n(y) = \bar{V}_n [e^{\mu_n y} - e^{-\mu_n y}]. \quad (38)$$

Below the cutoff frequencies given by Eqs. (12),  $\mu_m$  and  $\mu_n$  are real and the solution may be written in the form:

$$U_m(y) = \bar{U}_m \sinh \mu_m y, \quad V_n(y) = \bar{V}_n \sinh \mu_n y. \quad (39)$$

In this case, displacement boundary conditions  $u = v = 0$  at  $y = b$  can only be satisfied when  $\bar{U}_m = \bar{V}_n = 0$ . This means that the uncoupled modes do not exist below the cutoff frequencies. Above the cutoff frequencies,  $\mu_m$  and  $\mu_n$  are imaginary and the uncoupled mode shapes in  $Y$  direction take the form:

$$U_m(y) = \bar{U}_m \sin \mu_m y, \quad V_n(y) = \bar{V}_n \sin \mu_n y. \quad (40)$$

Values of  $\mu_m$  and  $\mu_n$  which satisfy  $u = v = 0$  at  $y = b$  are:

$$\mu_{m,p} = \frac{C_L}{C_T} \sqrt{\left[ \left( \frac{\omega^2}{C_L^2} \right) - \left( \frac{m \pi}{a} \right)^2 \right]} = \frac{p \pi}{b},$$

$$\mu_{n,s} = \frac{C_T}{C_L} \sqrt{\left[ \left( \frac{\omega^2}{C_T^2} \right) - \left( \frac{n \pi}{a} \right)^2 \right]} = \frac{s \pi}{b}. \quad (41)$$

The corresponding resonance frequencies are:

$$\omega_{m,p} = \sqrt{\left( \frac{m \pi}{a} \right)^2 C_L^2 + \left( \frac{p \pi}{b} \right)^2 C_T^2},$$

$$\omega_{n,s} = \sqrt{\left( \frac{n \pi}{a} \right)^2 C_T^2 + \left( \frac{s \pi}{b} \right)^2 C_L^2}. \quad (42)$$

Combining Eqs. (40), (41), and (2) gives rise to the following solution for the uncoupled modes in the frequency range of pure propagation:

$$u(x, y, t) = \sum_{m=1}^{\infty} \sum_{p=1}^{\infty} \bar{U}_{m,p} \sin \frac{m \pi x}{a} \sin \frac{p \pi y}{b} e^{-j \omega t}, \quad (43a)$$

$$v(x, y, t) = \sum_{n=1}^{\infty} \sum_{s=1}^{\infty} \bar{V}_{n,s} \sin \frac{n \pi x}{a} \sin \frac{s \pi y}{b} e^{-j \omega t}, \quad (43b)$$

where  $p$  and  $s$  are positive integers and, for uncoupled modes,  $(m + n)$  is an even integer.

It is important that the above solution, in addition to satisfying the displacement boundary conditions, also satisfies the force conditions at the edges  $y = 0$  and  $y = b$ . Substituting this solution into Eqs. (15) gives finite values for the longitudinal and shear forces at  $y = 0$  and  $y = b$ .

## B. Mode shapes and resonance frequencies of the coupled modes

The general form of the solution for coupled modes is given by Eq. (19). It is shown above that, to satisfy the boundary conditions of the C–C–C–C plate, the hyperbolic part must not exist. It follows that the solution given by Eqs. (43) is also a complete solution for the coupled modes where  $(m+n)$  is an odd integer in the frequency range of propagation. It satisfies both the displacement and force boundary conditions. Substituting this solution into Eqs. (1) and employing the orthogonal properties of the sine functions, the following equations are obtained:

$$\left[ \omega^2 - \left( \frac{m\pi C_L}{a} \right)^2 - \left( \frac{p\pi C_T}{b} \right)^2 \right] \bar{U}_{m,p} + \left( \frac{4}{ab} \right) (vC_L^2 + C_T^2) \times \sum_{n=1}^{\infty} \sum_{s=1}^{\infty} I_{m,n} I_{p,s} \bar{V}_{n,s} = - \frac{4X_e}{\rho hab} \sin \frac{m\pi x_e}{a} \sin \frac{p\pi y_e}{b}, \quad (44a)$$

$$\left[ \omega^2 - \left( \frac{s\pi C_L}{b} \right)^2 - \left( \frac{n\pi C_T}{a} \right)^2 \right] \bar{V}_{n,s} + \left( \frac{4}{ab} \right) (vC_L^2 + C_T^2) \times \sum_{m=1}^{\infty} \sum_{p=1}^{\infty} I_{n,m} I_{s,p} \bar{U}_{m,p} = - \frac{4Y_e}{\rho hab} \sin \frac{n\pi x_e}{a} \sin \frac{s\pi y_e}{b}, \quad (44b)$$

where  $I_{m,n}$  and  $I_{n,m}$  are given by Eqs. (4) and their evaluations by Eqs. (5). Similarly,  $I_{p,s}$  and  $I_{s,p}$  and their evaluation are given by the same expressions after replacing  $m$  with  $p$  and  $n$  with  $s$ . For the C–C–C–C plate, the in-plane modes are coupled when both  $(m+n)$  and  $(p+s)$  are odd integers.

To obtain an approximate expression for the resonance frequencies of the coupled modes, it is assumed that each pair of coupled modes  $U_{m,p}$  and  $V_{n,s}$  is uncoupled from all other modes. For free vibrations, this assumption leads to Eqs. (44) be simplified to the form:

$$\begin{bmatrix} (\omega^2 - \omega_{m,p}^2) & \left( \frac{4}{ab} \right) (vC_L^2 + C_T^2) I_{m,n} I_{p,s} \\ \left( \frac{4}{ab} \right) (vC_L^2 + C_T^2) I_{n,m} I_{s,p} & (\omega^2 - \omega_{n,s}^2) \end{bmatrix} \times \begin{bmatrix} \bar{U}_{m,p} \\ \bar{V}_{n,s} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}. \quad (45)$$

A nontrivial solution of the above equation represents the case of free vibration. The corresponding frequencies are the resonance frequencies which result from the solution of the characteristic equation. After substitution of the expressions for  $I_{m,n}$  and  $I_{p,s}$  from Eqs. (5), the approximate resonance frequencies for the coupled mode pair  $U_{m,p}$  and  $V_{n,s}$  are given by Eq. (33) where  $R^2$  is:

$$R^2 = 16 \left[ \frac{vC_L^2 + C_T^2}{ab} \right]^2 \left[ \frac{2mn}{m^2 - n^2} \right]^2 \left[ \frac{2ps}{p^2 - s^2} \right]^2. \quad (46)$$

The resonance frequencies of the coupled modes and their relation to those of the uncoupled modes are illustrated by Fig. 5. The general comments following Eq. (33) are still valid in this special case of C–C–C–C plate.

## C. The forced response of C–C–C–C plate

The approximate solution, assuming coupling of only one mode pair  $U_{m,p}$  and  $V_{n,s}$  is given by Eq. (35) for the general case of boundary conditions at  $y=0$  and  $y=b$ . It is given by the following relations for the special case of C–C–C–C plate:

$$\bar{U}_{m,p} \equiv \frac{4X_e}{\rho hab} \frac{(\omega_{n,s}^2 - \omega^2)}{(\omega_{c1}^2 - \omega^2)(\omega_{c2}^2 - \omega^2)} \left( \sin \frac{m\pi x_e}{a} \sin \frac{p\pi y_e}{b} \right) + \frac{4Y_e}{\rho hab} \frac{2C_L^2(1+v)I_{m,n}I_{p,s}}{ab(\omega_{c1}^2 - \omega^2)(\omega_{c2}^2 - \omega^2)} \times \left( \sin \frac{n\pi x_e}{a} \sin \frac{s\pi y_e}{b} \right), \quad (47a)$$

$$\bar{V}_{n,s} \equiv \frac{4Y_e}{\rho hab} \frac{(\omega_{m,p}^2 - \omega^2)}{(\omega_{c1}^2 - \omega^2)(\omega_{c2}^2 - \omega^2)} \left( \sin \frac{n\pi x_e}{a} \sin \frac{s\pi y_e}{b} \right) + \frac{4X_e}{\rho hab} \frac{2C_L^2(1+v)I_{n,m}I_{s,p}}{ab(\omega_{c1}^2 - \omega^2)(\omega_{c2}^2 - \omega^2)} \times \left( \sin \frac{m\pi x_e}{a} \sin \frac{p\pi y_e}{b} \right). \quad (47b)$$

An iteration process is required to improve the accuracy of the solutions for  $U_{m,p}$  and  $V_{n,s}$ .

## IV. THE IN-PLANE FORCED RESPONSE OF CLAMPED–CLAMPED–CLAMPED–FREE PLATES

### A. Mode shapes and resonance frequencies

It is proved in Sec. I (part ‘‘C’’) that for this case of boundary conditions the uncoupled modes, where  $(m+n)$  is an even integer, must not exist in order to obtain zero forces at the free edge. For coupled modes, the mode shapes should include a hyperbolic function in order to comply with the general solution represented by Eq. (19). The following assumed mode shapes satisfy the conditions of zero in-plane displacement at  $y=0$ :

$$U_{m,p}(y) = \bar{U}_{m,p} f_p(y) = \bar{U}_{m,p} \left[ \sin \left( p - \frac{1}{2} \right) \frac{\pi y}{b} + H_{ps} \sinh \beta_p y \right]. \quad (48a)$$

$$V_{n,s}(y) = \bar{V}_{n,s} f_s(y) = \bar{V}_{n,s} \left[ \sin \left( s - \frac{1}{2} \right) \frac{\pi y}{b} + H_{ps} \sinh \beta_s y \right]. \quad (48b)$$

Adopting Eqs. (15), the conditions of zero forces at a free edge are expressed as:

$$N_{yy} = \frac{-Eh}{1-\nu^2} \left[ \bar{V}_{n,s} \sin \frac{n\pi x}{a} f'_s(y) + \nu \bar{U}_{m,p} \left( \frac{m\pi}{a} \right) \cos \frac{m\pi x}{a} f_p(y) \right] = 0, \quad (49a)$$

$$N_{xy} = -Gh \left[ \bar{V}_{n,s} \left( \frac{n\pi}{a} \right) \cos \frac{n\pi x}{a} f_s(y) + \bar{U}_{m,p} \sin \frac{m\pi x}{a} f'_p(y) \right] = 0. \quad (49b)$$

Multiplying the first equation by  $\sin(n\pi x/a)$ , the second equation by  $\sin(m\pi x/a)$  and integrating both equations with respect to  $x$  from  $x=0$  to  $x=a$ , the following relations are obtained:

$$\left( \frac{a}{2} \right) f'_s(y) \bar{V}_{n,s} + \nu I_{n,m} f_p(y) \bar{U}_{m,p} = 0, \quad (50a)$$

$$I_{m,n} f_s(y) \bar{V}_{n,s} + \left( \frac{a}{2} \right) f'_p(y) \bar{U}_{m,p} = 0. \quad (50b)$$

The condition for a nontrivial solution for a coupled mode pair ( $U_{m,p}$  and  $V_{n,s}$ ) which produce zero in-plane forces at the free edge, as obtained from the above equations, is:

$$\left( \frac{a}{2} \right)^2 f'_p(y) f'_s(y) + \nu I_{m,n}^2 f_p(y) f_s(y) = 0. \quad (51)$$

Substituting the mode shapes from Eqs. (48), Eq. (51) becomes:

$$H_{ps}^2 \left[ \left( \frac{a}{2} \right)^2 \frac{\beta_p \beta_s}{\nu I_{m,n}^2} + \tanh \beta_p b \tanh \beta_s b \right] - H_{ps} \left[ (-1)^s \frac{\tanh \beta_p}{\cosh \beta_s b} + (-1)^p \frac{\tanh \beta_s}{\cosh \beta_p b} \right] + \left[ \frac{(-1)^{(p+s)}}{\cosh \beta_p b \cosh \beta_s b} \right] = 0. \quad (52)$$

Equation (52) is a quadratic equation in  $H_{ps}$ . The discriminator of Eq. (52) must be positive as a condition for the mode pair to exist. It can be shown that an even value for  $(p+s)$  will always produce a negative discriminator and an odd value produces a positive discriminator. Hence,  $(p+s)$  must be an odd integer for the coupled mode pair ( $U_{m,p}$  and  $V_{n,s}$ ) to exist.

The resonant characteristics of the coupled modes are investigated following the analysis presented in Sec. I. Equation (35) governs the approximate forced response of a coupled mode pair. In this equation,  $I_{ps}$  and  $I_{sp}$  have to be evaluated for the specific mode shapes under consideration. Approximate values for  $I_{ps}$  and  $I_{sp}$  may be obtained easily if the hyperbolic term in the mode shape [see Eqs. (48)] is neglected. This approximation is justifiable in view of the fact that the hyperbolic term is always very small compared to the trigonometric term. Thus considering only the trigono-

metric terms in the mode shapes  $f_p(y)$  and  $f_s(y)$  and noting that  $(p+s)$  is odd for the coupled modes,  $S_p$  and  $S_s$  of Eqs. (32)–(37) are given by  $S_p \cong S_s \cong 1$  and the evaluation of  $I_{ps}$  and  $I_{sp}$  becomes:

$$I_{ps} \cong \frac{(s - \frac{1}{2})}{(p-s)}, \quad I_{sp} \cong \frac{(p - \frac{1}{2})}{(s-p)}. \quad (53)$$

For the present boundary conditions, Eqs. (30) and (32) become:

$$\omega_{m,p}^2 = \left( \frac{m\pi}{a} \right)^2 C_L^2 + \left( p - \frac{1}{2} \right)^2 \left( \frac{\pi}{b} \right)^2 C_T^2, \quad (54)$$

$$\omega_{n,s}^2 = \left( \frac{n\pi}{a} \right)^2 C_T^2 + \left( s - \frac{1}{2} \right)^2 \left( \frac{\pi}{b} \right)^2 C_L^2,$$

$$R^2 = \left( \frac{4}{ab} \right)^2 (\nu C_L^2 + C_T^2)^2 I_{m,n}^2 \frac{(p - \frac{1}{2})(s - \frac{1}{2})}{(p-s)^2}. \quad (55)$$

The resonance frequencies for the coupled mode pair ( $U_{m,p}$  and  $V_{n,s}$ ) are then obtained utilizing Eq. (33). It is obvious that the value of  $R^2$  in Eq. (55) is always positive. The resonance frequencies  $\omega_{c1}$  and  $\omega_{c2}$  of the coupled mode pair  $U_{m,p}$  and  $V_{n,s}$  have the same characteristics as illustrated by Fig. 5 and can be explained by the general comments following Eq. (33).

To summarize, the resonance frequencies and mode shapes for a C–C–C–F plate are determined by the following procedure:

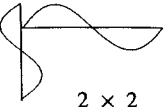

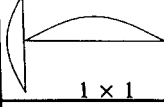
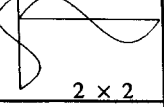
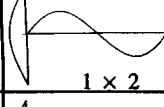
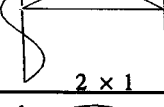
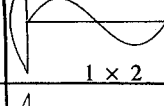
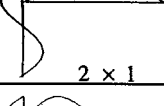
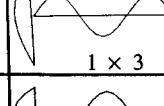
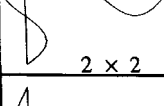
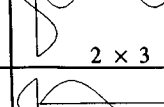
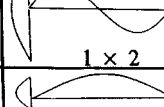
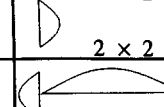
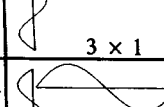
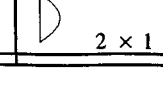
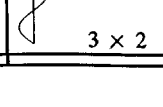
- (1) For each  $(m,n)$  pair, assume values for  $p$  and  $s$ . Note that for coupled modes  $(m+n)$  and  $(p+s)$  are odd integers.
- (2) Calculate the uncoupled ( $\omega_{m,p}^2$  and  $\omega_{n,s}^2$ ) and the coupled ( $\omega_{c1}$  and  $\omega_{c2}$ ) frequencies employing Eqs. (54), (55), and (33).
- (3) Use Eqs. (7) and (18) to obtain  $E_{m,n}$  and Eq. (21) to get  $\beta_p$  and  $\beta_s$ .
- (4) Substitute  $\beta_p$  and  $\beta_s$  into Eq. (52) and solve for  $H_{ps}$ . Use Eqs. (48) for the mode shapes.

The above procedure has to be repeated for all  $(m,n)$  and  $(p,s)$  combinations which produce resonance frequencies and corresponding mode shapes in the frequency range of interest.

## B. The forced response of C–C–C–F plates

The approximate solution, assuming coupling of only one mode pair  $U_{m,p}$  and  $V_{n,s}$ , is discussed in Sec. II for the general case of boundary conditions at  $y=0$  and  $y=b$ . The values of  $f_p(y)$ ,  $f_s(y)$ ,  $I_{m,n}$ ,  $I_{n,m}$ ,  $I_{p,s}$ ,  $I_{s,p}$ ,  $S_p$ ,  $S_s$ ,  $\omega_{m,p}$ ,  $\omega_{n,s}$ ,  $\omega_{c1}$ , and  $\omega_{c2}$  for one coupled mode pair  $U_{m,p}$  and  $V_{n,s}$  for a C–C–C–F plate are substituted into Eqs. (36) and then Eqs. (35) to obtain the approximate forced response  $\bar{U}_{m,p}$  and  $\bar{V}_{n,s}$ . Having obtained the approximate values for the coupled mode pairs, more accurate values are obtained by substituting the approximate values into Eqs. (37).

TABLE I. Comparison of resonance frequencies and mode shapes for in-plane vibration of C-C-C-C plate.

Mode no.	Mode shape for $u(x, y)$ $m \times p$	Mode shape for $v(x, y)$ $n \times s$	Resonance frequency (FE)	Resonance frequency Ref. 11	% diff. (Ref. 11 - FE)	Resonance frequency Modal	% diff. (Modal - FE)
1	 $2 \times 2$	 $1 \times 1$	2658	2666	0.3 %	2671	0.5%
2	 $1 \times 1$	 $2 \times 2$	2898	2906	0.3 %	2914	0.6%
3	 $1 \times 2$	 $2 \times 1$	3260	3279	0.6 %	3349	2.7%
4	 $1 \times 2$	 $2 \times 1$	4024	4052	0.7 %	4198	4.3%
5	 $1 \times 3$	 $2 \times 2$	4268	4308	1.0 %	4404	3.2%
6	 $2 \times 3$	 $1 \times 2$	4404	4431	0.6 %	4607	4.6%
7	 $2 \times 2$	 $3 \times 1$	4769	4820	1.1 %	4917	3.1%
8	 $2 \times 1$	 $3 \times 2$	5300	5350	0.9 %	5329	0.5%

**V. THE IN-PLANE FORCED RESPONSE OF CLAMPED-FREE-CLAMPED-FREE PLATES**

**A. Mode shapes and resonance frequencies**

It is proved in Sec. I that for this case of boundary conditions the uncoupled modes, where  $(m+n)$  is an even integer, must not exist in order to obtain zero forces at the free edges. For coupled modes, the mode shapes have to satisfy the conditions of zero in-plane forces, as expressed by Eqs. (49), at the free edges  $y=0$  and  $y=b$ .

The free vibrational modes may be classified into symmetric and antisymmetric modes. The symmetric modes in the  $Y$  direction have the properties:

$$f(0)=f(b), \tag{56a}$$

$$f'(0)=-f'(b), \tag{56b}$$

and the antisymmetric modes have the properties:

$$f(0)=-f(b), \tag{57a}$$

$$f'(0)=f'(b). \tag{57b}$$

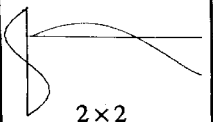
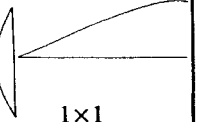
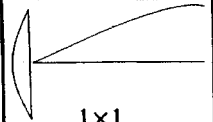
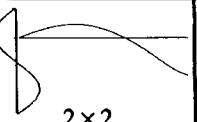
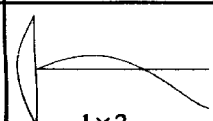
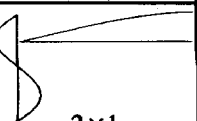
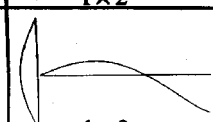
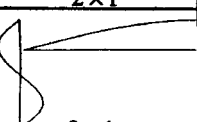
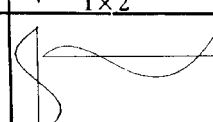
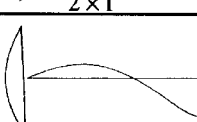
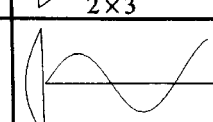
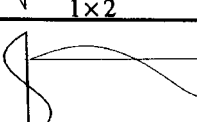
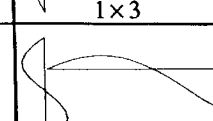
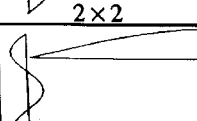
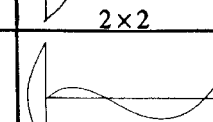
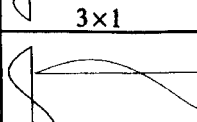
The symmetric and antisymmetric properties of the coupled modes can be investigated in view of Eqs. (49), (56), and (57). If  $f_p(y)$  and  $f_s(y)$  are both symmetric, Eq. (49a) will be satisfied either at  $y=0$  or at  $y=b$  but not at the two edges. The same thing applies for symmetric modes and Eq. (49b). This leads to a conclusion that symmetric modes of  $U$  cannot be coupled to symmetric modes of  $V$ . Similarly, antisymmetric modes of  $U$  cannot be coupled to antisymmetric modes of  $V$ .

Following the same argument, it can be shown that symmetric modes of  $U$  when coupled to antisymmetric modes of  $V$  satisfy the boundary conditions expressed by Eqs. (49) at  $y=0$  and at  $y=b$ . Also antisymmetric modes of  $U$  when coupled to symmetric modes of  $V$  satisfy these boundary conditions.

In view of the above discussion and the general expression for the complementary solution [Eq. (19)], a mode shape pair for symmetric  $U$  coupled to antisymmetric  $V$  is assumed in the form:



TABLE II. Comparison of resonance frequencies and mode shapes for in-plane vibration of C-C-C-F plate.

Mode number	Mode shape for $u(x, y)$ $m \times p$	Mode shape for $v(x, y)$ $n \times s$	Resonance frequency (FE)	Resonance frequency (Modal)	% diff. (Modal - FE)
1	 2x2	 1x1	1803	1892	4.9%
2	 1x1	 2x2	2656	2727	2.7%
3	 1x2	 2x1	2794	3026	8.4%
4	 1x2	 2x1	3392	3596	6.0%
5	 2x3	 1x2	3479	3624	4.2%
6	 1x3	 2x2	3704	3868	4.4%
7	 2x2	 3x1	4339	4704	8.4%
8	 1x3	 2x2	4589	4899	6.8%

$$U_{m,p}(y) = \bar{U}_{m,p} f_p^e(y) = \bar{U}_{m,p} \left[ \cos\left(\frac{p_e \pi y}{b}\right) + A_{ps} \cosh \beta_{p_e} \left(y - \frac{b}{2}\right) \right], \quad (58a)$$

$$U_{m,p}(y) = \bar{U}_{m,p} f_p^0(y) = \bar{U}_{m,p} \left[ \cos\left(\frac{p_0 \pi y}{b}\right) + A_{ps} \sinh \beta_{p_0} \left(y - \frac{b}{2}\right) \right], \quad (59a)$$

$$V_{n,s}(y) = \bar{V}_{n,s} f_s^0(y) = \bar{V}_{n,s} \left[ \cos\left(\frac{s_0 \pi y}{b}\right) + A_{ps} \sinh \beta_{s_0} \left(y - \frac{b}{2}\right) \right], \quad (58b)$$

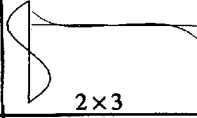
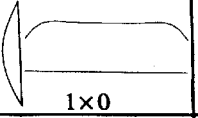
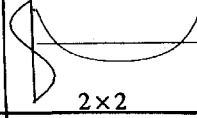
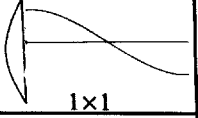
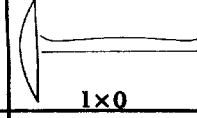
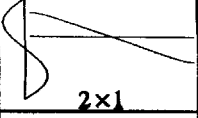
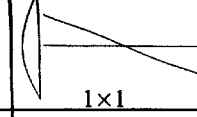
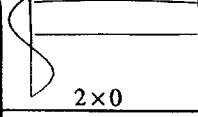
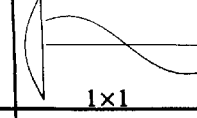
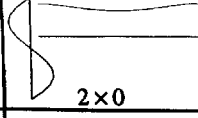
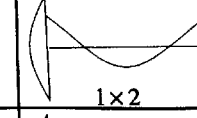
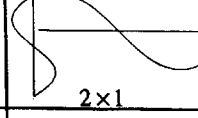
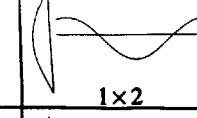
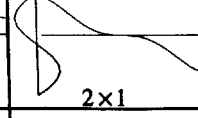
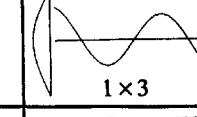
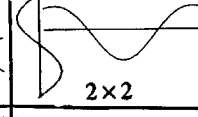
$$V_{n,s}(y) = \bar{V}_{n,s} f_s^e(y) = \bar{V}_{n,s} \left[ \cos\left(\frac{s_e \pi y}{b}\right) + A_{ps} \cosh \beta_{s_e} \left(y - \frac{b}{2}\right) \right], \quad (59b)$$

where  $p_0$  and  $s_0$  are odd integers assuming the values 1, 3, 5, etc.  $p_e$  and  $s_e$  are zero or even integers assuming the values 0, 2, 4, etc.

and a mode shape pair for antisymmetric  $U$  coupled to symmetric  $V$  is assumed in the form:

The condition for the free vibrational response to satisfy the condition of zero in-plane forces at the free edge is given by Eq. (51). Substituting the mode shapes from

TABLE III. Comparison of resonance frequencies and mode shapes for in-plane vibration of C-F-C-F plate.

Mode number	Mode shape for $u(x, y)$ $m \times p$	Mode shape for $v(x, y)$ $n \times s$	Resonance frequency (FE)	Resonance frequency (Modal)	% diff. (Modal - FE)
1	 2x3	 1x0	1449	1531	7%
2	 2x2	 1x1	2511	2682	6%
3	 1x0	 2x1	2567	2697	5%
4	 1x1	 2x0	2637	2994	12%
5	 1x1	 2x0	3037	3122	3%
6	 1x2	 2x1	3061	3390	10%
7	 1x2	 2x1	3917	4165	6%
8	 1x3	 2x2	4225	4451	5%

Eqs. (58) or (59) into Eq. (51), the following equation is obtained:

$$\begin{aligned}
 & A_{ps}^2 \left[ \left( \frac{a}{2} \right)^2 \frac{\beta_p \beta_s}{v I_{m,n}^2} \tanh \frac{\beta_p b}{2} + \tanh \frac{\beta_s b}{2} \right] \\
 & + A_{ps} \left[ \frac{\tanh \frac{\beta_s b}{2}}{\cosh \frac{\beta_p b}{2}} - \frac{1}{\cosh \frac{\beta_s b}{2}} \right] \\
 & - \left[ \frac{1}{\cosh \frac{\beta_p b}{2} \cosh \frac{\beta_s b}{2}} \right] = 0. \tag{60}
 \end{aligned}$$

The above equation is used to compute  $A_{ps}$  for each mode pair. It can be shown that the discriminator of the above equation is always positive indicating real roots.

The resonant characteristics of the coupled modes are

investigated following the analysis presented in Sec. I. Equation (29) governs the approximate forced response of a coupled mode pair. In this equation,  $I_{p,s}$  and  $I_{s,p}$  have to be evaluated for the specific mode shapes under consideration. Approximate values for  $I_{p,s}$  and  $I_{s,p}$  may be obtained without excessive computations if the hyperbolic terms in the mode shapes of Eqs. (58) and (59) are neglected. This approximation is justifiable in view of the fact that the hyperbolic term is always very small compared to the trigonometric term. The sketches of the mode shapes (see Table III) clearly show that the effect of the hyperbolic term is only noticeable very close to the free ends. It can also be observed that this effect is such that it increases the response at one edge and decreases it at the other end so that the area is approximately the same as that of the trigonometric term only.

The evaluation of  $I_{p,s}$  and  $I_{s,p}$ , based on the above simplification, is given by:

$$\text{and } \left. \begin{aligned} I_{p,s} &\cong \frac{2s^2}{(p^2-s^2)} \\ I_{s,p} &\cong \frac{2p^2}{(s^2-p^2)} \end{aligned} \right\} \text{ for odd values of } (p+s), \quad (61)$$

$$I_{p,s} \cong I_{s,p} \cong 0 \text{ otherwise.}$$

Also,  $S_p$  and  $S_s$  of Eqs. (25)–(32) are given by  $S_p \cong S_s \cong 1$ .

The result of Eq. (61) confirms that a coupled mode pair always consists of one odd (antisymmetric) mode coupled to an even (symmetric) mode. Substitution of the above values of  $I_{p,s}$  and  $I_{s,p}$  into Eq. (32) gives  $R^2$  for the present boundary conditions in exactly the same form as Eq. (46). The resonance frequencies for the coupled mode pair are obtained from Eq. (33). The illustration by Fig. 5 and the general comments following Eq. (33) are still valid in the special case of C–F–C–F plate.

To summarize, the resonance frequencies and mode shapes of in-plane vibration in a C–F–C–F plate are determined by the following procedure:

- (1) For each  $(m,n)$  pair, assume values for  $p$  and  $s$ . Note that for coupled modes  $(m+n)$  and  $(p+s)$  are odd integers. Calculate  $\alpha_p$  and  $\alpha_s$  from the relations:

$$\alpha_p = \frac{p\pi}{b}, \quad \alpha_s = \frac{s\pi}{b};$$

- (2) Calculate the uncoupled ( $\omega_{m,p}^2$  and  $\omega_{n,s}^2$ ) and the coupled frequencies ( $\omega_{c1}$  and  $\omega_{c2}$ ) employing Eqs. (30), (46), and (33);
- (3) Use Eqs. (7) and (18) to obtain  $E_{m,n}$  then Eq. (21) to get  $\beta_p$  and  $\beta_s$ ;
- (4) Substitute  $\beta_p$  and  $\beta_s$  into Eq. (60) and solve for  $A_{ps}$ .

The above procedure has to be repeated for all  $(m,n)$  and  $(p,s)$  combinations which produce resonance frequencies and corresponding mode shapes in the frequency range of interest.

## B. The forced response of C–F–C–F plates

The approximate solution, assuming coupling of only one coupled mode pair  $U_{m,p}$  and  $V_{n,s}$ , is discussed in Sec. II for the general case of boundary conditions at  $y=0$  and  $y=b$ . The values of  $f_p(y)$ ,  $f_s(y)$ ,  $I_{m,n}$ ,  $I_{n,m}$ ,  $I_{p,s}$ ,  $I_{s,p}$ ,  $S_p$ ,  $S_s$ ,  $\omega_{m,p}$ ,  $\omega_{n,s}$ ,  $\omega_{c1}$ , and  $\omega_{c2}$  for one coupled mode pair  $U_{m,p}$  and  $V_{n,s}$  for a C–F–C–F plate are substituted into Eqs. (36) and Eqs. (35) to obtain the approximate forced response  $\bar{U}_{m,p}$  and  $\bar{V}_{n,s}$ . Having obtained the approximate values for the coupled mode pairs, more accurate values can be obtained by substituting the approximate values into Eqs. (37).

## VI. COMPUTATIONAL EXAMPLES

The resonance frequencies and mode shapes of in-plane vibration in a rectangular plate are computed using the procedures given in the last three sections for the three cases of

boundary conditions of Fig. 1. The accuracy of the present formulation is assessed by comparing its results with Finite Element (FE) predictions for these three cases of boundary conditions. The predictions of the present method are also compared against the predictions of the method of Ref. 11 for C–C–C–C edges. The comparisons are presented in Tables I, II, and III.

The rectangular plate used in the simulation has the same configuration as that used in Ref. 11. Plate dimensions are  $1.2 \times 1.0 \times 0.025 \text{ m}^3$  and its aluminum material has a Young's modulus of  $70 \times 10^9 \text{ N/m}^2$ , density of  $2700 \text{ kg/m}^3$ , and Poisson's ratio of 0.33. To obtain FE predictions, the plate was modeled, using MSC NASTRAN, by 480 plate elements ( $0.05 \times 0.05 \times 0.025 \text{ m}$  each) of the 4-node type.

For the C–C–C–C case, Table I presents the mode shapes and compares the resonance frequencies for the first eight modes of in-plane vibration. The results of the present method are compared against the FE predictions and the results of Ref. 11. Tables II and III present the mode shapes and compares resonance frequencies obtained from the present method and FE for the first eight modes for C–C–C–F and C–F–C–F edge conditions, respectively. The plots of mode shapes in Tables I, II, and III are those of FE predictions.

The following observations are worth noting:

- (1) Predictions of the mode shapes by the different methods are very similar. This is clear from subjective comparison of the mode shapes predicted by FE and the assumed mode shapes of Eqs. (43), (48), (58), and (59).
- (2) Predictions of resonance frequencies by the present method are always higher than FE predictions. They are also higher than predictions of the method of Ref. 11 for C–C–C–C boundaries (for the first seven modes). This is attributed to the fact that the present method considers only the coupling between one pair of modes and neglects the coupling to all other modes. This also suggests that, in view of Eqs. (32) and (33), the coupling term  $R^2$  decreases as more modes are considered in computing the coupled resonance frequencies ( $\omega_{c1}$  and  $\omega_{c2}$ ).
- (3) The deviation of the predictions of the present method from those of FE is within 4.6%, 8.4%, and 12% for C–C–C–C, C–C–C–F, and C–F–C–F boundary conditions, respectively. This suggests that the coupling to other modes has its highest strength in the C–F–C–F case and its lowest strength in the C–C–C–C case. This observation agrees with the expected physical behavior of the rectangular panels with these cases of boundary conditions. The clamped edges at  $y=0$  and  $y=b$  (see Fig. 1) add a constraint to the response in the  $Y$  direction due to excitation in the  $X$  direction while the free edges at  $y=0$  and/or  $y=b$  (see Fig. 1) give more freedom for the coupling to occur.
- (4) The method of Ref. 11 produces high accuracy on the expense of high computational cost. The present method produces reasonable accuracy with a very low computational effort. It can be used for quick and approximate estimation of the in-plane resonances. It gives more insight into the physical behavior of in-plane vibrations in

thin plates. The accuracy of the present method can be increased by considering the coupling between many modes.

- (5) Further work is recommended to investigate the effect of number of mode pairs considered and the boundary conditions on the accuracy of the predictions.

## VII. CONCLUSIONS

The equations of in-plane vibration of thin plates are solved for rectangular panels with two parallel edges clamped while the other two edges are either both clamped, both free or one clamped and one free. The propagation characteristics of in-plane waves are investigated and three frequency ranges are identified for full attenuation, partial propagation and full propagation. Simple expressions for the cutoff frequencies are derived. The nature of the coupling between in-plane longitudinal and in-plane shear waves is illustrated to be due to the shear waves accompanying the longitudinal waves and due to Poisson's effect. The resonant modes are classified into uncoupled and coupled modes. It is proved that the uncoupled modes only exist when all edges are clamped. Simple expressions are derived for approximate estimation of the resonance frequencies of coupled and uncoupled modes. Resonant mode shapes are given, for each case of edge conditions, which satisfy both the displacement and force conditions at the plate edges. A simple procedure is given for the determination of resonance frequencies and mode shapes, in any frequency range of interest, without excessive computations. The forced response to in-plane force excitations is expressed as modal receptance functions. The predicted resonance frequencies and mode shapes are compared to the finite element results for the three cases of boundary conditions considered in this paper. The maximum deviation is 12% in the frequency range considered in the example problems. This deviation is mainly due to the approximation of considering the coupling between two modes only. The accuracy can be improved if the coupling to other modes is considered. The mode shapes of in-plane vibration are depicted for the first eight resonant modes for each case of edge conditions. The results of the present method are also compared against the results of the method of Ref. 11 for the same plate panel to illustrate the consequences of neglecting the coupling to other modes in the present method.

The solutions presented in this paper can be used in the plate coupling procedures. This is because the mode shapes are satisfying both the displacement and force boundary conditions. Also, the cases of one or two free edges are considered. These cases are important in the plate coupling procedures where the coupling edges are considered free and the coupling forces are applied to these free edges. The results of the present analysis lay grounds for further investigations of in-plane vibrational power flow characteristics in single and coupled plates.<sup>14</sup>

## ACKNOWLEDGMENT

Support for this work from the Australian Research Council, the Engineering Foundation of Western Australia and Transfield (Australian Shipbuilding Industries W.A.) is gratefully acknowledged.

## APPENDIX: LIST OF SYMBOLS

$a, b, h$	plate dimensions along $X, Y, Z$ directions
$\rho$	material density (mass per unit volume)
$\nu$	Poisson's ratio
$\eta$	structural (hysteretic) damping coefficient
$\omega$	angular frequency (radian per second)
$C_L^2 = E/\rho(1 - \nu^2)$	square of the longitudinal wave speed in plates
$C_T^2 = E/2\rho(1 + \nu)$	square of the transverse (in-plane) shear wave speed in plates
$E$	Young's modulus of elasticity of the plate material
$N_{xx}, N_{yy}, N_{xy}$	Intensities of in-plane internal forces (force per unit length) in the directions shown in Fig. 3
$m, n$	number of half-wave lengths in $X$ direction for in-plane response $u$ and $v$ , respectively
$p, s$	indicators of the number of half-wave lengths for $u$ and $v$ in-plane response in $Y$ direction
$u, v$ and $\ddot{u}, \ddot{v}$	components of in-plane response displacement in the directions of the coordinate axes and corresponding components of the acceleration response
$U, V$	mode shape functions of $u$ and $v$ in the $Y$ direction
$U', V'$	first derivatives of $U$ and $V$ with respect to $y$
$(X_e, Y_e)$	components of in-plane excitation force vector in $X$ and $Y$ directions
$(x_e, y_e)$	excitation position on the middle plane of the plate

<sup>1</sup>A. Leissa, *Vibration of Plates* (Acoustical Society of America, Woodbury, NY, 1993).

<sup>2</sup>A. Leissa, "The free vibration of rectangular plates," *J. Sound Vib.* **31**, 257–295 (1973).

<sup>3</sup>L. Cremer, M. Heckl, and E. E. Ungar, *Structure-Borne Sound* (Springer-Verlag, Berlin, 1988).

<sup>4</sup>K. F. Graff, *Wave Motion in Elastic Solids* (Clarendon, Oxford, 1975).

<sup>5</sup>W. Soedel, *Vibrations of Shells and Plates* (Marcel Dekker, New York, 1981).

<sup>6</sup>J. L. Guyader, C. Boisson, and C. Lesueur, "Energy transmission in finite coupled plates, Part 1: Theory," *J. Sound Vib.* **81**, 81–92 (1982).

<sup>7</sup>R. H. Lyon, "In-plane contribution to structural noise transmission," *Noise Control Eng. J.* **26**, 22–27 (1986).

<sup>8</sup>B. A. Ovunc, "In-plane vibration of plates under external disturbances applied at singular points," *J. Thin-Walled Structures.* **18**, 83–95 (1994).

<sup>9</sup>J. S. Kim, H. S. Kim, H. J. Kang, and S. R. Kim, "Effect of inplane modes in SEA on structure-borne noise transmission in ship structures," 4th international congress on sound and vibration, St. Petersburg, Russia, pp. 217–222 (1996).

<sup>10</sup>J. M. Cuschieri and M. D. McCollum, "In-plane and out-of-plane waves"

- power transmission through L-plate junctions using the mobility power flow approach," J. Acoust. Soc. Am. **100**, 857–870 (1996).
- <sup>11</sup>N. H. Farag and J. Pan, "Free and forced in-plane vibration of rectangular plates," J. Acoust. Soc. Am. **103**, 408–413 (1998).
- <sup>12</sup>S. Azimi, J. F. Hamilton, and W. Soedel, "The receptance method applied to the free vibration of continuous rectangular plates," J. Sound Vib. **93**, 9–29 (1984).
- <sup>13</sup>S. Ljunggren, "Forced vibrations of infinite plates," J. Sound Vib. **121**, 221–236 (1988).
- <sup>14</sup>N. H. Farag and J. Pan, "On the free and forced vibration of single and coupled rectangular plates," J. Acoust. Soc. Am. **104**, 204–216 (1998).
- <sup>15</sup>C. Hwang and W. S. Pi, "Investigation of vibrational energy transfer in coupled structures," NASA CR-124450, Final report, 1973.

# Frequency constants for transverse vibration of annular disks

Thomas B. Gabrielson

Applied Research Laboratory, The Pennsylvania State University, P.O. Box 30, State College, Pennsylvania 16804

(Received 9 October 1998; revised 12 February 1999; accepted 25 February 1999)

Frequency constants for the transverse vibration of thin, annular disks are presented. All nine combinations of the boundary conditions, free, simply supported, and clamped, are included for Poisson's ratios of 0.0, 0.3, and 0.5. Many of these cases have been treated earlier in the literature but with occasional, significant error. Consequently, several tests of the validity of the present solutions are described. © 1999 Acoustical Society of America. [S0001-4966(99)01206-0]

PACS numbers: 43.40.Dx, 46.70.De, 43.38.Ar [CBB]

## INTRODUCTION

Solution for the frequency constants for transverse vibration of annular disks is not new.<sup>1-8</sup> Frequency constants for many cases are given in a number of references, including the excellent compilation by Leissa.<sup>5</sup> Unfortunately, there are errors in many of the original works that have propagated into standard references. In many cases, three or four digits are published but only two of the digits are reliable. In a few cases, there are errors of 20% to 40%. Some of these errors may be simple typographic errors in the original references. Some are probably errors from inaccurate approximations for the relevant Bessel functions. Accurate values for Bessel functions have been available<sup>9-11</sup> since the late 1800's, but lower-quality approximations (by polynomial approximation,<sup>11,12</sup> poorly converging series solutions, or propagated errors from repeated application of recursion formulas, for example) may be used in the actual computer implementation. Other errors may be the result of inaccurate calculation of the roots of the eigenvalue determinant.

It is the intent of this paper to present a reasonable survey of frequency constants for the transverse vibration of annular (circular) disks with all combinations of the boundary conditions: free, clamped, and simply supported. Interpolation will handle many intermediate cases but a strategy for mechanizing the computation is also outlined so that specific cases can be computed directly if necessary (hopefully, after checking a few nearby values against the values presented here).

How many digits should be published for the frequency constants? Two or three are sufficient for most engineering design purposes but engineering design is not the only application for such data. Careful measurement of resonances in well controlled circumstances can produce agreement in frequency ratios (e.g., ratio of  $n$ th-mode frequency to first-mode frequency) to three digits so at least four digits should be published. Approximation of resonance frequencies by Rayleigh-Ritz can also achieve three-digit accuracy with relatively simple, adjustable guess functions (and greater accuracy with more sophisticated functions<sup>8</sup>), so at least four digits are useful. Consequently, in the tables included in this paper, five digits will be given. The quality checks discussed below indicate that the worst case solution underlying any tabulated value is still reliable to at least six digits.

## I. FORMULATION

The governing differential equation and its origins are well known for thin, circular plates<sup>1-5</sup> and will not be repeated here. The general solution for the transverse displacement,  $w$ , as a function of radius,  $r$ , and angle,  $\theta$ , is

$$w(r, \theta) = \cos(n\theta)[A_{n1}J_n(kr) + A_{n2}Y_n(kr) + A_{n3}I_n(kr) + A_{n4}K_n(kr)], \quad (1)$$

where

$$k^4 = \frac{12(1-\nu^2)\rho\omega^2}{Eh^2}, \quad (2)$$

and  $E$ ,  $\rho$ , and  $\nu$  are the elastic modulus, density, and Poisson's ratio for the disk material;  $h$  is the disk thickness, and  $\omega$  is the angular frequency.  $J_n$  and  $Y_n$  are the ordinary Bessel functions of the first and second kind, while  $I_n$  and  $K_n$  are the modified Bessel functions. The definition of  $K_n$  differs (in sign of the odd orders) between Abramowitz and Stegun<sup>11</sup> and Whittaker and Watson.<sup>10</sup> Here, the convention of Abramowitz and Stegun will be followed; some older papers<sup>3</sup> use the Whittaker and Watson convention.

The boundary conditions are satisfied by specifying two conditions at the outer edge and two conditions at the inner edge. At each edge, the two conditions are selected from a set of four functions: (1) the displacement,  $D$ ; (2) the radial slope,  $S$ ; (3) the radial bending moment,  $M$ , and (4) the radial Kelvin-Kirchhoff shear,  $V$ . For a free edge,  $M$  and  $V$  are zero; for a clamped edge,  $D$  and  $S$  are zero; and for a simply supported edge,  $D$  and  $M$  are zero.

For the purposes of making the calculations systematic, the following definitions will be made: First, three auxiliary functions,  $p_n$ ,  $q_n$ , and  $t_n$ , are defined as

$$p_n(kr) = n/(kr), \quad (3)$$

$$q_n(kr) = (1-\nu)/(kr), \quad (4)$$

$$t_n(kr) = n(n-1)(1-\nu)/(kr)^2. \quad (5)$$

Then, omitting the arguments, functional forms for the boundary condition terms are introduced:

$$\begin{aligned} D_{n1} &= J_n; & D_{n2} &= Y_n, \\ D_{n3} &= I_n; & D_{n4} &= K_n, \end{aligned} \quad (6)$$

$$S_{n1} = p_n J_n - J_{n+1}; \quad S_{n2} = p_n Y_n - Y_{n+1}, \quad (7)$$

$$S_{n3} = p_n I_n + I_{n+1}; \quad S_{n4} = p_n K_n - K_{n+1},$$

$$M_{n1} = (t_n - 1)J_n + q_n J_{n+1}; \quad M_{n2} = (t_n - 1)Y_n + q_n Y_{n+1}, \\ M_{n3} = (t_n + 1)I_n - q_n I_{n+1}; \quad M_{n4} = (t_n + 1)K_n + q_n K_{n+1}, \quad (8)$$

$$V_{n1} = p_n(t_n + 1)J_n - (np_n q_n + 1)J_{n+1}, \\ V_{n2} = p_n(t_n + 1)Y_n - (np_n q_n + 1)Y_{n+1}, \\ V_{n3} = p_n(t_n - 1)I_n + (np_n q_n - 1)I_{n+1}, \quad (9)$$

$$V_{n4} = p_n(t_n - 1)K_n - (np_n q_n - 1)K_{n+1}.$$

These functions are combined to form the appropriate boundary conditions. For example, the condition for zero displacement at  $r = a$  would be written:

$$\sum_{i=1}^4 A_{ni} D_{ni}(ka) = 0. \quad (10)$$

The frequency constants are found by searching for the zeroes of the determinant of the fourth-order matrix formed by the boundary-condition functional forms. For an annular disk with inner radius,  $b$ , and outer radius,  $a$ , define a frequency constant,  $\lambda^2$ , and a radius ratio,  $\alpha$ :

$$\lambda^2 = (ka)^2; \quad \alpha = b/a. \quad (11)$$

If, for example, the inner edge is clamped and the outer edge is free, the matrix,  $F$ , of boundary-condition functions would be

$$F = \begin{bmatrix} D_{n1}(\alpha\lambda) & D_{n2}(\alpha\lambda) & D_{n3}(\alpha\lambda) & D_{n4}(\alpha\lambda) \\ S_{n1}(\alpha\lambda) & S_{n2}(\alpha\lambda) & S_{n3}(\alpha\lambda) & S_{n4}(\alpha\lambda) \\ M_{n1}(\lambda) & M_{n2}(\lambda) & M_{n3}(\lambda) & M_{n4}(\lambda) \\ V_{n1}(\lambda) & V_{n2}(\lambda) & V_{n3}(\lambda) & V_{n4}(\lambda) \end{bmatrix}. \quad (12)$$

The frequency constants are then found from the roots of the determinant:

$$|F| = 0. \quad (13)$$

The frequency constants,  $\lambda_{nm}^2$ , are independent of the disk dimensions and the disk material except for a modest dependence on Poisson's ratio. The actual frequencies can be found from the frequency constants as follows:

$$f_{nm} = \frac{\lambda_{nm}^2 h}{2\pi a^2} \sqrt{\frac{E}{12\rho(1-\nu^2)}}. \quad (14)$$

These frequencies describe modes with  $n$  nodal diameters and  $m$  nodal circles. (Throughout this paper,  $\rho$  is the mass per unit volume; many references use mass per unit area.)

## II. QUALITY CHECKS

With modern mathematical analysis software, solutions for the zeroes of the frequency equation are simple to obtain. It is not, however, trivial to ensure that the software algorithms produce *accurate* solutions. Because of the impor-

tance of avoiding the introduction of still more erroneous values into the literature, quality checks were performed.

As a basic test of the mathematical analysis packages, many values of the Bessel functions,  $J$ ,  $Y$ ,  $I$ , and  $K$  were calculated for orders 0, 1, and 2 and compared to tabulated values.<sup>11</sup> Beyond this obvious test, a number of other checks were applied to ensure accurate solutions. Several of these checks are sufficiently general that they can be applied to many problems in numerical root finding.

### A. Wronskians of the Bessel functions

Point-by-point comparison of calculated Bessel-function values with tabulated values is tedious at best. The critical values of comparison are generally in the vicinity of those arguments that are at points of transition in the form of approximation used by the analysis software. However, the forms of approximation and the arguments at which transitions from one form to another occur are often not available to the user. Consequently, it is useful to have a means for scanning large ranges of argument in fine steps with reasonable speed. In order to do this, the Wronskians<sup>11</sup> of the Bessel functions were computed. For each pair of functions, the true Wronskian was subtracted from the computed result to form two test functions,  $W_{JY}$  for  $J_n$  and  $Y_n$  and  $W_{IK}$  for  $I_n$  and  $K_n$ :

$$W_{JY}(z) = J_{n+1}(z)Y_n(z) - J_n(z)Y_{n+1}(z) - 2/(\pi z), \quad (15)$$

$$W_{IK}(z) = I_{n+1}(z)K_n(z) + I_n(z)K_{n+1}(z) - 1/z. \quad (16)$$

The logarithm to the base 10 of either quantity roughly indicates the number of reliable decimal places in the Wronskian calculation. For example, if the  $\log_{10}(W_{IK})$  is  $-14$  for some  $z$ , then about 14 significant figures are reliable. This is actually a harsh test of the software calculation because, particularly for  $W_{IK}$ , the result depends on subtraction of two nearly equal, large numbers to form a small result. Such differences are always suspect operations insofar as loss of precision is concerned.

Plotting the  $\log_{10}$  of  $W_{IK}$  revealed a significant loss of precision in one of the software packages for  $z$  in the vicinity of 8. Since the Wronskian calculation may introduce loss of precision even if the Bessel functions are calculated correctly, at any points of concern, values for the individual Bessel functions were calculated and compared to 15-digit values in Abramowitz and Stegun. This isolated the problem to the calculation of  $K_n$  and a correction was developed. The Wronskian test was not used as a substitute for manually checking computed values against tabulated values. Instead, the test was used to scan quickly (by generating a plot of the error) over the entire range of relevant argument and order of each type of Bessel function with a much finer resolution in argument than is possible with checks against tables.

### B. Subtraction of nearly equal numbers

Most loss-of-precision problems in numerical algorithms can be traced to results that depend on subtraction of two

TABLE I. Frequency constants for annular disks with free inner boundary and free outer boundary.

$\nu$	$m$	$n$	Radius ratio, $a$				
			0	0.1	0.3	0.5	0.7
0.0	0	0	-	-	-	-	-
0.0	0	1	-	-	-	-	-
0.0	0	2	6.1530	6.0670	5.4768	4.6416	3.8262
0.0	0	3	13.956	13.952	13.657	12.393	10.473
0.0	0	4	24.124	24.124	24.026	22.835	19.784
0.0	0	5	36.600	36.600	36.572	35.677	31.694
0.0	1	0	8.2439	8.1823	8.2895	9.5655	13.721
0.0	1	1	20.060	20.024	18.932	18.906	24.733
0.0	1	2	35.442	34.936	33.240	33.970	43.603
0.3	0	0	-	-	-	-	-
0.3	0	1	-	-	-	-	-
0.3	0	2	5.3583	5.3034	4.9060	4.2711	3.5725
0.3	0	3	12.439	12.437	12.266	11.425	9.8589
0.3	0	4	21.835	21.835	21.783	21.067	18.697
0.3	0	5	33.495	33.495	33.481	32.982	30.025
0.3	1	0	9.0031	8.7745	8.3535	9.3135	13.163
0.3	1	1	20.475	20.406	18.292	17.198	21.914
0.3	1	2	35.260	34.931	32.973	31.115	37.841
0.5	0	0	-	-	-	-	-
0.5	0	1	-	-	-	-	-
0.5	0	2	4.6404	4.6016	4.3107	3.8052	3.2039
0.5	0	3	10.941	10.939	10.829	10.221	8.9035
0.5	0	4	19.415	19.415	19.385	18.908	16.965
0.5	0	5	30.027	30.027	30.020	29.710	27.351
0.5	1	0	9.4137	8.9544	7.9537	8.5969	12.003
0.5	1	1	20.722	20.610	17.313	15.394	19.212
0.5	1	2	35.145	34.917	32.373	28.159	32.801

nearly equal numbers.<sup>13</sup> Such situations arise in two places in evaluation of the frequency constants: (1) in the calculation of the constituent functions,  $S$ ,  $M$ , and  $V$ ; and (2) in the calculation of the determinant. Since the inner workings of the determinant evaluation are not available to the user in most software packages, an explicit, step-by-step expansion of the determinant was programmed so that each differencing operation could be examined individually.

Each subtraction (actually, each sum or difference since the sign of the terms is not predetermined) is evaluated in the following manner. Consider the subtraction of two terms,  $A - B$ . The following quantity is computed:

$$T_{AB} = \log_{10} \left| \frac{|A| + |B|}{A - B} \right|. \quad (17)$$

This test value gives (roughly) the number of digits *lost* in a fixed-precision, floating-point operation. This test was performed for each step of each calculation. The integer sum of all the values gives an overestimate of the number of digits lost in the calculation because summing assumes that each operation has the same significance in the final result. Calculations were performed with 16 decimal-digit computations; in a few cases in which the test indicated a loss of more than 10 digits, the calculations were repeated using 20-digit computations. As expected, the actual loss in precision was not as great as predicted from the sum all of the values for  $T_{AB}$  from a given calculation.

TABLE II. Frequency constants for annular disks with simply supported inner boundary and simply supported outer boundary.

$\nu$	$m$	$n$	Radius ratio, $a$			
			0.1	0.3	0.5	0.7
0.0	0	0	14.693	21.357	40.261	110.23
0.0	0	1	16.774	23.495	41.985	111.61
0.0	0	2	25.636	30.246	47.200	115.73
0.0	0	3	39.626	41.714	55.973	122.61
0.0	0	4	56.497	57.262	68.304	132.25
0.0	0	5	75.864	76.108	84.106	144.67
0.0	1	0	52.674	82.329	159.05	439.49
0.0	1	1	57.027	85.163	160.96	440.90
0.0	1	2	71.612	93.778	166.71	445.14
0.3	0	0	14.485	21.079	40.043	110.06
0.3	0	1	16.776	23.317	41.797	111.44
0.3	0	2	25.936	30.273	47.089	115.59
0.3	0	3	39.976	41.910	55.957	122.49
0.3	0	4	56.842	57.546	68.379	132.17
0.3	0	5	76.203	76.427	84.257	144.63
0.3	1	0	51.781	81.737	158.64	439.18
0.3	1	1	56.507	84.635	160.57	440.59
0.3	1	2	71.687	93.417	166.35	444.84
0.5	0	0	14.110	20.776	39.827	109.90
0.5	0	1	16.654	23.103	41.606	111.29
0.5	0	2	26.110	30.237	46.959	115.44
0.5	0	3	40.197	42.013	55.904	122.37
0.5	0	4	57.063	57.721	68.399	132.08
0.5	0	5	76.421	76.630	84.336	144.56
0.5	1	0	50.912	81.268	158.33	438.95
0.5	1	1	55.992	84.215	160.27	440.37
0.5	1	2	71.696	93.122	166.07	444.62

### C. Location of true zeros

Although it is rare that modern analysis packages fail in this manner, each root was tested to ensure that the root actually represented a zero of the function. This was done by adding and subtracting one part per million and recalculating the value of the determinant. The two values should have opposite sign and nearly equal magnitude. If a root passes this test, it is not guaranteed to be correct but a failure would indicate a problem. No cases ultimately failed this test but the test did indicate several instances in which a tighter convergence tolerance was required in the root finder.

### D. Energy-method verification

Once the frequency constant is determined, the corresponding deflection function can be determined. The frequency constants come from the eigenvalue solution; the deflection function comes from the eigenfunction solution. The eigenfunction can be computed using an eigenfunction routine in the analysis package or the appropriate coefficients,  $A_{ni}$ , can be determined by assuming a value for one of the coefficients, by reducing the  $4 \times 4$  homogeneous system to a  $3 \times 3$  inhomogeneous system and by solving the resulting matrix equation.

Having the displacement function permits an independent check on the frequency constant. The frequency of a



TABLE III. Frequency constants for annular disks with clamped inner boundary and clamped outer boundary. These solutions are independent of Poisson's ratio.

$\nu$	$m$	$n$	Radius ratio, $a$			
			0.1	0.3	0.5	0.7
-	0	0	27.281	45.346	89.251	248.40
-	0	1	28.916	46.644	90.230	249.16
-	0	2	36.617	51.139	93.321	251.48
-	0	3	51.219	60.033	98.928	255.44
-	0	4	69.678	73.945	107.57	261.20
-	0	5	90.740	92.495	119.70	268.92
-	1	0	75.366	125.36	246.34	684.99
-	1	1	78.635	127.38	247.74	686.04
-	1	2	90.448	133.67	251.97	689.19

mode can be expressed in terms of the stored potential energy of flexure and the kinetic energy of motion of the disk. Both of these energy terms can be obtained by integrals involving the displacement function and its derivatives. The potential energy,  $U$ , resulting from flexure is<sup>3-5</sup>

$$U = \frac{Eh^3}{24(1-\nu^2)} \hat{U}, \quad (18)$$

where

TABLE IV. Frequency constants for annular disks with free inner boundary and simply supported outer boundary.

$\nu$	$m$	$n$	Radius ratio, $a$				
			0	0.1	0.3	0.5	0.7
0.0	0	0	4.4436	4.4279	4.4994	5.1040	7.1366
0.0	0	1	13.501	13.488	12.947	12.485	14.841
0.0	0	2	25.245	24.909	23.522	23.268	27.209
0.0	0	3	39.603	39.569	38.003	35.954	40.755
0.0	0	4	56.496	56.494	55.647	51.726	55.624
0.0	0	5	75.864	75.864	75.529	70.907	72.289
0.0	1	0	29.364	29.751	38.174	67.118	177.23
0.0	1	1	48.136	47.898	48.036	72.030	180.22
0.0	1	2	69.782	68.499	66.850	85.023	188.92
0.3	0	0	4.9351	4.8533	4.6641	5.0769	6.9310
0.3	0	1	13.898	13.872	12.816	11.607	13.311
0.3	0	2	25.613	25.394	24.116	22.357	24.329
0.3	0	3	39.957	39.935	38.775	35.636	37.097
0.3	0	4	56.842	56.840	56.250	52.032	51.713
0.3	0	5	76.203	76.203	75.975	71.638	68.502
0.3	1	0	29.720	29.438	37.042	65.842	175.64
0.3	1	1	48.479	48.012	45.837	69.891	178.07
0.3	1	2	70.117	69.260	65.038	81.100	185.24
0.5	0	0	5.2127	5.0367	4.5412	4.7588	6.3736
0.5	0	1	14.141	14.097	12.410	10.564	11.769
0.5	0	2	25.844	25.693	24.284	20.951	21.510
0.5	0	3	40.182	40.167	39.167	34.414	33.238
0.5	0	4	57.062	57.061	56.595	51.209	47.084
0.5	0	5	76.421	76.421	76.248	71.225	63.335
0.5	1	0	29.946	28.861	35.951	64.805	174.46
0.5	1	1	48.699	47.923	43.612	68.168	176.50
0.5	1	2	70.333	69.712	62.378	77.757	182.55

TABLE V. Frequency constants for annular disks with free inner boundary and clamped outer boundary.

$\nu$	$m$	$n$	Radius ratio, $a$				
			0	0.1	0.3	0.5	0.7
0.0	0	0	10.216	10.311	12.030	18.626	44.503
0.0	0	1	21.260	21.226	20.508	23.737	47.440
0.0	0	2	34.877	34.349	32.580	34.383	55.354
0.0	0	3	51.030	50.966	48.523	47.593	66.676
0.0	0	4	69.666	69.662	68.148	63.906	80.456
0.0	0	5	90.739	90.739	90.080	83.894	96.448
0.0	1	0	39.771	40.445	53.537	95.873	254.36
0.0	1	1	60.829	60.435	62.700	100.17	256.91
0.0	1	2	84.583	82.952	82.176	112.09	264.44
0.3	0	0	10.216	10.159	11.424	17.715	43.142
0.3	0	1	21.260	21.195	19.540	22.015	45.332
0.3	0	2	34.877	34.535	32.594	32.116	51.585
0.3	0	3	51.030	50.990	49.069	45.812	61.290
0.3	0	4	69.666	69.663	68.584	63.018	74.009
0.3	0	5	90.739	90.739	90.288	83.814	89.579
0.3	1	0	39.771	39.521	51.745	93.847	251.64
0.3	1	1	60.829	60.061	59.759	97.376	253.72
0.3	1	2	84.583	83.478	79.061	107.49	259.91
0.5	0	0	10.216	9.970	10.827	16.958	42.133
0.5	0	1	21.260	21.152	18.447	20.485	43.734
0.5	0	2	34.877	34.642	32.182	29.682	48.533
0.5	0	3	51.030	51.003	49.197	43.235	56.526
0.5	0	4	69.666	69.664	68.767	60.827	67.717
0.5	0	5	90.739	90.739	90.387	82.211	82.115
0.5	1	0	39.771	38.451	50.233	92.337	249.74
0.5	1	1	60.829	59.559	57.093	95.272	251.49
0.5	1	2	84.583	83.763	75.348	103.84	256.71

$$\hat{U} = \int \int \left\{ [\nabla^2 w]^2 - 2(1-\nu) \left[ \frac{\partial^2 w}{\partial r^2} \left( \frac{1}{r} \frac{\partial w}{\partial r} + \frac{1}{r^2} \frac{\partial^2 w}{\partial \theta^2} \right) - \left( \frac{\partial}{\partial r} \left( \frac{1}{r} \frac{\partial w}{\partial \theta} \right) \right)^2 \right] \right\} dS \quad (19)$$

in polar coordinates. The integration over angle was done analytically. Furthermore, the integration over angle produces the same factor for the potential energy and the kinetic energy so that factor cancels in the frequency-constant calculation. All of the derivatives were expanded by hand using either the differential equations satisfied by the Bessel functions (Eqs. 9.1.1 and 9.6.1 in Abramowitz and Stegun) or the recursion relations (Eqs. 9.1.27 and 9.6.26 in Abramowitz and Stegun). Then the integration over the radial coordinate was done numerically. For a few cases, namely those for which the inner edge was free,  $\alpha=0.1$ , and  $n=4$  or  $5$ , the convergence of the integration was slow. In these cases, there is a strong concentration of bending potential energy near the inner edge. By splitting the integration into two sections—a small inner section at the strain-energy concentration and another section covering the remainder of the radial coordinate—the integrations were performed without difficulty.

The kinetic energy,  $T$ , is

TABLE VI. Frequency constants for annular disks with simply supported inner boundary and free outer boundary.

$\nu$	$m$	$n$	Radius ratio, $a$			
			0.1	0.3	0.5	0.7
0.0	0	0	3.4119	3.5694	4.3562	6.5416
0.0	0	1	2.6401	3.7903	5.5502	9.5752
0.0	0	2	6.2422	7.0680	9.4160	15.959
0.0	0	3	13.959	14.203	16.088	24.110
0.0	0	4	24.124	24.187	25.459	33.988
0.0	0	5	36.600	36.615	37.400	45.768
0.0	1	0	20.894	31.480	60.602	168.59
0.0	1	1	24.290	34.572	63.129	170.73
0.0	1	2	36.063	43.576	70.455	177.04
0.3	0	0	3.4497	3.4223	4.1210	6.1867
0.3	0	1	2.4377	3.3738	4.8616	8.3506
0.3	0	2	5.4277	6.0804	7.9861	13.427
0.3	0	3	12.441	12.611	14.035	20.454
0.3	0	4	21.835	21.876	22.788	29.403
0.3	0	5	33.495	33.504	34.047	40.387
0.3	1	0	20.889	31.603	61.009	169.49
0.3	1	1	24.244	34.457	63.227	171.31
0.3	1	2	35.774	42.937	69.766	176.70
0.5	0	0	3.3360	3.1367	3.7283	5.5878
0.5	0	1	2.2246	2.9782	4.2337	7.2425
0.5	0	2	4.6963	5.2284	6.8091	11.391
0.5	0	3	10.942	11.067	12.181	17.442
0.5	0	4	19.415	19.442	20.119	25.394
0.5	0	5	30.027	30.033	30.417	35.343
0.5	1	0	20.590	31.541	61.180	170.01
0.5	1	1	24.063	34.274	63.211	171.62
0.5	1	2	35.572	42.470	69.264	176.44

$$T = \frac{1}{2} \rho h \omega^2 \hat{T}, \quad (20)$$

where

$$\hat{T} = \int \int w^2 dS. \quad (21)$$

Using these definitions, the frequency constant is

$$k^2 = (\lambda/a)^2 = \sqrt{\hat{U}/\hat{T}}. \quad (22)$$

If the deflection function is exact, the resulting frequency will be exact (for any mode, not just the fundamental). All of the frequency constants were verified to six digits in this manner. Selected values were verified to eight digits.

As a further check, several solutions were examined to ensure that the energy-derived frequency constants represented local stationary points with respect to perturbations of the deflection function. For the stationarity property to hold, the perturbed deflection function must still satisfy any constraining boundary conditions (zero displacement or zero slope) but need not satisfy the “natural” boundary conditions<sup>14</sup> (zero bending moment or zero Kelvin–Kirchhoff shear). Consequently, the simplest case to examine is the free–free case because the true deflection function must represent a stationary point with respect to any perturbation. Once what is assumed to be the exact deflection function [i.e., Eq. (1)] is determined, each of the four constants,

TABLE VII. Frequency constants for annular disks with simply supported inner boundary and clamped outer boundary.

$\nu$	$m$	$n$	Radius ratio, $a$			
			0.1	0.3	0.5	0.7
0.0	0	0	23.541	34.691	64.979	175.77
0.0	0	1	25.834	36.717	66.453	176.86
0.0	0	2	35.521	43.277	70.998	180.18
0.0	0	3	51.074	54.896	78.897	185.78
0.0	0	4	69.668	71.185	90.413	193.79
0.0	0	5	90.739	91.259	105.66	204.33
0.0	1	0	67.011	105.31	203.13	559.45
0.0	1	1	71.680	108.06	204.89	560.71
0.0	1	2	87.053	116.44	210.22	564.51
0.3	0	0	22.701	33.765	63.973	174.41
0.3	0	1	25.283	35.906	65.486	175.52
0.3	0	2	35.406	42.731	70.136	178.87
0.3	0	3	51.065	54.609	78.182	184.53
0.3	0	4	69.667	71.063	89.857	192.61
0.3	0	5	90.739	91.216	105.25	203.23
0.3	1	0	65.639	104.22	202.07	558.10
0.3	1	1	70.689	107.03	203.85	559.37
0.3	1	2	86.716	115.58	209.21	563.17
0.5	0	0	21.899	33.055	63.255	173.47
0.5	0	1	24.792	35.288	64.796	174.59
0.5	0	2	35.315	42.324	69.523	177.97
0.5	0	3	51.058	54.400	77.677	183.67
0.5	0	4	69.667	70.976	89.467	191.80
0.5	0	5	90.739	91.185	104.97	202.48
0.5	1	0	64.460	103.44	201.33	557.19
0.5	1	1	69.861	106.29	203.12	558.45
0.5	1	2	86.452	114.96	208.52	562.27

$A_{ni}$ , and the frequency constant,  $k^2$ , can be varied (one at a time) and the modified  $k^2$  computed from Eq. (22). For both positive and negative perturbation of the  $A_{ni}$ 's or  $k^2$ , the modified  $k^2$  computed from Eq. (22) must be stationary with respect to the zero-perturbation value, otherwise the initial solution is incorrect. Each of the instances examined in this manner proved to correspond to proper stationary points. This process was also applied to several instances of free-clamped and free–simply supported disks by forcing the constraining boundary conditions to be satisfied under perturbation and examining variations in  $k^2$  about the assumed exact value.

### E. Convergence on alternate variables

Normally, the exact solutions are found by applying the root finder to the parameter,  $\lambda$ . The root finder convergence can be tested by allowing either  $\alpha$  or  $\nu$  be the free parameter instead. This fixes the values of the Bessel functions and forces the convergence to depend on another variable in the determinant. The procedure is as follows: (1) the “exact” solution is found for a particular set of  $\alpha$  and  $\nu$ ; (2) that value for  $\lambda$  and the previous value for  $\nu$  are set and the root finder is permitted to vary  $\alpha$ ; and (3) the values for  $\lambda$  and  $\alpha$  are set and the root finder is applied to  $\nu$ . In each case, the same values should result. This test passed for each case to at least eight digits.

TABLE VIII. Frequency constants for annular disks with clamped inner boundary and free outer boundary.

$\nu$	$m$	$n$	Radius ratio, $a$			
			0.1	0.3	0.5	0.7
0.0	0	0	3.9633	6.2383	12.377	35.842
0.0	0	1	3.4605	6.4757	13.063	36.855
0.0	0	2	6.4182	8.7637	15.518	39.891
0.0	0	3	13.969	14.888	20.502	44.966
0.0	0	4	24.125	24.415	28.436	52.139
0.0	0	5	36.600	36.681	39.319	61.508
0.0	1	0	24.436	41.592	83.657	237.64
0.0	1	1	27.167	43.922	85.630	239.36
0.0	1	2	37.111	51.006	91.480	244.49
0.3	0	0	4.2374	6.6604	13.024	36.953
0.3	0	1	3.4781	6.5523	13.290	37.498
0.3	0	2	5.6227	7.9565	14.704	39.277
0.3	0	3	12.451	13.276	18.562	42.654
0.3	0	4	21.836	22.074	25.596	48.071
0.3	0	5	33.495	33.557	35.730	55.880
0.3	1	0	25.262	42.614	85.033	239.84
0.3	1	1	27.673	44.631	86.706	241.28
0.3	1	2	36.941	50.947	91.738	245.59
0.5	0	0	4.3794	6.8904	13.398	37.632
0.5	0	1	3.4883	6.5974	13.427	37.898
0.5	0	2	4.9156	7.3057	14.115	38.862
0.5	0	3	10.951	11.734	16.915	40.974
0.5	0	4	19.415	19.621	22.862	44.840
0.5	0	5	30.027	30.077	31.943	51.004
0.5	1	0	25.750	43.237	85.894	241.26
0.5	1	1	27.979	45.068	87.383	242.52
0.5	1	2	36.833	50.910	91.903	246.29

### F. Hand checks of special values

Tabulated values<sup>11</sup> are available for the Bessel functions for arguments in increments of 0.1. Cases in which  $\lambda$  and  $\alpha\lambda$  are “even” to a 0.1 increment can be calculated manually from tabulated values. Such cases were isolated by picking trial values for  $\lambda$  and  $\alpha$  and root finding on  $\nu$ . Only cases in which the resulting value for  $\nu$  was between 0 and 0.5 were used for these checks.

### G. Transcription checks

Once the frequency-constant tables were constructed as spreadsheets, text files were generated from those tables. These text files were read by another routine that extracted the appropriate  $\nu$ ,  $n$ , and  $\alpha$  for each  $\lambda^2$  and performed the energy-method check. In this manner, the values in the tables were tested for accuracy. Since it is vital to avoid transcription errors in publication, these table files were transmitted electronically to the publisher.

## III. RESULTS

All of the calculations were performed using either MathCad<sup>15</sup> 4.0 or Mathematica<sup>16</sup> 3.0 on a Pentium-based PC with the default machine precision in numerical calculations (16 decimal digits for Mathematica). For most calculations, both packages were used to compare and cross-check results.

TABLE IX. Frequency constants for annular disks with clamped inner boundary and simply supported outer boundary.

$\nu$	$m$	$n$	Radius ratio, $a$			
			0.1	0.3	0.5	0.7
0.0	0	0	17.332	29.425	59.091	167.38
0.0	0	1	18.960	30.869	60.271	168.35
0.0	0	2	26.334	35.761	63.952	171.29
0.0	0	3	39.706	45.036	70.481	176.29
0.0	0	4	56.503	58.895	80.239	183.47
0.0	0	5	75.864	76.781	93.483	192.96
0.0	1	0	59.770	99.953	197.41	551.47
0.0	1	1	62.872	102.02	198.91	552.63
0.0	1	2	74.277	108.47	203.43	556.11
0.3	0	0	17.789	29.978	59.820	168.52
0.3	0	1	19.394	31.403	60.987	169.49
0.3	0	2	26.717	36.243	64.631	172.41
0.3	0	3	40.062	45.459	71.107	177.38
0.3	0	4	56.848	59.273	80.802	184.51
0.3	0	5	76.203	77.133	93.988	193.95
0.3	1	0	60.144	100.42	198.05	552.51
0.3	1	1	63.243	102.48	199.54	553.67
0.3	1	2	74.631	108.92	204.05	557.14
0.5	0	0	18.072	30.326	60.286	169.27
0.5	0	1	19.663	31.739	61.445	170.23
0.5	0	2	26.957	36.549	65.067	173.14
0.5	0	3	40.287	45.728	71.509	178.09
0.5	0	4	57.068	59.515	81.166	185.19
0.5	0	5	76.421	77.359	94.314	194.60
0.5	1	0	60.385	100.73	198.47	553.19
0.5	1	1	63.482	102.78	199.96	554.35
0.5	1	2	74.859	109.21	204.46	557.82

The resulting frequency constants,  $\lambda_{mn}^2$ , are given in Tables I–IX. Each of the permutations of free, clamped, and simply supported boundary conditions are presented for  $\alpha$  from 0.1 to 0.7 in steps of 0.2 (including  $\alpha=0$  for cases in which this is appropriate) and for Poisson’s ratios of 0.0, 0.3, and 0.5. The  $\alpha=0$  cases<sup>5</sup> were calculated from the  $2 \times 2$  determinants formed from Eq. (1) with  $A_{n2}$  and  $A_{n4}$  set identically to zero and the clamped, simply supported, and free outer edge conditions. The special cases described in Sec. IIF above (for the clamped-inside/free-outside case) are given in Table X. A small set of solutions is given in Table XI to eight digits so that workers developing their own solutions can test them.

The symbology in this paper is generally consistent with Leissa.<sup>5</sup> Here,  $\lambda^2$  is identical to  $(ka)^2$ . In Young,<sup>17</sup> the fre-

TABLE X. Special values of parameters for manual checking of frequency-constant solutions. These frequency constants are for disks with clamped inner edge and free outer edge.

$n$	$\alpha$	$\lambda$	$\alpha\lambda$	$\nu$
0	0.10	2.0	0.2	0.034 953
0	0.40	3.0	1.2	0.286 850
0	0.64	5.0	3.2	0.147 685
0	0.70	6.0	4.2	0.040 742
1	0.40	3.0	1.2	0.040 307
2	0.20	2.5	0.5	0.362 239

TABLE XI. Selected frequency constants with eight digits ( $\nu=0.3$ ). The notation in the BC column gives the boundary conditions. The first letter is for the inner boundary; the second letter is for the outer boundary: C=clamped, F=free, and S=simply supported.

BC	$\alpha=0.1$		$\alpha=0.5$	
	$n=0$	$n=2$	$n=0$	$n=2$
CC	27.280 550	36.617 289	89.250 751	93.321 200
FF	8.774 547 9	5.303 384 5	9.313 480 6	4.271 113 8
SS	14.484 728	25.935 789	40.043 131	47.088 724
CF	4.237 379 8	5.622 703 3	13.024 257	14.703 813
FS	4.853 274 8	25.394 401	5.076 871 3	22.356 569

quency constants are given as  $K_n$  which are also  $(ka)^2$ . Vogel and Skinner<sup>2</sup> use  $\lambda$  to denote their frequency constants but this  $\lambda$  is related to  $ka$  as follows:

$$\lambda_{VS} = \frac{(ka)^2}{\sqrt{3(1-\nu^2)}}. \quad (23)$$

Southwell<sup>3</sup> simply tabulates  $ka$ . Singh and Chakraverty<sup>18</sup> use  $\lambda^2$  and present some comparisons between different authors for the fundamental mode.

A note about the order of the resonances is appropriate. Without treating each case separately, it is not practical to make each list of frequency constants complete from lowest to highest. In some cases, the frequency of one mode is higher than another for some range of  $\alpha$  and lower for some other range. In other cases, some of the modes shift so much with  $\alpha$  that they are beyond the upper range of the table for some  $\alpha$  but in the middle of the range for some other  $\alpha$ . Some of these modes are missing from the tables so it should not be assumed that a table including frequency constants up to some value is complete up to that value.

## ACKNOWLEDGMENTS

Professor Arthur Leissa (Ohio State University) provided several useful suggestions and encouragement throughout this work. The late Charles L. Bartberger taught

the author many techniques for verifying the accuracy of numerical solutions and engendered a healthy skepticism of computer solutions. This study was partially funded by the Naval Sea Systems Command.

- <sup>1</sup>J. W. Strutt, *The Theory of Sound*, Vol. I (Dover, New York, 1945), Paragraphs 218, 219.
- <sup>2</sup>S. M. Vogel and D. W. Skinner, "Natural frequencies of transversely vibrating uniform annular plates," *J. Appl. Mech.* **32**, 926–931 (1965).
- <sup>3</sup>R. V. Southwell, "On the free transverse vibrations of a uniform circular disk clamped at its center; and on the effects of rotation," *Proc. R. Soc. London, Ser. A* **101**, 133–153 (1922).
- <sup>4</sup>H. Lamb and R. V. Southwell, "The vibrations of a spinning disk," *Proc. R. Soc. London, Ser. A* **99**, 272–280 (1921).
- <sup>5</sup>A. Leissa, *Vibration of Plates*, NASA SP-160, U. S. Government Printing Office, 1969; reprinted by the Acoustical Society of America, NY, 1993.
- <sup>6</sup>O. G. McGee, C. S. Huang, and A. W. Leissa, "Comprehensive exact solutions for free vibrations of thick annular sectorial plates with simply supported radial edges," *Int. J. Mech. Sci.* **37**, 537–566 (1995).
- <sup>7</sup>T. Irie, G. Yamada, and K. Takagi, "Natural frequencies of thick annular plates," *J. Appl. Mech.* **49**, 633–637 (1982).
- <sup>8</sup>D. Avalos, H. A. Larrondo, V. Sonzogni, and P. A. A. Laura, "A general approximate solution of the problem of free vibrations of annular plates of stepped thickness," *J. Sound Vib.* **196**, 275–283 (1996).
- <sup>9</sup>W. S. Aldis, "Tables for the solution for the equation  $d^2y/dx^2 + (1/x)*dy/dx - (1+n^2/x^2)*y=0$ ," *Proc. R. Soc. London* **64**, 203–223 (1899).
- <sup>10</sup>E. T. Whittaker and G. N. Watson, *A Course of Modern Analysis*, 4th ed. (Cambridge University Press, Cambridge, 1963), Sec. 17.71.
- <sup>11</sup>M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions* (Dover, New York, 1972), Chap. 9.
- <sup>12</sup>A. J. M. Hitchcock, "Polynomial approximations to Bessel functions of order zero and one and to related functions," in *Mathematical Tables and Other Aids to Computation*, Vol. 11 (National Research Council, 1957), pp. 86–88.
- <sup>13</sup>F. S. Acton, *Real Computing Made Real* (Princeton University Press, Princeton, 1996).
- <sup>14</sup>R. Weinstock, *Calculus of Variations* (Dover, New York, 1974), Section 10-9.
- <sup>15</sup>MathSoft, Inc., Cambridge, MA. Specific mention of any commercial product should not be construed to be an endorsement or recommendation of that product by The Pennsylvania State University.
- <sup>16</sup>Wolfram Research, Inc., Champaign, IL.
- <sup>17</sup>W. C. Young, *Roark's Formulas for Stress and Strain*, 6th ed. (McGraw-Hill, New York, 1989), Table 36.
- <sup>18</sup>B. Singh and S. Chakraverty, "Transverse vibration of annular circular and elliptic plates using the characteristic orthogonal polynomials in two dimensions," *J. Sound Vib.* **162**, 537–546 (1993).

# Analysis of acousto-ultrasonic characteristics for an isotropic thin plate

T. Liu<sup>a)</sup>

Department of Aircraft Engineering, Nanjing University of Aeronautics & Astronautics, Nanjing, 210016, People's Republic of China

K. M. Liew

Centre for Advanced Numerical Engineering Simulations (CANES), School of Mechanical and Production Engineering, Nanyang Technological University, Singapore 639798

S. Kitipornchai<sup>b)</sup> and G. Wang

Department of Civil Engineering, The University of Queensland, Brisbane, Qld. 4072, Australia

(Received 13 March 1998; revised 9 January 1999; accepted 11 March 1999)

For the purpose of understanding and improving the acousto-ultrasonic (AU) technique, the AU characteristics for an isotropic plate is investigated in this paper. Unlike the previous wave tracing method, a spectral analysis approach is presented, which can take into account the effects of the transducers' characteristics. Using this approach and introducing the point source/point receiver assumption, the input-output frequency spectrum relationships for the cases of one transmitter/one receiver and one transmitter/multiple receivers coupled to an isotropic thin plate are expressed in explicit forms. Based on this, a wave scattering parameter to bridge two different sensing types of the receiving transducers is presented, and a new type stress wave parameter for the AU technique is developed, which eliminates the influence of the complex mechanic-electron transduction characteristics of the receivers and also separates to some extent the effects of some other external factors such as transducer masses and locations. Finally, some numerical simulations are carried out and conclusions drawn concerning the AU technique. © 1999 Acoustical Society of America.

[S0001-4966(99)03906-5]

PACS numbers: 43.40.Dx, 43.35.Gg, 43.20.Jr [CBB]

## INTRODUCTION

In recent years, an analytical nondestructive evaluation technique, known as the acousto-ultrasonic (AU) technique or stress wave factor technique, has been receiving considerable attention (Vary, 1991). This technique is especially useful for providing a measurable quantitative parameter which correlates well with mechanical properties of various advanced materials. This technique is a combination of acoustic emission and ultrasonic material characterization methodology. It involves exciting ultrasonic interrogation pulses at one position on a material surface by means of a transmitting transducer and sensing the resulting disturbance stress waves at another position on the same material surface using a receiving transducer. In principle, this technique consists of three physical processes: wave generation, wave propagation, and wave reception.

To increase the theoretical understanding and to enable the development of measurement systems for general applications, several aspects of the basic physical processes mentioned above have been studied in previous theoretical investigations. The major attention has clearly been paid to investigate the propagation behavior of stress waves in various homogeneous and inhomogeneous materials, e.g., Duke *et al.* (1989) and Datta *et al.* (1988, 1992). Many of these

studies have resulted in the development of special purpose measurement systems which were used to determine, for example, the elastic properties of composite materials or characterize the mechanical behavior of specific materials and structures, e.g., Veidt *et al.* (1990,1994). In order to specifically describe the wave generation and reception processes, some theoretical models for the transducers have been examined, e.g., Rose *et al.* (1994 and Pelts and Rose, 1996). With respect to synthetic studies of the whole three stages of the AU technique, Williams (1982) presented a wave tracing method. In his work the ultrasonic input-output for transmitting and receiving transducers coupled to an isotropic elastic plate was studied through examining the multiple reflections of waves in the plate excited by a single tone burst input.

Understanding of the technique is still limited, since many problems related to the technique have unclear solutions. For example, the effects of transducer characteristics, mass and mounting pressure, type of coupling agent used between transducer and material surface, and the interrogation characteristics have not been modeled. And because of this, the stress wave factor used as a measure of efficiency of the stress wave energy transmission in the AU technique does not yet have a standard definition (dos Reis Henrique, 1991). It is traditionally evaluated as the number of oscillations higher than a chosen threshold in the ring down oscillations in the output signal from the receiving transducer (Henneke *et al.*, 1987). Although several investigators have presented alternative parameters for the stress wave factor

<sup>a)</sup>Visiting scholar, the Department of Civil Engineering, The University of Queensland, Australia.

<sup>b)</sup>Corresponding author.

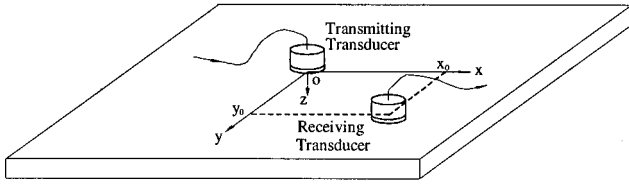


FIG. 1. Schematic of acousto-ultrasonic technique configuration.

such as peak amplitude, event duration, energy and various spectral moments (Lorenzo and Hahn, 1988; Duke *et al.*, 1989; Kautz *et al.*, 1991), there is no analytical model to analyze the effects of the various external factors.

In this paper as an initial attempt to model the above problems, the acousto-ultrasonic characteristics for an isotropic plate are investigated. Unlike Williams' wave tracing method, a spectral analysis approach is presented, which can take into account the transducers' characteristics. Using this approach the input-output frequency spectrum relationships for the cases of one transmitter/one receiver and one transmitter/multiple receivers coupled to an isotropic plate are expressed in explicit forms. Based on this a wave scattering parameter to bridge two different sensing types of the receiving transducers is presented, and a new type of stress wave parameter for AU technique is developed which can eliminate the influence of various external factors.

## I. THEORETICAL CONSIDERATION

A schematic of the problem under consideration is shown in Fig. 1. A transmitting transducer and a receiving transducer are coupled to an isotropic plate. The input pulse in the form of a normal contact stress is first injected on the surface of the plate through the transmitting transducer to excite ultrasonic interrogation pulses. As the pulse wave front arrives at the receiving transducer, the output can be sensed in a form of normal contact stress. Here the electron-mechanic and the mechanic-electron transfer segments are not considered. This does not affect the investigation in this paper at all, since such consideration only requires introduction of two transduction ratios (Williams *et al.*, 1982).

We assume the dominant wavelength of the interrogation pulse excited by the transmitting transducer is larger enough than the thickness of the plate, so the classical plate theory can be used (Medick, 1961; Gorman, 1996). From this theory, the equations governing the motion of the plate can be written as

$$D \left( \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} \right) w(x, y, t) + \rho h \frac{\partial^2 w(x, y, t)}{\partial t^2} = q(x, y, t) \quad (1)$$

with

$$D = \frac{Eh^3}{12(1-\nu^2)}, \quad (2)$$

where  $t$  is the time variable,  $x$  and  $y$  are the space variables referring to the coordinate system with the  $x$  and  $y$  axes lying in the plane of the plate and the  $z$  axis perpendicular to the

plane of the plate, and  $\rho$ ,  $E$ ,  $\nu$  and  $h$  are the mass density, Young's modulus, Poisson's ratio and the thickness of the plate, respectively.  $w(x, y, t)$  denotes the transverse displacement of the plate, and  $q(x, y, t)$  the distributed forces exerted on the surface of the plate by the transmitting and the receiving transducers.

As the fundamental study of the problem considered, the point source/point receiver assumption is introduced in this investigation, that is, the radius of each transducer is assumed to be so small that the normal contact pressures between the transducer and the plate can be considered as a point force. Thus,  $q(x, y, t)$  in Eq. (1) can be expressed as

$$q(x, y, t) = \delta(x) \delta(y) f(t) + \delta(x - x_0) \delta(y - y_0) \sigma(t), \quad (3)$$

where  $\delta(x)$  stands for the Dirac delta function,  $f(t)$  is the normal force exerted on the plate by the transmitting transducer located at the origin, and  $\sigma(t)$  is the normal force sensed by the receiving transducer located at  $x = x_0$ ,  $y = y_0$ , which is an unknown function to be determined.

In view of the motion of the receiving transducer, the following equation is obtained

$$M \frac{d^2 w(x_0, y_0, t)}{dt^2} = -\sigma(t), \quad (4)$$

where  $M$  stands for the mass of the receiving transducer, and the displacement continuity condition between the receiving transducer and the plate is used.

In order for the analysis method in this paper to also be applicable to the composite plate cases, a multi-dimensional integral transform method referring to the rectangular coordinate system is presented to solve the above equations.

The following Fourier transforms for time domain and space domain (Sneddon, 1951) are introduced:

$$\bar{g}(x, y, \omega) = \int_0^\infty g(x, y, t) e^{-i\omega t} dt, \quad (5)$$

$$\bar{\bar{g}}(\kappa_x, \kappa_y, \omega) = \int_{-\infty}^\infty \int_{-\infty}^\infty \bar{g}(x, y, \omega) e^{-i(\kappa_x x + \kappa_y y)} dx dy. \quad (6)$$

Applying the transform (5) and (6) to Eqs. (1), (3), and (4) with respect to the variable  $t$  and  $x, y$ , respectively, the unknown  $\bar{\sigma}(\omega)$  can be finally expressed as follows:

$$\bar{\sigma}(\omega) = \frac{\bar{f}(\omega) I_2}{4\pi^2 / M\omega^2 - I_1}, \quad (7)$$

in which

$$I_1 = \int_{-\infty}^\infty \int_{-\infty}^\infty 1/\Phi(\kappa_x, \kappa_y, \omega) d\kappa_x d\kappa_y, \quad (8)$$

$$I_2 = \int_{-\infty}^\infty \int_{-\infty}^\infty e^{i(\kappa_x x_0 + \kappa_y y_0)} / \Phi(\kappa_x, \kappa_y, \omega) d\kappa_x d\kappa_y, \quad (9)$$

$$\Phi(\kappa_x, \kappa_y, \omega) = D(\kappa_x^2 + \kappa_y^2)^2 - \rho h \omega^2. \quad (10)$$

Using the contour integration technique to evaluate the integrals  $I_1$  and  $I_2$  and setting  $y_0 = 0$  yields

$$\bar{\sigma}(\omega) = -\frac{M\omega\bar{f}(\omega)}{4\pi D\alpha(1+i\gamma)}(T_R+iT_I), \quad (11)$$

in which

$$\gamma = \frac{M\omega}{8D\alpha}, \quad (12)$$

$$T_R = K_0(\sqrt{\alpha\omega}x_0) + \frac{\pi}{2}Y_0(\sqrt{\alpha\omega}x_0), \quad (13)$$

$$T_I = \frac{\pi}{2}J_0(\sqrt{\alpha\omega}x_0), \quad (14)$$

where  $J_0(x)$  and  $Y_0(x)$  stand for the Bessel function of order zero of the first kind and second kind, respectively, and  $K_0(x)$  is the modifying Bessel function of the first kind and order zero (Gradshteyn, 1980).

Although Eq. (11) only corresponds to the case when the receiving transducer is located on the  $x$  axis, obviously, for the isotropic plate case this does not affect the generality of the result obtained.

## II. SOLUTION FOR ONE TRANSMITTER/MULTIPLE RECEIVERS CASE

In this section, the problem of ultrasonic input–output for one transmitting transducer and two receiving transducers coupled to an isotropic plate is further examined.

Without loss of generality of the analysis method, here we only consider the case for two receivers in order to simplify the expressions. We assume a transmitting transducer is located at the origin on the plate, and two receiving transducers are, respectively, at the two points  $(x_1, y_1)$  and  $(x_2, y_2)$  on the plate. Equation (1) is still used to simulate the motion of the plate, but  $q(x, y, t)$  is rewritten in the following form:

$$q(x, y, t) = \delta(x)\delta(y)f(t) + \delta(x-x_1)\delta(y-y_1)\sigma_1(t) + \delta(x-x_2)\delta(y-y_2)\sigma_2(t), \quad (15)$$

where  $\sigma_1(t)$  and  $\sigma_2(t)$  are the normal forces sensed by the receiving transducers.

Like Eq. (4), the governing equations for the two receiving transducers can be expressed as follows:

$$M_1 \frac{d^2w(x_1, y_1, t)}{dt^2} = -\sigma_1(t), \quad (16)$$

$$M_2 \frac{d^2w(x_2, y_2, t)}{dt^2} = -\sigma_2(t), \quad (17)$$

in which  $M_1, M_2$  stand for the masses of the two receivers, respectively, and the displacement continuity conditions between the two receivers and the plate have been used.

Applying the integral transforms (5) and (6) to Eqs. (1) and (15)–(17), after some manipulations, leads to

$$\frac{4\pi^2}{M_1\omega^2}\bar{\sigma}_1(\omega) = I_{10}\bar{f}(\omega) + I_{11}\bar{\sigma}_1(\omega) + I_{12}\bar{\sigma}_2(\omega), \quad (18)$$

$$\frac{4\pi^2}{M_2\omega^2}\bar{\sigma}_2(\omega) = I_{20}\bar{f}(\omega) + I_{21}\bar{\sigma}_1(\omega) + I_{22}\bar{\sigma}_2(\omega), \quad (19)$$

with

$$I_{j0} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{i(\kappa_x x_j + \kappa_y y_j)} / \Phi(\kappa_x, \kappa_y, \omega) d\kappa_x d\kappa_y, \quad (20)$$

$$j = 1, 2,$$

$$I_{jj} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} 1 / \Phi(\kappa_x, \kappa_y, \omega) d\kappa_x d\kappa_y, \quad j = 1, 2, \quad (21)$$

$$I_{21} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{i[(\kappa_x(x_2-x_1) - \kappa_y(y_2-y_1))]} / \Phi(\kappa_x, \kappa_y, \omega) d\kappa_x d\kappa_y, \quad (22)$$

$$I_{12} = I_{21}^*, \quad (23)$$

in which  $\Phi(\kappa_x, \kappa_y, \omega)$  is seen in Eq. (10), and the symbol “\*” denotes the complex conjugate.

Following the same procedure for the evaluation of  $I_1$  and  $I_2$ , the integrals  $I_{11}, I_{22}, I_{10}, I_{20}, I_{21}$  in Eqs. (20)–(22) can all be determined. After that, according to Eqs. (18) and (19) the spectral responses  $\bar{\sigma}_1(\omega)$  and  $\bar{\sigma}_2(\omega)$  of the output of the two receiving transducers can be finally evaluated. Due to the limitations on the length of this paper, the final results are only given for the case  $y_1 = y_2 = 0$ , i.e., the transmitter and the two receivers lying on a straight line, as follows:

$$\bar{\sigma}_1(\omega) = -\frac{\Delta_1}{\Delta} \bar{f}(\omega), \quad (24)$$

$$\bar{\sigma}_2(\omega) = -\frac{\Delta_2}{\Delta} \bar{f}(\omega), \quad (25)$$

in which

$$\Delta = \frac{\pi^4}{\omega^2} \left\{ \frac{16}{M_1 M_2 \omega^2} - \frac{1}{4D^2 \alpha^2} [1 + J_0^2(\sqrt{\alpha\omega}x_{12})] - \frac{1}{4D^2 \alpha^2} \left[ \frac{2}{\pi} K_0(\sqrt{\alpha\omega}x_{12}) + Y_0(\sqrt{\alpha\omega}x_{12}) \right]^2 + i \frac{2}{D\alpha\omega} \left( \frac{1}{M_1} + \frac{1}{M_2} \right) \right\}, \quad (26)$$

$$\Delta_1 = \frac{\pi^2}{D^2 \alpha^2 \omega^2} \left[ \frac{4\pi D\alpha}{M_2 \omega} I_{10}^{(0)} - I_{12}^{(0)} I_{20}^{(0)} + i \frac{\pi}{2} I_{10}^{(0)} \right], \quad (27)$$

$$\Delta_2 = \frac{\pi^2}{D^2 \alpha^2 \omega^2} \left[ \frac{4\pi D\alpha}{M_1 \omega} I_{20}^{(0)} - I_{21}^{(0)} I_{10}^{(0)} + i \frac{\pi}{2} I_{20}^{(0)} \right], \quad (28)$$

with

$$I_{j0}^{(0)} = K_0(\sqrt{\alpha\omega}x_j) + \frac{\pi}{2} Y_0(\sqrt{\alpha\omega}x_j) + i \frac{\pi}{2} J_0(\sqrt{\alpha\omega}x_j), \quad j = 1, 2, \quad (29)$$

$$I_{12}^{(0)} = K_0(\sqrt{\alpha\omega}x_{12}) + \frac{\pi}{2} Y_0(\sqrt{\alpha\omega}x_{12}) - i \frac{\pi}{2} J_0(\sqrt{\alpha\omega}x_{12}), \quad (30)$$

$$I_{21}^{(0)} = K_0(\sqrt{\alpha\omega}x_{12}) + \frac{\pi}{2}Y_0(\sqrt{\alpha\omega}x_{12}) + i\frac{\pi}{2}J_0(\sqrt{\alpha\omega}x_{12}), \quad (31)$$

where  $x_{12} = x_2 - x_1$ , and  $x_2 > x_1$  is required.

### III. A WAVE SCATTERING PARAMETER FOR THE RECEIVER

In order to examine the scattering effect of the receiving transducer on the AU results, we first consider the output of receiver when the scattering effect is omitted. In this case, Eq. (3) is rewritten as follows:

$$q(x, y, t) = \delta(x)\delta(y)f(t). \quad (32)$$

Combining Eq. (32) with Eqs. (1), (2), and (4) and following the same procedure as presented in Sec. II, we obtain the spectral response of the output in the form of the normal contact force between the receiving transducer and the plate as follows:

$$\bar{\sigma}(\omega) = -\frac{M\omega\bar{f}(\omega)}{4\pi D\alpha}(T_R + iT_I). \quad (33)$$

Comparing Eqs. (33) and (11), it is observed that for two equations corresponding to the cases with and without consideration of the scattering effect of receiving transducer, the only difference is in  $\gamma$ . Hence, we can deduce that  $\gamma$  is, in fact, a parameter reflecting the receiver's scattering effect on the output of itself. It is a ratio of the impedance of a lumped mass (the transducer) and a thin plate.

Examining this parameter, two different sensing types of the receiver can be found. First, we assume

$$\gamma = \frac{M\omega}{8D\alpha} \ll 1, \quad (34)$$

i.e., the product of the receiver's mass and frequency is very small with respect to the value of  $8D\alpha$ . Then it is seen that Eq. (11) can be approximately replaced by Eq. (33). That is, the scattering effect of the receiver is so small that it can be neglected. This implies that in this case the output of the receiver is completely a result of its vibration.

Second, we assume  $M$  is so large that

$$\gamma = \frac{M\omega}{8D\alpha} \gg 1 \quad (35)$$

then from Eq. (11), the following result can be obtained:

$$\bar{\sigma}(\omega) \approx \frac{2}{\pi}\bar{f}(\omega)(-T_I + iT_R). \quad (36)$$

It can be found that in this case, the output of the receiver is no longer related to  $M$ , which is similar to the case of a wave being reflected by a fixed boundary.

From the above analysis, it can be concluded that when  $\gamma \ll 1$ , the output of the receiver is, in fact, sensed through a vibration process; and when  $\gamma \gg 1$ , the output of the receiver is sensed through a wave reflection process. This conclusion is of much significance to understanding the AU results.

### IV. A NEW STRESS WAVE PARAMETER FOR THE AU TECHNIQUE

There exist many other external factors influencing the AU measurement results apart from material properties or the internal conditions, such as the transducer's spectral characteristics, mass, aperture, location, coupling agent, and so on. Hence, in order to accurately assess the internal condition of materials, the effects of the irrelevant signal components caused by these external factors must be eliminated or reduced substantially if possible. In previous investigations, several stress wave parameters, or alternatives to the so-called stress wave factor, have been presented. However, since they are all based on the simple measurement configuration, i.e., one transmitter/one receiver, it is difficult to eliminate or determine the influence of the external factors. In view of this, a new type of stress wave parameter is developed based on a measurement configuration of one transmitter/two receivers.

Assume a harmonic electrical voltage is applied to the transmitting transducer. Then the contact forces sensed by the two receiving transducers, or in other words, the stress waves impinging on the two receivers, can be expressed as

$$\sigma_1(t) = A_1 e^{-i(\omega t + \phi_1)}, \quad (37)$$

$$\sigma_2(t) = A_2 e^{-i(\omega t + \phi_2)}, \quad (38)$$

where  $A_i$  and  $\phi_i$  ( $i=1,2$ ) are the amplitudes and the phase angles, respectively. When two mechanic-electron transduction ratios  $T_1(\omega)$  and  $T_2(\omega)$  for the two receivers are introduced (Williams *et al.*, 1982), the two output voltages, denoted by  $v_1(t)$  and  $v_2(t)$ , from the two receivers can be written as follows:

$$v_1(t) = T_1(\omega)A_1 e^{-i(\omega t + \phi_1)}, \quad (39)$$

$$v_2(t) = T_2(\omega)A_2 e^{-i(\omega t + \phi_2)}, \quad (40)$$

where  $\phi_1, \phi_2$  are the phase angles of the two output voltages.

With the above assumptions, the amplitudes of the spectra of the output voltages corresponding to the two contact force spectral responses  $\bar{\sigma}_1(\omega)$  and  $\bar{\sigma}_2(\omega)$  can be evaluated from Eqs. (24) and (25) as follows:

$$|\bar{v}_1(\omega)| = T_1(\omega) \frac{|\Delta_1|}{|\Delta|} |\bar{f}(\omega)|, \quad (41)$$

$$|\bar{v}_2(\omega)| = T_2(\omega) \frac{|\Delta_2|}{|\Delta|} |\bar{f}(\omega)|, \quad (42)$$

where  $||$  denotes the module of a complex number. From Eqs. (41) and (42) evaluating the ratio of  $|\bar{v}_2(\omega)|$  to  $|\bar{v}_1(\omega)|$ , we have

$$\frac{|\bar{v}_2(\omega)|}{|\bar{v}_1(\omega)|} = \frac{T_2(\omega)|\Delta_2|}{T_1(\omega)|\Delta_1|}. \quad (43)$$

Like the case of one transmitter/one receiver, we introduce two receiver scattering parameters as

$$\gamma_1 = \frac{M_1\omega}{8D\alpha}, \quad (44a)$$



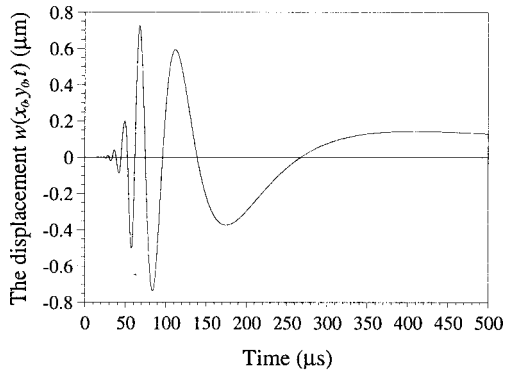


FIG. 2. A displacement response of the plate without receiving transducers ( $x_0=0.1$  m,  $y_0=0$ , and  $T=20$   $\mu$ s).

$$\gamma_2 = \frac{M_2 \omega}{8D\alpha}. \quad (44b)$$

Substituting Eqs. (27) and (28) into Eq. (43), and then applying the expressions (44a) and (44b) to the resultant equation, yields

$$\frac{|\bar{v}_2(\omega)|}{|\bar{v}_1(\omega)|} = \frac{M_2 T_2(\omega) \left| I_{20}^{(0)}(1+i\gamma_1) - \frac{2}{\pi} \gamma_1 I_{21}^{(0)} I_{10}^{(0)} \right|}{M_1 T_1(\omega) \left| I_{10}^{(0)}(1+i\gamma_2) - \frac{2}{\pi} \gamma_2 I_{12}^{(0)} I_{20}^{(0)} \right|}. \quad (45)$$

If the two receiving transducers chosen are the same, then the expression (45) reduced to the following form:

$$\frac{|\bar{v}_2(\omega)|}{|\bar{v}_1(\omega)|} = \frac{\left| I_{20}^{(0)}(1+i\gamma_1) - \frac{2}{\pi} \gamma_1 I_{21}^{(0)} I_{10}^{(0)} \right|}{\left| I_{10}^{(0)}(1+i\gamma_2) - \frac{2}{\pi} \gamma_2 I_{12}^{(0)} I_{20}^{(0)} \right|}, \quad (46)$$

in which  $\gamma_1 = \gamma_2$ .

So far, an analytical expression for a measurable ratio which reflects the wave attenuation properties of the plate has been obtained. Using this expression, the effects of various external factors with respect to the internal condition of the material can be evaluated. It can be found that the parameter eliminates the influence of the interrogation pulses excited by the transducer and, especially, avoids the complex problem of the mechanic-electron transduction properties of the receivers which cannot, in general, be exactly determined theoretically.

In order to evaluate the influence of the receivers' self-scattering and mutual interaction, and the attenuation of waves spread in plane of the plate, first consider a fundamental case for  $\gamma_1 \rightarrow 0$ ,  $\gamma_2 \rightarrow 0$ . From Eq. (46), it can be found that in this case

$$\frac{|\bar{v}_2(\omega)|}{|\bar{v}_1(\omega)|} = \frac{|I_{20}^{(0)}|}{|I_{10}^{(0)}|}. \quad (47)$$

Using Eq. (29) and noticing the asymptotic relations, it can be obtained that when  $\beta x_1$  and  $\beta x_2$  are both large,  $I_{20}^{(0)}$  and  $I_{10}^{(0)}$  in Eq. (47) are of the following form:

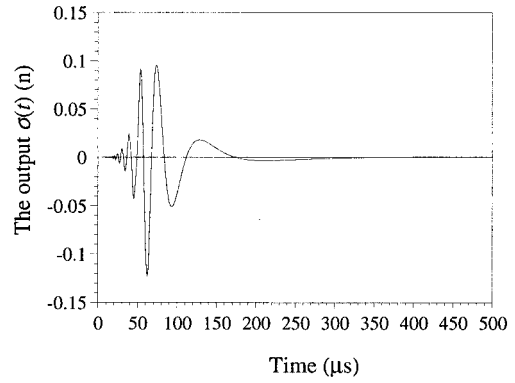


FIG. 3. A time history of the output for  $T=20$   $\mu$ s,  $M=0.1$  kg,  $x_0=0.1$  m, and  $y_0=0$ .

$$|I_{20}^{(0)}| \approx \sqrt{\frac{\pi}{2\beta x_2}}, \quad (48a)$$

$$|I_{10}^{(0)}| \approx \sqrt{\frac{\pi}{2\beta x_1}}. \quad (48b)$$

Substituting Eqs. (48a) and (48b) into Eq. (47), yields

$$\frac{|\bar{v}_2(\omega)|}{|\bar{v}_1(\omega)|} \approx \sqrt{\frac{x_1}{x_2}}. \quad (49)$$

Following the preceding analysis procedure for the scattering parameter  $\gamma$ , it can be found that Eqs. (47) and (49) are, in fact, the results corresponding to the case without consideration of the receivers' self-scattering and mutual interaction influence. It can also be seen that Eq. (49) in fact stands for the influence of the wave spreading attenuation.

With regard to the above features, in order to specify the influence of the receivers' self-scattering and mutual interaction we introduce an influence coefficient as follows:

$$F = \sqrt{\frac{x_2}{x_1} \frac{|I_{20}^{(0)}(1+i\gamma_1) - (2/\pi) \gamma_1 I_{21}^{(0)} I_{10}^{(0)}|}{|I_{10}^{(0)}(1+i\gamma_2) - (2/\pi) \gamma_2 I_{12}^{(0)} I_{20}^{(0)}|}}. \quad (50)$$

Using this coefficient, the influence of receivers self-scattering and mutual interaction can be evaluated to some extent.

A normalized stress wave parameter can be finally constructed as follows

$$\varepsilon(\omega) = \sqrt{\frac{x_2}{x_1}} \cdot \frac{1}{F} \cdot \frac{|\bar{v}_2(\omega)|}{|\bar{v}_1(\omega)|}. \quad (51)$$

The first term on the right hand side of Eq. (51) reflects the influence of the wave spreading attenuation, while the second term reflects the influence of receivers self-scattering and mutual interaction. The third term is a practical measurement value, which is related not only to the internal condition of the material but also to the above external factors. It can be seen that for a perfect plate, this parameter is equal to 1, while for a plate that is damaged or has some internal imperfections between the two receivers will have a parameter value less than 1. Therefore this parameter can be used to evaluate and define the internal condition of materials.

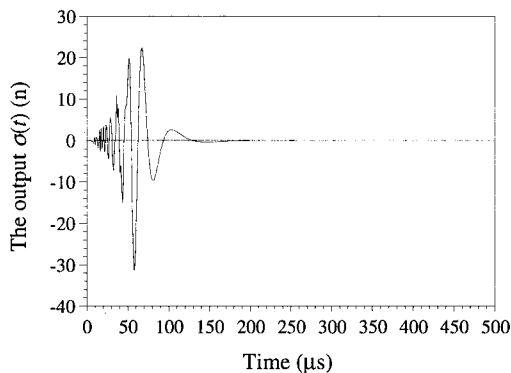


FIG. 4. A time history of the output for  $T=20 \mu\text{s}$ ,  $M=0.1 \text{ kg}$ ,  $x_0=0.1 \text{ m}$ , and  $y_0=0$  calculated without consideration of the receiver's scattering effects.

## V. NUMERICAL RESULTS AND DISCUSSIONS

In this section, a series of numerical results are presented and some discussions concerning the AU technique are detailed.

In all calculations, the applied force  $f(t)$  is assumed to be the delayed sine pulse of time duration, i.e.,

$$f(t) = \begin{cases} \sin(2\pi t/T), & 0 \leq t < T \\ 0, & t \geq T. \end{cases}$$

In the above equation  $T$  is called the pulse width and, unless otherwise noted, it is given a fixed value  $20 \mu\text{s}$ . The parameters for the plate are chosen as:  $\rho=2700 \text{ kg/m}^3$ ,  $E=69 \times 10^9 \text{ N/m}^2$ ,  $\nu=0.33$ , and  $h=0.002 \text{ m}$ .

Figure 2 shows a time history of the output of the receiving transducer for  $M=0.1 \text{ kg}$ ,  $x_0=0.1 \text{ m}$ , and  $y_0=0$ , while the corresponding displacement response of the receiver location point on the plate is shown in Fig. 3. From these two figures it can be seen that the output of the receiver is different from the displacement response of its location point. Two differences are observed: one is the instant corresponding to the peak point, the other is the event duration. This implies that examination of only the surface displacement response of materials due to input pulses is not sufficient to simulate the AU results.

For the purpose of examination of the receiver self-scattering effects on its output, Fig. 4 gives a time history of

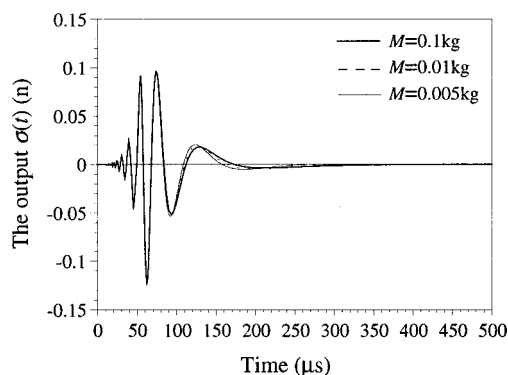


FIG. 5. Influence of receiver mass on its output ( $x_0=0.1 \text{ m}$ ,  $y_0=0$ , and  $T=20 \mu\text{s}$ ).

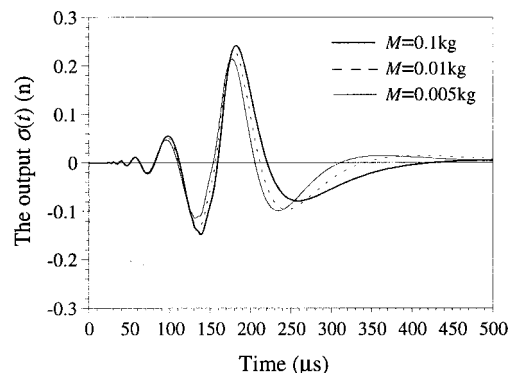


FIG. 6. Influence of receiver mass on its output for  $T=100 \mu\text{s}$  ( $x_0=0.1 \text{ m}$ , and  $y_0=0$ ).

the output of the receiver for  $M=0.1 \text{ kg}$ ,  $x_0=0.1 \text{ m}$ ,  $y_0=0$  calculated according to Eq. (33), i.e., without consideration of the receiver scattering effects. Comparing this figure and Fig. 3 it is observed that the effects on output are very large. Evaluating the scattering parameter  $\gamma$  presented in the above section we can observe that in this case  $\gamma \gg 1$ . Hence it can be concluded that when  $\gamma \gg 1$ , i.e., when the output is sensed mainly through wave reflection process, ignoring the receiver scattering influence causes a large error in the evaluation of the receiver output.

In order to examine the effects of receiving transducer mass on the output, the numerical results of the output of the receiver ( $x_0=0.1 \text{ m}$ ,  $y_0=0$ ) for three different masses ( $M=0.1, 0.01, 0.005 \text{ kg}$ ) are illustrated in Fig. 5. It can be seen that the difference between these three output results is very small. But it is necessary to point out that this difference, to some extent, varies with the width of the input pulse. If the width of the input pulse increases, then this difference becomes large. That means for low-frequency input or less wave scattering effect of a receiver, the output is sensitive to the receiving transducer mass. In Fig. 6, the numerical results for  $T=100 \mu\text{s}$  are given for the above three mass cases.

The investigation of the features of output for  $\gamma \leq 1$  is illustrated in Fig. 7, in which the time history of the output for  $x_0=0.1 \text{ m}$ ,  $y_0=0$ ,  $M=0.1 \text{ kg}$ ,  $T=200 \mu\text{s}$  is calculated, respectively, according to Eqs. (11) and (33), i.e., with and

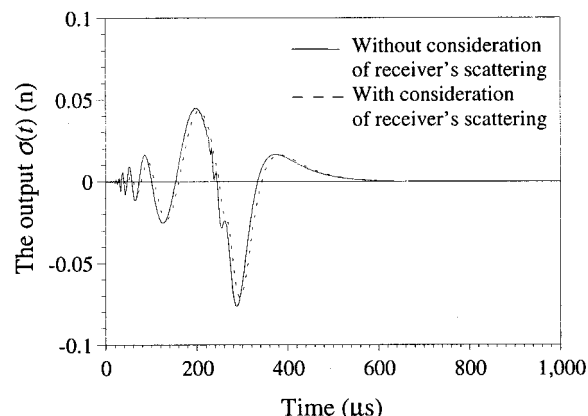


FIG. 7. The receiver scattering effects on the output for  $\gamma \leq 1$  ( $T=200 \mu\text{s}$ ,  $M=0.001 \text{ kg}$ ,  $x_0=0.1 \text{ m}$ , and  $y_0=0$ ).

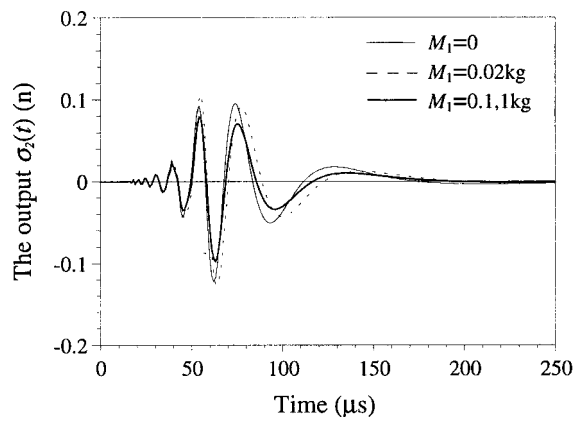


FIG. 8. The influence of the inner receiver on the output of the outer receiver for  $x_1=0.05$  m and  $x_2=0.1$  m ( $T=20$   $\mu$ s and  $M_2=0.1$  kg).

without consideration of the receiver scattering influence. Comparing the two results, it can be seen that the receiver scattering influence on the output of itself can be neglected.

The problem of interaction between two receiving transducers for the measurement configuration of one transmitter/two receivers is examined in Figs. 8 and 9. The influences of the inner receiver, which is near to the transmitter, on the output of the outer receiver are shown in Fig. 8, where the corresponding results for the single outer receiver case are also given and denoted by  $M_1=0$ . From this figure it is evident that: (1) adding an inner receiver makes the output of the outer receiver decrease and (2) with increasing mass of the inner receiver, the output of the outer receiver decreases, but after the mass reaches a certain value, further increases in mass do not cause further reduction of the output of the outer receiver. The influence of the outer receiver on the output of the inner one is illustrated in Fig. 9. Compared with the influence of the inner on the outer one, the effect of the outer on the inner one is insignificant.

Finally, some numerical results are given in Fig. 10 for the coefficient  $F$  [see Eq. (50)] which is used to reflect the influences of receiver self-scattering and mutual interaction on the ratio of the two spectrum amplitude of the output voltages measured from the two receiving transducers. In the calculations we chose  $M_1=M_2=0.1$  kg. Using this figure we can readily determine the influence coefficient required for

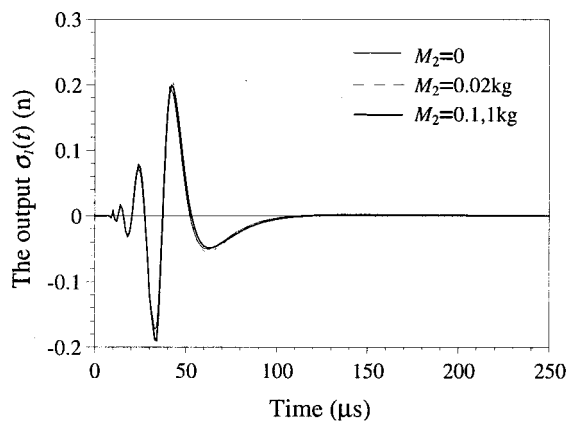


FIG. 9. The influence of the outer receiver on the output of the inner receiver for  $x_1=0.05$  m and  $x_2=0.1$  m ( $T=20$   $\mu$ s and  $M_1=0.1$  kg).

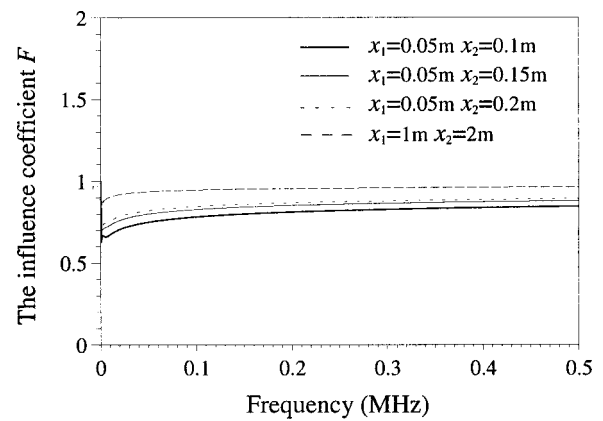


FIG. 10. The influence coefficient  $F$  in the expression of the stress wave parameter presented in this paper ( $M_1=M_2=0.1$  kg and  $T=20$   $\mu$ s).

further determining the stress wave parameter when the AU technique of one transmitter/two receiver is used. In addition, we can easily specify the influences of the receivers self-scattering and mutual interaction. It is evident that the closer the curves are to the line  $F=1$ , the smaller the influences will be.

## VI. CONCLUDING REMARKS

The acousto-ultrasonic characteristics for isotropic thin plates have been investigated in this paper. The following conclusions are drawn:

(1) Unlike the wave tracing method, a spectral analysis approach is presented, which can take into account the influences of the receiving transducer's characteristics. Using this approach and the point source/point receiver assumption, the input-output frequency spectrum relationships for the cases of one transmitter/one receiver and one transmitter/multiple receivers coupled to an isotropic thin plate are expressed in explicit forms. Based on this, using the fast Fourier transform (FFT) technique, a series of numerical results in time domain are presented for the discussion of various influence on the AU technique. Since the analysis approach is based on a general multiple-dimensional integral transform method instead of Green's function method, it can be extended to complex plate cases and transducer cases.

(2) For the AU technique with contact-type transducers, examination of only the displacement response is insufficient to simulate its output. The transducers' own characteristics must be taken into account.

(3) There exists a wave scattering parameter for contact-type receiving transducers which can bridge two important sensing types: one is a vibration sensing in which the transducer scattering effect can be neglected, and the other is a wave-reflecting sensing in which the transducer scattering effect is dominant. The former corresponds to the general vibration measurement case, while the latter applies to the ultrasonic measurement case as long as the mass of the receiver is not very small and the dominant frequency of the interrogation pulse is not too low.

(4) Extending the traditional one transmitter/one receiver configuration of the AU technique to the one transmitter/two receiver case, a new type of stress wave pa-

parameter has been developed which avoids the complex problem of evaluation of the mechanic-electron transduction characteristics of receivers, and also to some extent can be used to specify the influences of receiver self-scattering and mutual interaction, and the natural attenuation of wave spreading in the plate plane.

- Datta, S. K., Bratton, R. L., Chakraborty, T., and Shah, A. H. (1988). "Wave propagation in laminated composite plates," *J. Acoust. Soc. Am.* **83**, 2020–2026.
- Datta, S. K., Ju, T. H., and Shah, A. H. (1992). "Scattering of an impact wave by a crack in a composite plate," *J. Appl. Mech.* **59**, 596–603.
- dos Reis Henrique, L. M. (1991). "Acousto-ultrasonic evaluation of ceramic matrix composite materials," NASA CR 187073.
- Duke, J. C., Henneke, E. G., Kiernan, M. T., and Grosskopf, P. P. (1989). "A study of the stress wave factor technique for evaluation of composite materials," NASA CR 4195.
- Gorman, M. R., and Prosser, W. H. (1996). "Application of normal mode expansion to acoustic emission waves in finite plates," *J. Appl. Mech.* **63**, 555–557.
- Gradshteyn, I. S., and Ryzhik, I. M. (1980). *Table of Integrals, Series, and Products* (Academic, New York).
- Henneke, E. G., Duke, J. C., and Stiffler, R. C. (1987). "Characterizing the damage state of composite laminates via the acousto-ultrasonic technique," in *Solid Mechanics Research for Quantitative Non-Destructive Evaluation* (Martinus Nijhoff, Dordrecht, Netherlands), pp. 217–235.
- Kautz, H. E., and Lerch, B. A. (1991). "Preliminary investigation of acousto-ultrasonic evaluation of metal-matrix composite specimens," *Mater. Eval.* **49**, 607–612.
- Lorenzo, L., and Hahn, H. T. (1988). "Damage assessment by acousto-ultrasonic technique in composites," *Composite Materials: Testing and Design, ASTM STP 972* (ASTM, Philadelphia), pp. 380–397.
- Medick, M. A. (1961). "On classical plate theory and wave propagation," *ASME J. Appl. Mech.* **28**, 223–228.
- Pelts, S. P., and Rose, J. L. (1996). "Source influence parameters on elastic guided waves in an orthotropic plate," *J. Acoust. Soc. Am.* **99**, 2124–2129.
- Rose, J. L., Ditri, J. J., and Aleksander, P. (1994). "Wave mechanics in acousto-ultrasonic nondestructive evaluation," *J. Acoust. Emiss.* **12**, 23–26.
- Sneddon, *Fourier Transform* (McGraw-Hill, New York, 1951).
- Vary, A. (1991). "Acousto-ultrasonics: retrospective exhortation with bibliography," *Mater. Eval.* **49**, 581–591.
- Veidt, M., and Sayir, M. B. (1990). "Experimental evaluation of global composite laminate stiffnesses by structural wave propagation," *J. Compos. Mater.* **24**, 688–706.
- Veidt, M., and Sachse, W. (1994). "Ultrasonic point source/point receiver measurements in thin specimens," *J. Acoust. Soc. Am.* **96**, 2318–2326.
- Williams, J. H., Karagulle, J. H., and Lee, S. S. (1982). "Ultrasonic input-output for transmitting and receiving longitudinal transducers coupled to same face of isotropic elastic plate," *Mater. Eval.* **40**, 655–662.

# Efficiency of a noise barrier with an acoustically soft cylindrical edge for practical use

Tomonao Okubo<sup>a)</sup> and Kyoji Fujiwara

*Department of Acoustic Design, Kyushu Institute of Design, 4-9-1 Shiobaru, Minami-ku, Fukuoka 815-8540, Japan*

(Received 20 October 1998; accepted for publication 7 March 1999)

This paper examines the sound shielding efficiency of a noise barrier having an acoustically soft cylindrical edge, as it compares to that provided by the well-known absorptive cylindrical edge. It has been reported that the waterwheel-shaped cylinder (acoustic tubes in a radial arrangement) approximates a soft surface cylinder, and that the sound shielding efficiency of a noise barrier is improved by placing the cylinder on its edge. The efficiency of the waterwheel-shaped edge barrier is strongly frequency dependent, thus the improvement in overall sound pressure level is smaller than expected when the source is broadband noise. The present study investigates the use of varied tube depths to improve the efficiency of the waterwheel-shaped edge. It is shown that adding tubes of different depths can flatten the frequency dependence, and that such tubes are only needed in the upper half of the cylinder. These findings led to the design of a new edge device for controlling road traffic noise, whose numerical simulations suggest that it is twice as effective in overall sound pressure level as the original waterwheel with uniform-depth channels. © 1999 Acoustical Society of America. [S0001-4966(99)03806-0]

PACS numbers: 43.50.Gf, 43.20.Fn [MRS]

## INTRODUCTION

Noise barriers are the most common solution for controlling traffic noise from expressways, and several methods have been developed for improving their efficiency without increasing their height. It is well known that the sound pressure measured at the barrier edge can be regarded as an imaginary line source for a diffracted field in the back of the barrier. Thus, suppression of the sound pressure at the edge reduces this imaginary source and decreases the diffracted field behind the barrier. Installing an absorptive obstacle on the barrier's edge has also been proposed, based on the same concept. According to one report,<sup>1</sup> the installed absorber reduced the sound pressure around the edge and improved the sound shielding efficiency of the noise barrier. Installation of an absorber around the barrier edge to control expressway noise is currently in use.

If the sound pressure at the edge is considered as an imaginary source, then additional suppression of the sound pressure should further improve sound shielding efficiency. As such, installing an obstacle with an "acoustically soft surface"<sup>2</sup> on the barrier's edge should prove effective. The characteristic impedance of a "soft surface" is sufficiently less than that of air; consequently, the soft-surface sound pressure is much less than that of an absorptive surface. It has been theorized that barriers with soft cylindrical edges<sup>3</sup> and those that are soft only around the edge<sup>4</sup> are both efficient. Despite their significant effect, these barriers are not widely used because of the difficulty in obtaining materials whose impedance is significantly less than that of air. Therefore, the authors previously tried to develop a soft-surface cylinder consisting of open ends of tubes arranged radially,<sup>5,6</sup>

as shown in Fig. 1. Hereafter, this design is referred to as a "waterwheel-shaped cylinder" because of its cross-sectional shape.

It has been shown that the waterwheel-shaped cylinder approximates a soft surface and that installing the waterwheel on the edge of a noise barrier improves its sound shielding efficiency.<sup>5</sup> Due to interference inside the tubes, however, the efficiency of a noise barrier with a waterwheel-shaped edge is unavoidably frequency dependent.<sup>6</sup> Although the improvement in the efficiency of the barrier is remarkable for those frequencies in which the waterwheel can develop a soft surface, its improvement is zero, or even negative, outside these effective frequencies; therefore, the efficiency of the waterwheel-shaped edge against a broadband noise, such as road traffic, is unexpectedly small. This paper investigates the relationship between the cross-section design of the waterwheel-shaped edge and the sound shielding efficiency of the noise barrier against broadband noise, and presents the design of a new edge device for efficiently reducing road traffic noise.

## I. NUMERICAL ANALYSIS

### A. Boundary element method

In this section, the boundary integral equation and the boundary element method (BEM) are described briefly. Details of the theory and accuracy analyses are shown in a previous report.<sup>6</sup> A two-dimensional sound field is assumed throughout this paper because the three-dimensional simulations require many more computational resources and calculation time. It is already shown by scale-model experiments that the two-dimensional analyses can be applied to a practical situation; that is, the three-dimensional situation that a point source and a receiver are in the vertical plane which is perpendicular to the barrier.<sup>6</sup>

<sup>a)</sup>Current address: Kobayasi Institute of Physical Research, 3-20-41 Higashi-motomachi, Kokubunji, Tokyo 185-0022, Japan. Electronic mail: okubo@kobayasi-riken.or.jp

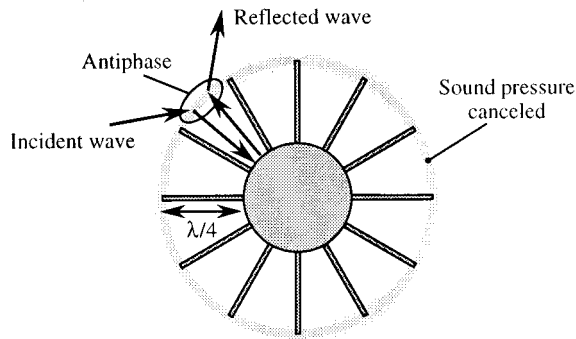


FIG. 1. Development of a soft cylinder with a waterwheel-shaped cylinder.

Figure 2 shows the cross section of the situation for which the sound field was calculated. A noise barrier of infinite length lies on a rigid plane, i.e., on a reflective ground. The barrier and a monofrequency line source of sound are parallel to the  $z$  axis. Let  $\mathbf{r}_0$  denote the source position,  $\mathbf{r}$  denote the receiver position, and  $\beta(\mathbf{r}_s)$  denote the normalized surface admittance at point  $\mathbf{r}_s$  on the barrier surface  $S$ . The sound pressure at the receiver,  $p(\mathbf{r}, \mathbf{r}_0)$ , satisfies the following boundary integral equation:<sup>7</sup>

$$\varepsilon(\mathbf{r})p(\mathbf{r}, \mathbf{r}_0) = G(\mathbf{r}_0, \mathbf{r}) - \int_S \left( \frac{\partial G(\mathbf{r}_s, \mathbf{r})}{\partial n(\mathbf{r}_s)} + jk\beta(\mathbf{r}_s)G(\mathbf{r}_s, \mathbf{r}) \right) \times p(\mathbf{r}_s, \mathbf{r}_0) ds(\mathbf{r}_s), \quad (1)$$

where  $ds(\mathbf{r}_s)$  denotes the arc-length of the barrier surface  $S$  at point  $\mathbf{r}_s$ ,  $\partial/\partial n(\mathbf{r}_s)$  denotes the normal derivative at  $\mathbf{r}_s$ ,  $k$  denotes the wave number, and the time dependence factor  $\exp(j\omega t)$  is understood.  $\varepsilon(\mathbf{r})=1$  when  $\mathbf{r}$  is in the propagating medium and not on  $S$ ;  $\varepsilon(\mathbf{r})=\frac{1}{2}$  when  $\mathbf{r}$  is on  $S$ .  $G(\mathbf{r}, \mathbf{r}_0)$  is the sound pressure at  $\mathbf{r}$  in the absence of the barrier:

$$G(\mathbf{r}, \mathbf{r}_0) = \frac{1}{4j} \{ H_0^{(2)}(k|\mathbf{r}_0 - \mathbf{r}|) + H_0^{(2)}(k|\mathbf{r}'_0 - \mathbf{r}|) \}, \quad (2)$$

where  $\mathbf{r}'_0$  denotes the position of the imaginary line source in the rigid ground, and  $H_0^{(2)}$  is the Hankel function of the second kind of order zero. In order to solve Eq. (1) numerically, the BEM was carried out. The barrier surface  $S$  was divided into straight elements, then Eq. (1) is written in a discrete form by the approximation that sound pressure on an element surface is constant and equal to that at the midpoint of the element. Then by setting  $\mathbf{r}$  to the midpoints of the elements, a set of linear equations is obtained in the unknowns of the surface sound pressure. When the equations are solved, the sound pressure at any point  $\mathbf{r}$  can be calculated by substituting the pressure at the midpoint of each element into the

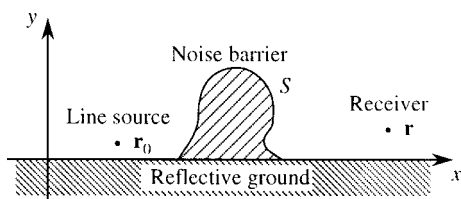


FIG. 2. Cross section of the two-dimensional sound field investigated by the boundary element method.

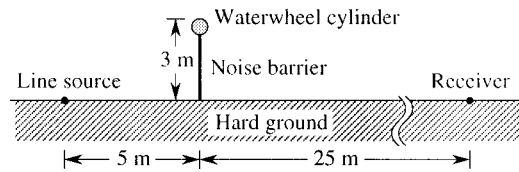


FIG. 3. Configuration of the two-dimensional sound field simulated by the boundary element method.

discrete equation. To solve the integral equation with sufficient accuracy, a maximum element length is smaller than  $\lambda/8$  (Ref. 8) in all of the calculations carried out in this paper.

It is well known that the BEM analysis has numerical difficulty at some frequencies: the linear equations have more than one solution at frequencies which are close to the eigenfrequencies of the interior boundary value problem. If the frequency where the BEM calculation is carried out happens to be very close to one of the eigenfrequencies, the result of the BEM is unreliable. That is, when a calculated result at a certain frequency protrudes significantly from the frequency characteristic curve of the sound pressure level (SPL), it might be caused by the numerical difficulty, especially in higher frequency range. Although different kinds of methods to improve this numerical difficulty around the eigenfrequencies have been proposed (most of them are based on the CHIEF method by Schenck<sup>9</sup> or the linear combined integral equation method by Burton and Miller<sup>10</sup>), the improvement of the difficulty was not adopted in this paper. This is because some of the methods are computationally expensive to implement, and others are quite simple but do not always improve successfully. Instead, if calculated results protrude extremely from the SPL curve and they are definitely not related to the essential characteristics of the sound field, the calculated peak or dip at the frequency was modified to flatten the SPL curve. The protrusions were never removed when authors could not distinguish the numerical problem and the acoustical properties of the waterwheel-shaped cylinder which is strongly frequency dependent.

## B. Configuration of the sound field and the spectrum of traffic noise

Figure 3 shows the cross section of the sound field to be calculated numerically. It is assumed that the cross-sectional shape of a noise barrier does not vary along its length. The positions of a line source, noise barrier, and receiver were determined to simulate the sound field around a noise barrier built along road traffic. The relationship between their positions and the efficiency of the edge device is investigated later (in Figs. 20 and 21). A rigid noise barrier of infinite length, 3-m height, and 0.03-m thickness was positioned on the reflecting ground. The 0.03-m thickness is unrealistically small, as 0.095-m-thick barriers are typically used in Japan. The 0.03-m dimension was used, however, to avoid the numerical difficulty described in the previous section. A coherent monofrequency line source was placed at a horizontal distance of 5 m in front of the barrier, and a receiver was placed 25 m behind the barrier. The line source and the re-

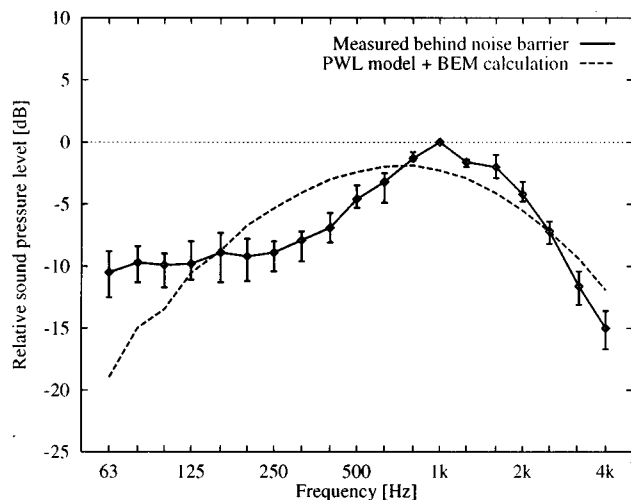


FIG. 4. Road traffic noise spectrum observed at the receiver behind the barrier shown in Fig. 3.

ceiver were placed on the ground surface to avoid complications due to interference caused by reflections from ground surface in the monofrequency sound field. Were the interference to minimize the sound pressure at a receiver above the ground, it would be impossible to distinguish the sound shielding efficiency of the barrier from the effect of the interference.

The shape of the spectrum of the noise targeted for suppression by barriers appears considerably significant when designing a frequency-dependent noise-control device, such as the waterwheel-shaped edge. In general, the traffic noise behind a noise barrier can be predicted by subtracting the numerically calculated efficiency of the barrier from the power level spectrum of the noise source. However, in a realistic sound field, the predicted spectrum sometimes differs from the measured spectrum, probably because properties of the field, such as the effect of the ground surface, are not dealt with successfully in the numerical simulations. Therefore, the traffic noise spectrum measurements were taken behind noise barriers that had been built along expressways, and were used in this paper as the target of barrier noise control. Measurements of  $\frac{1}{3}$ -octave-band SPLs were taken at five points along the Kyushu expressway, which is in southwestern Japan. A microphone was placed on the ground at a distance of 25 m behind a 3-m-high noise barrier. The ground surface was asphalt at four of the observation points and grass at the fifth. Two spectra of five differ from the others obviously: at a point on the asphalt ground SPL around 500 Hz is emphasized by the sound of engine braking from large vehicles running down the long slope, and at the point on the grass SPL in higher frequency range decreases because of the ground absorption. Thus the two spectra were not considered and three other spectra were averaged to represent the traffic noise. The solid line in Fig. 4 shows averaged results of the three spectra as a  $\frac{1}{3}$ -octave-band SPL relative to the noise level at 1 kHz, and the variation of the three spectra is indicated by error bars. The broken line shows the spectrum simulated by the combination of a measured power level spectrum of vehicles<sup>11</sup> and the BEM calculation; the

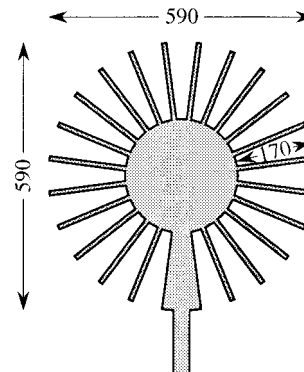


FIG. 5. Waterwheel cylinder installed on the edge of a barrier.

curve is adjusted to equal its overall SPL to that of the measured spectrum. It is shown that the peak of the measured spectrum is narrower as compared with the simulated spectrum. In the following sections, it is assumed that the traffic noise spectrum of solid line in Fig. 4 was observed at a receiver that was 25 m behind the 3-m-high barrier, as shown in Fig. 3.

## II. RELATIONSHIP BETWEEN THE PROFILE OF THE WATERWHEEL-SHAPED EDGE AND ITS SOUND SHIELDING EFFICIENCY

The frequency characteristics of the SPL behind barriers with and without the waterwheel-shaped edge were calculated using the BEM. The complicated cross-sectional shape of the barrier with the waterwheel was divided into many straight-line elements, and the admittance is zero on the surface of all elements. Both of the barriers with and without the waterwheel were 3 m high in order to exclude the effect of extra barrier height from the effect of the waterwheel-shaped edge. The waterwheel increases the thickness of the edge, even as the height remains constant; thus, the geometrical boundary of diffraction moves upward and increases the so-called "effective height." This increase in effective height, however, is sufficiently small that its efficiency changes only slightly with the increase in height.

The SPLs behind the barriers were calculated using the BEM at  $\frac{1}{15}$ -octave frequency intervals. The five results around a  $\frac{1}{3}$ -octave-band center frequency were averaged energetically to approximate  $\frac{1}{3}$ -octave-band SPL. The SPL difference between barriers with and without the waterwheel-shaped edge is referred to as the effect of the waterwheel. The traffic-noise spectrum observed in back of the waterwheel-equipped barrier was predicted by subtracting this waterwheel effect from the representative spectrum of road traffic noise shown in Fig. 4.

### A. Waterwheel diameter and channel depth

A waterwheel cylinder with a diameter of 590 mm and a channel depth of 170 mm, evaluated in previous reports, was installed on the edge of a 3-m-high barrier as shown in Fig. 5. The opening angle of the channels is 15 degrees (the circumference is divided into 24 arcs). Figure 6 shows the analysis based on a 70 dB SPL at a 1 kHz band before

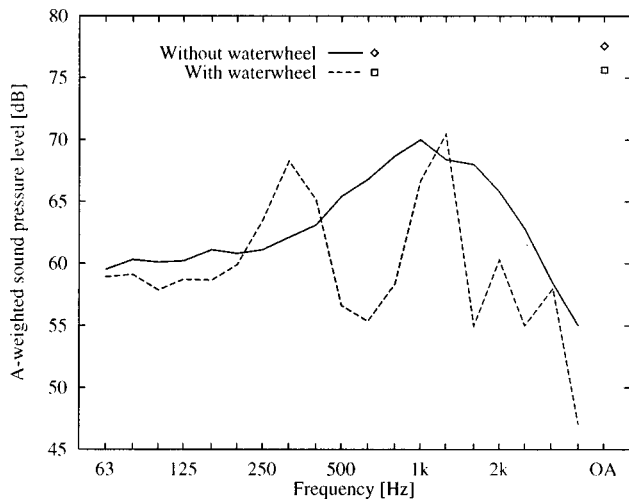


FIG. 6. Effect of the waterwheel-shaped edge; 590-mm diameter and 170-mm channel depth.

installation of the waterwheel-shaped edge. The two points on the right-hand side of Fig. 6 indicate the overall SPL behind the barrier with and without the cylinder. The overall SPL was reduced by only 1.7 dB, whereas the waterwheel brings noticeably large effect, more than 10 dB, at 630 Hz, mainly because those peaks remaining at 315 Hz and 1.2 kHz affect the overall SPL. In particular, because the peak at 1.2 kHz prevents the reduction of the original peak of the observed spectrum without the cylinder, it prevents the reduction of the overall SPL. Moreover, the sound shielding efficiency of the barrier decreases (i.e., its sound pressure level increases) remarkably in the broad frequency range around 315 Hz. In the following discussion, the decrease in efficiency is referred to as the “negative effect” of the waterwheel-shaped edge.

Changing the depth of the channels while keeping their diameter at 590 mm, the waterwheel effect changes as shown in Fig. 7. A waterwheel with a channel depth of 110 mm (a quarter of the wavelength at 770 Hz) shows that the effect shifts toward a higher frequency range, as compared to the results from a depth of 170 mm. Furthermore, this effect

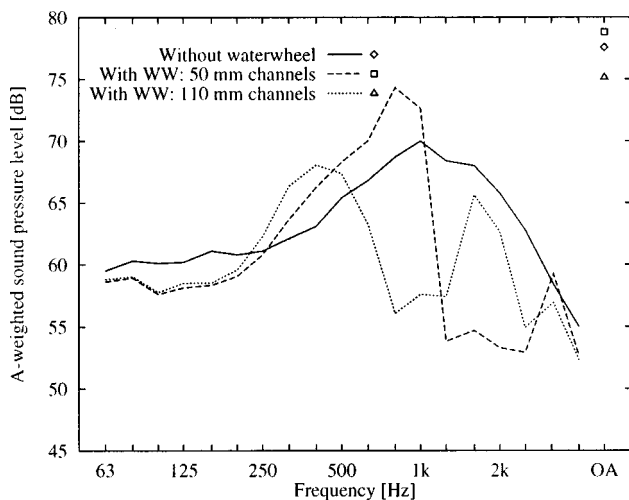


FIG. 7. Effect of the waterwheel-shaped edge; 590-mm diameter and different channel depths.

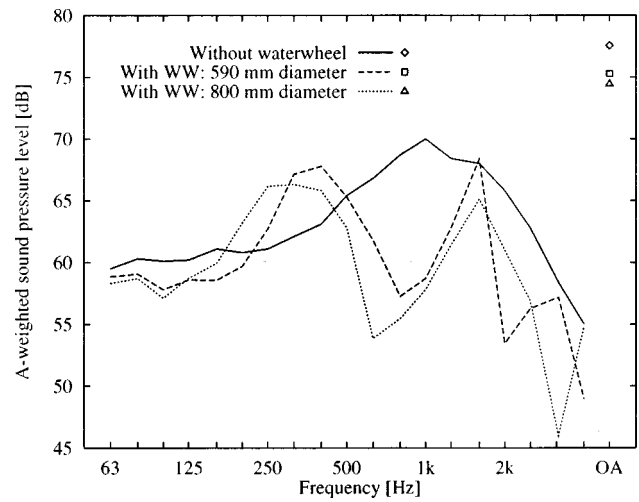


FIG. 8. Effect of the waterwheel-shaped edge; different diameters and 130-mm channel depth.

shifts to a much higher range when the depth is changed to 50 mm (a quarter of the wavelength at 1.7 kHz). The “negative-effect” range is widened when the depth is shortened, while the lower limit of the range remains almost the same, around 200 Hz. The relationship between channel depth and improvement in overall SPL reveals that 110-mm channels reduce the overall SPL more than do 170-mm channels, because the effective range of the waterwheel matches the peak of the traffic noise spectrum. In the case of 50-mm channels, the peak of the traffic noise corresponds to the negative effect of the waterwheel; installing the waterwheel unfortunately increases the noise at 800 Hz and 1 kHz which strongly affects the overall SPL. Consequently, it is the waterwheel-shaped edge with 50-mm channels that decreases the sound shielding efficiency of a barrier. The change of diameter affects the waterwheel effect, as shown in Fig. 8. The effects of waterwheel-shaped edges with a channel depth of 130 mm (a quarter of the wavelength at 650 Hz) and diameters of 590 and 800 mm were compared. The result indicates that increasing the diameter while keeping the channel depth constant shifts the lower limit of the improvement range downward, even while the upper limit remains at the same frequency. Hence, the improvement range becomes wider and the ineffective range shifts to a lower-frequency range.

Figure 9 shows the improvement by the waterwheel-shaped edge in overall SPL as a function of channel depth. Calculations were carried out based on two waterwheel-shaped edges, with diameters of 590 and 800 mm, and variable channel depth at 10-mm intervals. The channel depths that maximize overall improvement are 140 and 130 mm; however, the maxima of improvement are not large: 2.3 and 3.1 dB, respectively. The overall effect of the waterwheel-shaped edge against broadband noises, such as traffic noise from expressways, is relatively small because of the frequency characteristics of the waterwheel-shaped effect, that is, its negative effect and the narrowness of the effective range. The following discussion of the relationship between the negative effect and the cross-sectional shape of the waterwheel, its diameter and channel depth, is based on the 29



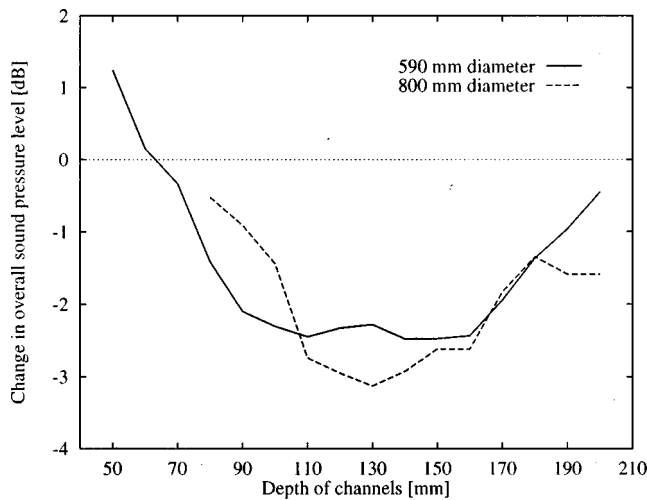


FIG. 9. Change in overall sound pressure level as a function of the depth of the channels.

spectra calculated to draw Fig. 9, but the spectra were not included due to page considerations. (Please refer to Figs. 7 and 8.) Diameter determines the lower limit frequency of the ineffective range, and the larger-diameter waterwheel has a negative effect in the lower frequency range. The lower limit is hardly affected by the channel depth; rather, the upper limit of the ineffective range (i.e., the lower limit of the effective range) shifts with varying channel depth. In other words, the waterwheel-shaped edge with a large diameter and shallow channels decreases the sound shielding efficiency of the barrier across a broad frequency range.

### B. Combining channels of various depths

As described in the previous section, installing a waterwheel with uniform-depth channels does not obviously decrease the overall SPL, because the peaks of the noise spectrum remaining on both sides of the effective range prevent a decrease in the overall level, even though the sound shielding efficiency is greatly improved in the effective frequency range. In this section, channels of various depths are combined with channels of uniform depth to improve the overall efficiency of the waterwheel-shaped edge. As shown in Fig. 8, when the waterwheel with 800-mm diameter and 130-mm channels is installed, two peaks remain in the noise spectrum: one between 315 Hz and 400 Hz, and the other at 1.6 kHz. In the following, channels with depths corresponding to the remaining peaks are added to reduce the sound pressure of the peaks.

First, the channel depth necessary to reduce the sound pressure around 400 Hz is investigated. It had already been indicated that sound pressure at the open end of the channel is minimized at a little higher frequency than the frequency at which the channel depth corresponds to a quarter of the wavelength when the open end is larger than the bottom in the cross-sectional shape of the channel.<sup>5</sup> Therefore, when the depth of the newly added channel is 260 mm corresponding to a quarter of the wavelength at 330 Hz, this additional channel is expected to minimize sound pressure at a higher frequency, around 400 Hz. Figure 10 shows the additional

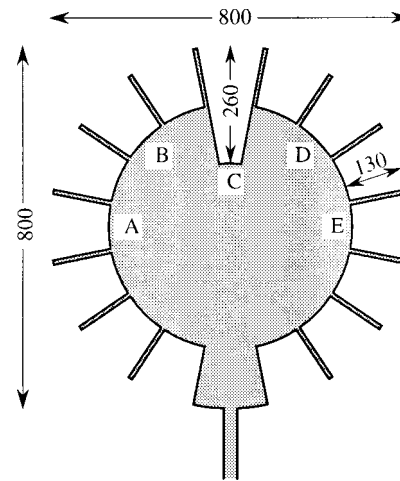


FIG. 10. An additional channel of 260-mm depth to reduce the negative effect of the waterwheel.

channel of 260 mm is placed on the top of the original waterwheel-equipped barrier. The opening angle of the channels is 22.5 degrees (circumference divided into 16 arcs). Figure 11 shows the change in the sound shielding efficiency caused by the addition of the 260-mm channel. At 400 Hz, the SPL decreases 5.2 dB, even though the overall SPL still is not improved because of the remaining two peaks. Between 630 Hz and 1.2 kHz, the improvement obtained by the original waterwheel-shaped edge with a uniform channel depth is decreased, probably because the addition of the 260-mm channel decreases the number of open ends of 130-mm channels in the cylinder surface.

Next, the relationship between the position of the additional channel and its effect is investigated. The position of the 260-mm channel is changed to one of five positions (A–E, Fig. 10), and the changes in sound shielding efficiency are calculated. Only one channel is 260 mm, and the other channels are 130 mm. Figure 12 shows the result of (a) effective positions and (b) ineffective positions. Placing the 260-mm channel at positions B, C, or D, which are close to the top of the barrier, reduces the SPL peak at 400 Hz; at

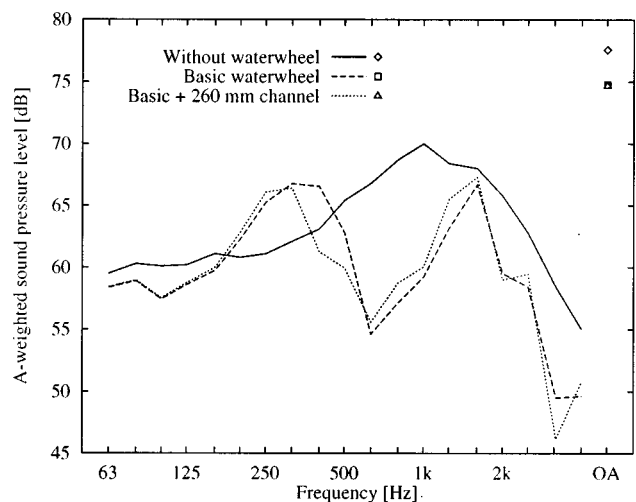
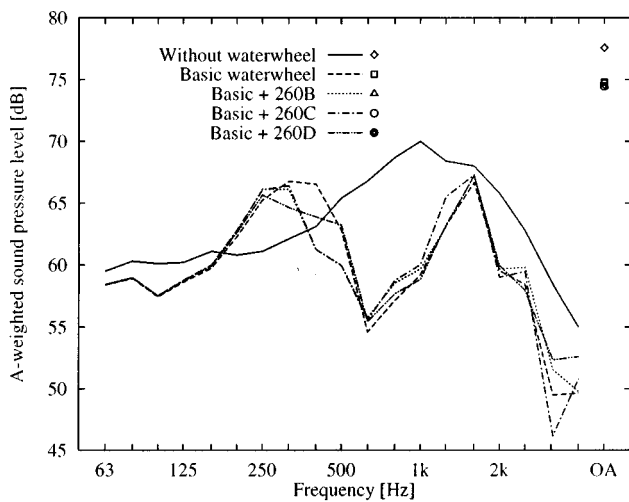
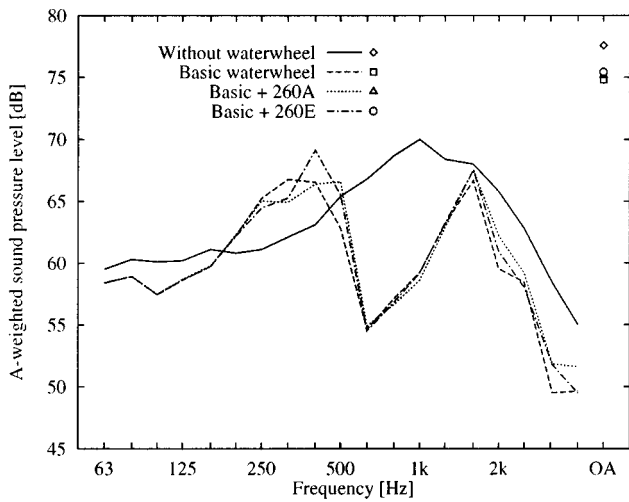


FIG. 11. Improvement in the waterwheel's negative effect by the additional channel of 260 mm.



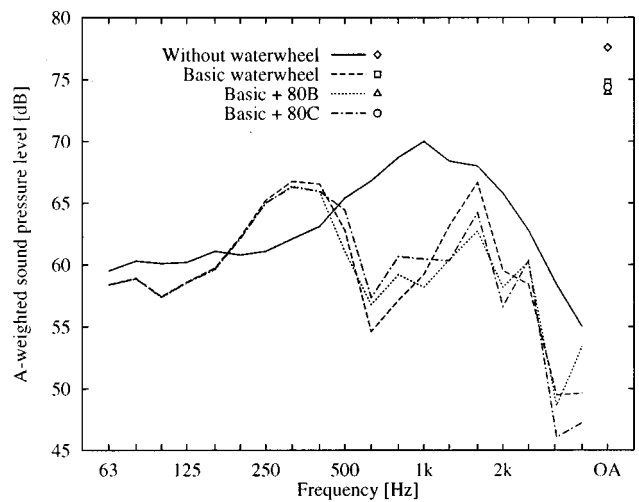
(a) successfully improved



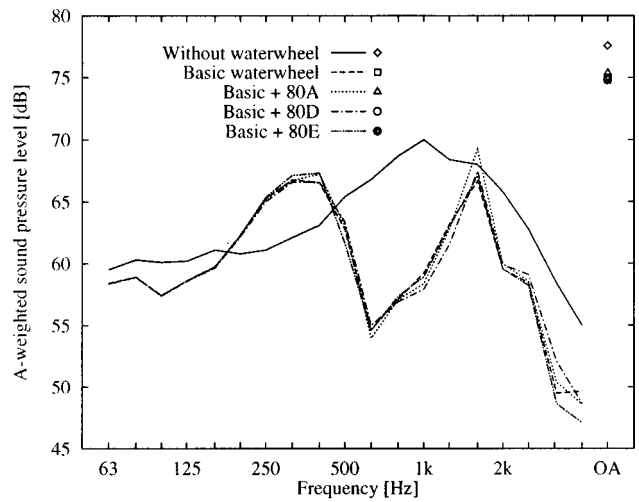
(b) not improved

FIG. 12. Effect of the position of the additional 260-mm channel.

position B or C by more than 5 dB, and at position D by 3 dB. However, placing the 260-mm channel at position A or E does not cause intended improvements at 400 Hz. Similar investigations were carried out on the remaining 1.6-kHz peak. A channel at one of the positions of A–E was changed to 80 mm and the efficiency of the barrier calculated. It has been reported that a channel does not work successfully when the open end is much larger than the wavelength (i.e., the opening angle of the channel is too wide) because the sound wave propagation is not only in the direction of the channel depth,<sup>6</sup> thus, for the 80-mm channel to work successfully, two 80-mm channels with opening angles of 11.25 degrees (half the opening angle of the other channels) are assigned to one of five positions. Figure 13 shows that placing the 80 mm channels at position B or C successfully reduces the peak at 1.6 kHz, and that placing the channels at other positions does not affect the noise spectrum. Also, when 80-mm channels are located at position B or C, the SPL is increased in the 630-Hz to 1-kHz frequency range, where a large improvement was obtained originally. The addition of 80-mm channels increases the SPL in the range between 630 Hz and 1 kHz more than does the addition of a



(a) successfully improved



(b) not improved

FIG. 13. Effect of the position of the additional 80-mm channel.

260-mm channel. This reduction of improvement for the original 130-mm channels is probably due to the reduction in the number of 130-mm channels and to the negative effect of the 80-mm channels. In summary, there are effective and ineffective positions for channels with different depth to be assigned and the sound shielding efficiency is not affected so long as the additional channel is not close to the top of the barrier. Moreover, 80-mm channels on the position D do not affect the efficiency at all while a 260-mm channel on the position D works effectively. This indicates that ideal positioning for additional channels is related to the wavelength. In other words, when the target frequency of the additional channel is high, the channels should be assigned close to the top of the barrier; when the target frequency is low, the channels can be effective even if assigned far from the barrier top.

Based on these results, the next step is to determine the cross-sectional shape of the waterwheel-shaped edge that will be appropriate for broadband traffic noise. Additional channels with depths of 260, 80 (described previously), and 430 mm (corresponding to a quarter of the wavelength at 200 Hz) were added to a waterwheel of 800-mm diameter and 130-mm basic channels, as shown in Fig. 14. The cross sec-

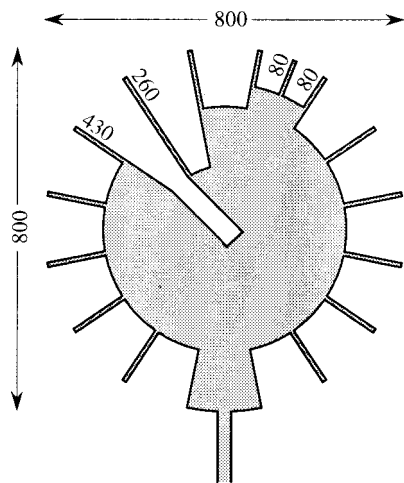


FIG. 14. The waterwheel-shaped edge with the basic 130-mm channels and the additional channels of 80, 260, and 430 mm.

tion of the deeper half of the 430-mm channel is rectangular because keeping a channel wedge shaped as it becomes deeper results in the bottom of the channel becoming too small. Thus, neither the volume of the air in the channel nor the frequency minimizing of the open-end sound pressure are affected, even by large changes of channel depth. Therefore, the channel bottom must be large enough to relate the change of the depth with the change of the frequency that is developing the soft surface. Figure 15 shows the effect of a waterwheel-shaped edge that has channels of four different depths. The two peaks that remain with uniform channels of 130 mm are reduced although the improvement between 630 Hz and 1 kHz that is obtained by uniform channels is decreased. Consequently, the shape of the noise spectrum becomes rather flat and the improvement in overall SPL changes from 2.8 to 4.3 dB by the additional channels.

As discussed previously, shallow channels should be assigned near the top of the barrier, so BEM calculations were performed by switching 80- and 130-mm channels at the top of the barrier in Fig. 14. The resulting spectrum of improvement is not the expected shape, probably due to either the

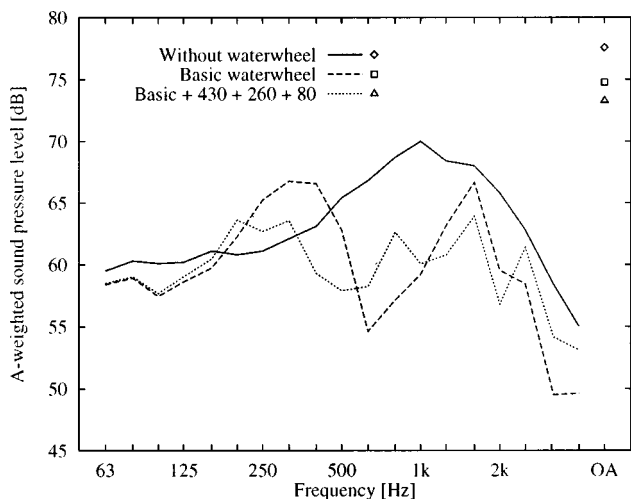


FIG. 15. Improvement of the negative effect by the combination of the channel depths.

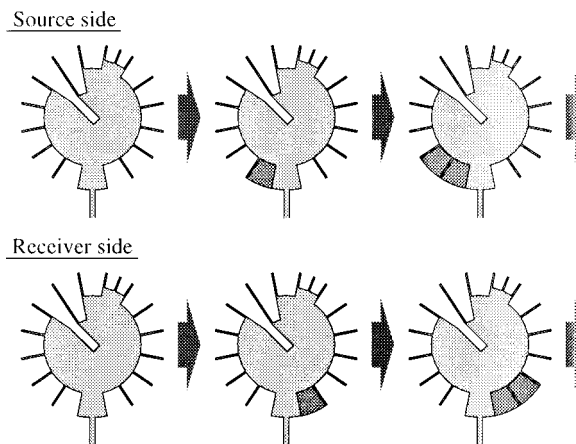


FIG. 16. Channels filled up gradually to determine the minimum number of channels to obtain the same improvement.

strong effect of the 80-mm channels or the decrease in the improvement by the 130-mm channel. Although this result suggests that the channel on the highest position of the waterwheel-equipped barrier affects efficiency improvement very strongly, it must be remembered that this discussion is based on the assumptions that each channel independently affects the improvement of the efficiency, and that the interaction between adjacent channels of different depths is not considered.

### C. Minimum number of channels

The discussions in the previous section show that four kinds of channels are needed on the top of the waterwheel-shaped barrier. Now it is investigated how 130-mm channels other than the four kinds of channels (i.e., five channels on the source side and six channels on the receiver side) affect the efficiency of the waterwheel shapes, in order to determine the minimum number of the channels to obtain the same improvement. Change in the improvement by the waterwheel-shaped edge is calculated by filling up the channels one by one, starting with the channel furthest from the top of the barrier, as shown in Fig. 16. Six channels on the receiver side are not filled up when the source side channels are filled up, and similarly five channels on the source side are not filled up when the receiver side channels are filled up. Figure 17 shows the change in the overall SPL of the noise spectrum as a function of the number of filled channels. As regards both source and receiver side channels, the overall SPL remarkably increases when the fourth channel is filled up. Namely, the channels arrayed on the upper half of the cylinder surface are necessary and those on the lower half do not affect the efficiency. In other words, even if all channels of the source or receiver sides are filled up, the SPL behind the waterwheel-equipped barrier increases by only 0.8 dB.

## III. WATERWHEEL FOR PRACTICAL USE

When one applies the waterwheel investigated in this paper to the practical noise barrier, some modifications are still needed. The horizontal width of the cylinder sticking out of the barrier must not exceed legal building restrictions on expressway facilities. The cylinder had better weigh as little

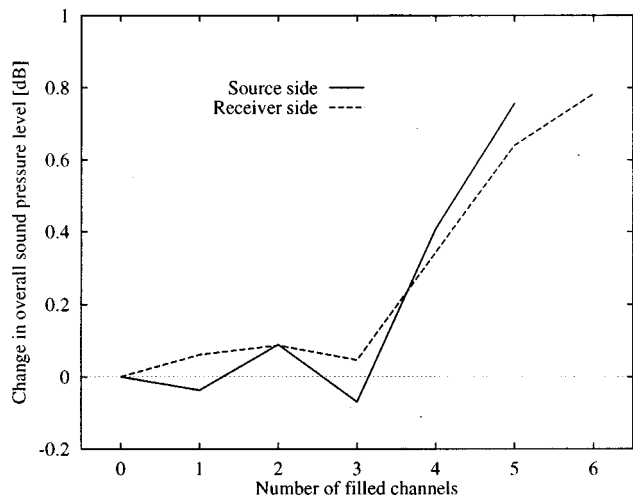


FIG. 17. Effect of the filled channels: change in the overall sound pressure level of the expressway noise.

as possible for the structural strength of the noise barrier. Furthermore, a smaller cylinder can suppress the cost of its material. These imply that the waterwheel of 800-mm diameter is not practical enough to be applied.

Thus, the cross-sectional design is based on the discussions in the previous section: the combination of various-depth channels at the top of the barrier, and the omission of the channels arranged on the lower half of the cylinder. Furthermore, the minimization of the overall volume of the waterwheel-shaped edge is considered. The decrease in overall SPL of the noise spectrum shown in Fig. 4 was evaluated, resulting in the cross-section diagram shown in Fig. 18. There are five kinds of channels, 90, 160, 260, 340, and 430 mm; they correspond to a quarter of the wavelength at 940, 530, 330, 250, and 200 Hz, respectively. This combination of channels is obtained by the procedure described in the previous section: the remaining peaks from a waterwheel-shaped edge with uniform 160-mm channels are reduced by the addition of four kinds of channels. Three of these additional channels are bent to minimize the volume of the waterwheel. Results of some numerical calculations indicated that it is appropriate for the channels to be bent twice to avoid reflection in the channel. If the channel is bent only once, the bending angle becomes acute and the reflection caused around the bend harmfully reduces efficiency; if the channel is bent more than three times, the results are the

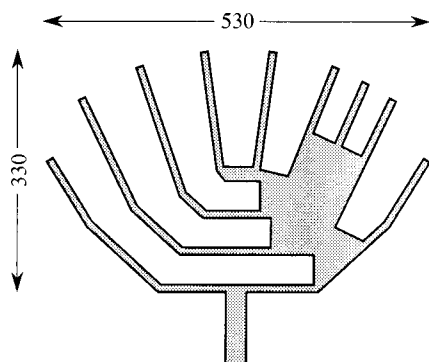


FIG. 18. New waterwheel-shaped edge designed for practical use.

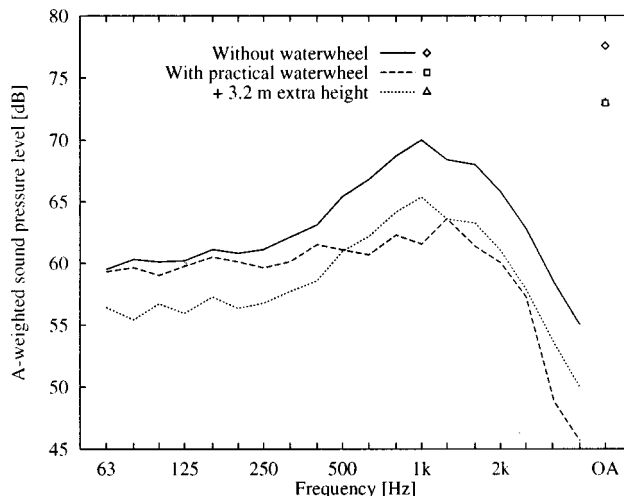


FIG. 19. Improvement by the practical waterwheel-shaped edge.

same as bending it twice. By bending the channels, the size of the waterwheel can be minimized to 530 mm wide and 330 mm high. This width is almost the same as the absorbing cylinder already being used, thus it will not be against the building regulations of expressway.

Figure 19 shows the change in the noise spectrum that results from installation of the practical waterwheel-shaped edge on the barrier. The spectrum with the waterwheel (dashed line) becomes rather flat and the negative effect of the waterwheel, which had harmed overall efficiency, is completely removed, resulting in an improvement of 4.7 dB in overall SPL. This improvement is approximately twice that of the uniform depth waterwheel with a 590-mm diameter (2.3 dB; see Fig. 9). The dotted line in Fig. 19 shows the spectrum behind a simple barrier of 6.2-m height. The 6.2-m simple barrier is less effective than a 3.0-m barrier with the waterwheel around 1 kHz, although it is efficient in the lower frequency range. Its decrease in overall SPL is almost the same as that of the practical waterwheel, thus installation of the waterwheel is equivalent to 3.2 m of extra height of a simple barrier. Notice that the cross section shown in Fig. 18 is designed only to control the noise source which has the spectrum shown in Fig. 4, and that a different cross section would be needed for a different noise spectrum. In other words, precise prediction of the traffic noise spectrum to be suppressed is quite significant in designing a noise control device, such as the waterwheel-shaped edge, whose effect is strongly dependent on frequency.

Spatial distribution of the effect of the practical waterwheel-shaped edge was calculated in the back of the barrier. The calculation was carried out in the region of 5–30 m horizontally from the barrier and ground surface to 5 m high, as shown in Fig. 20, with the receivers set in an array (26×6) at intervals of 1.0 m. Figure 21 shows the results behind a barrier of 3.0-m height placing a line source in front of the barrier at a horizontal distance of (a) 20.0 m, (b) 12.5 m, and (c) 5.0 m. Positive values in Fig. 21 indicate improvement of the sound shielding efficiency, i.e., decrease in overall SPL. The source positions correspond to the farthest lane, the central reservation, and the nearest lane, respectively, of an expressway with four traffic lanes. The improvement by

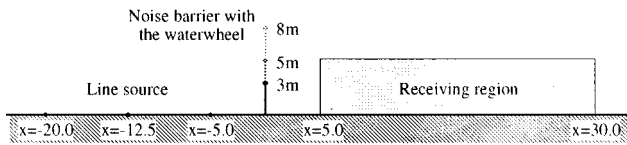
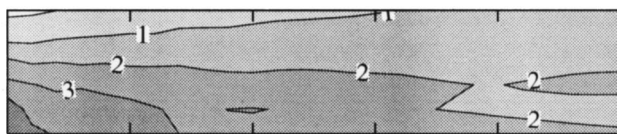
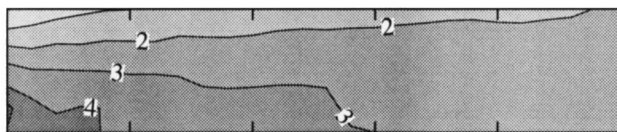


FIG. 20. Configuration of the two-dimensional sound field to calculate spatial distribution of the overall effect of the practical waterwheel.

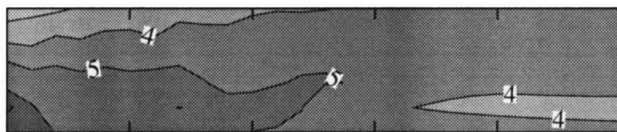
the waterwheel is large when the source is close to the barrier. Figure 21(c)–(e) shows the results behind a barrier of 3.0-, 5.0-, and 8.0-m height, respectively, placing a line source in front of the barrier at a distance of 5.0 m. It is shown that the improvement increases with an increasing height of the barrier. There are some dips of the improvement, for example, in the region of 22–30 m and 1 m high in Fig. 21(c). In this region, SPL around 1 kHz is minimized before installation of the waterwheel, because of the interference caused by the receiver height. Thus the effect of the waterwheel is wasteful because the sound pressure which is already minimized cannot decrease any more, although the waterwheel-shaped edge is quite efficient in this frequency range. As a result, overall SPL is less affected by the waterwheel-shaped edge. The average of the improvement in the calculated region ranges approximately from 2 dB [Fig. 21(a)] to 7 dB [Fig. 21(e)]: the waterwheel is very effective when a noise barrier is high and a source is close to the barrier. Therefore, the waterwheel-shaped edge should be



(a) 3.0 m barrier, source at -20.0 m



(b) 3.0 m barrier, source at -12.5 m



(c) 3.0 m barrier, source at -5.0 m



(d) 5.0 m barrier, source at -5.0 m



(e) 8.0 m barrier, source at -5.0 m

FIG. 21. Spatial distribution of the decrease in the overall SPL by the practical waterwheel-shaped edge.

equipped on the edge of a higher barrier which is efficient even without the waterwheel, whereas it is against our policy to improve the efficiency of a barrier without increasing its height.

#### IV. CONCLUSIONS

The sound shielding efficiency of a noise barrier with a waterwheel-shaped edge was investigated, and the relationship between the efficiency and the cross section of a waterwheel installed on the barrier edge was discussed. The original waterwheel-shaped edge, with uniform depth channels, was less effective than expected because the sound shielding efficiency is not changed or even decreased in any frequency range where the soft surface is not developed. Thus, the assignment of additional channels corresponding to the ineffective range extends the effective range of the waterwheel-shaped edge. The assignment of channels with various depths decreases the number of basic channels, and consequently decreases the large improvement originally obtained in the effective frequency range. Nevertheless, the reduction of the “negative effect” successfully results in an improvement in overall SPL. Also, additional channels work effectively when they are located close to the top of the barrier.

Next, the minimum number of channels that would be required to equal the efficiency of the waterwheel-shaped edge having channels arranged around the full circumference was investigated. The results show that channels located far from the edge hardly affect the efficiency; hence, channels are only needed in the upper half of the cylinder. As a result, the characteristics of the waterwheel-shaped edge are determined by the channels close to the top of the barrier. These results were used to design the practical waterwheel-shaped edge, with channels of various depth on the upper half of a cylinder. In general, smaller devices are preferred for practical applications, and the waterwheel-shaped device described in this paper is quite effective, despite its small (30 cm by 150 cm) size.

Discussions in this paper are highly dependent on the BEM model. It has already been shown by experiments using 1/10 scale model that the two-dimensional BEM model can be used to evaluate the efficiency of noise barriers in a three-dimensional sound field,<sup>6</sup> thus the waterwheel-shaped device designed in this paper should be efficient as predicted by the BEM. Currently full-scale experiments in a hemi-anechoic chamber and in an outdoor field are testing the efficiency of the practical waterwheel-shaped edge made of polycarbonate to confirm the BEM results.

<sup>1</sup>K. Fujiwara and N. Furuta, “Sound shielding efficiency of a barrier with a cylinder at the edge,” *Noise Control Eng. J.* **37**, 5–11 (1991).

<sup>2</sup>K. Fujiwara, “Sound shielding efficiency of a barrier with soft surface,” *Proc. Internoise '90*, 343–346 (1990).

<sup>3</sup>J. B. Keller and D. G. Magiros, “Diffraction by a semi-infinite screen with a round end,” *Commun. Pure Appl. Math.* **14**, 457 (1961).

<sup>4</sup>A. D. Rawlins, “Diffraction of sound by a rigid screen with a soft or perfectly absorbing edge,” *J. Sound Vib.* **45**, 53–67 (1976).

<sup>5</sup>T. Okubo and K. Fujiwara, “Efficiency of a noise barrier with an acoustically soft cylindrical edge,” *J. Acoust. Soc. Jpn.* **52**, 957–966 (1996) [English translation: *J. Acoust. Soc. Jpn. (E)* **19**, 187–197 (1998)].

- <sup>6</sup>T. Okubo and K. Fujiwara, "Efficiency of a noise barrier on the ground with an acoustically soft cylindrical edge," *J. Sound Vib.* **216**, 771–790 (1998).
- <sup>7</sup>D. C. Hothersall, S. N. Chandler-Wilde, and M. N. Hajmirzae, "Efficiency of single noise barriers," *J. Sound Vib.* **146**, 303–322 (1991).
- <sup>8</sup>R. Seznec, "Diffraction of sound around barriers: use of the boundary elements technique," *J. Sound Vib.* **73**, 195–209 (1980).
- <sup>9</sup>H. A. Schenck, "Improved integral formulation for acoustic radiation problems," *J. Acoust. Soc. Am.* **44**, 41–58 (1968).
- <sup>10</sup>A. J. Burton and G. F. Miller, "The application of integral equation methods to the numerical solutions of some exterior boundary value problems," *Proc. R. Soc. London, Ser. A* **323**, 201–210 (1971).
- <sup>11</sup>T. Sone, S. Kono, and T. Iwase, "Power levels and their spectra of automobile noise," *J. Acoust. Soc. Jpn.* **50**, 233–239 (1994) (in Japanese).

# Demographic and attitudinal factors that modify annoyance from transportation noise

Henk M. E. Miedema<sup>a)</sup> and Henk Vos

*TNO Prevention and Health, P.O. Box 2215, 2301 CE Leiden, The Netherlands*

(Received 4 September 1998; revised 11 February 1999; accepted 14 February 1999)

The effect of demographic variables (sex, age, education level, occupational status, size of household, homeownership, dependency on the noise source, and use of the noise source) and two attitudinal variables (noise sensitivity and fear of the noise source) on noise annoyance is investigated. It is found that fear and noise sensitivity have a large impact on annoyance (DNL equivalent equal to [at most] 19 and 11 dB, respectively). Demographic factors are much less important. Noise annoyance is not related to gender, but age has an effect (DNL equivalent equal to 5 dB). The effects of the other demographic factors on noise annoyance are (very) small, i.e., the equivalent DNL difference is equal to 1–2 dB, and, in the case of dependency, 3 dB. The results are based on analyses of the original data from various previous field surveys of response to noise from transportation sources (number of cases depending on the variable between 15 000 and 42 000).

© 1999 Acoustical Society of America. [S0001-4966(99)03905-3]

PACS numbers: 43.50.Qp, 43.50.Sr [MRS]

## INTRODUCTION

Various authors (Schultz, 1978; Fidell, Barber, and Schultz, 1991; Miedema and Vos, 1998) established for transportation noise simple relationships between noise exposure (descriptor: DNL) and annoyance (descriptor: percentage highly annoyed persons, %HA). They did not investigate the causes of variation in individual reactions to exposures with equal DNL level. Apart from random factors, demographic and attitudinal variables may have a systematic effect on annoyance. Fields (1992, 1993) investigated the effect of personal and situational factors on annoyance by a meta-analysis. He concluded that the demographic variables studied (age, sex, socioeconomic status, income, education, homeownership, type of dwelling, use of the noise source, dependency on the source) do not have an (important) effect on annoyance, and that the attitudinal variables fear of the source, feeling that noise annoyance is preventable, and noise sensitivity have an important effect on annoyance. Fields (1993) stated that for most variables, the chief impediments to forming strong conclusions are not inherent differences in data, but rather differences in published presentations of data. According to Fields, for these issues the most efficient advances in our knowledge can come from secondary analyses of existing datasets. Such analyses are presented in this paper, so that stronger conclusions can be formulated with respect to the effect of demographic and attitudinal variables on noise annoyance.

## I. DATA

The last decennium TNO in Leiden has been building an archive of original datasets from studies on annoyance caused by environmental noise. Investigations on different modes of transportation (aircraft, road traffic, and railway) are included. They were carried out in Europe, North

America, and Australia. As far as possible, a common set of variables is derived from the studies, which includes, among others, noise-exposure measures, noise-annoyance measures, and various demographic and attitudinal variables. Much effort has been put into the consistent derivation of the common variables from different studies. Studies are included in the archive if and only if DNL [or  $L_{Aeq}(24h)$ ] and %HA can be derived in such a way that they satisfy certain minimal criteria (see Miedema and Vos, 1998). Currently, the data archive contains 49 701 annoyance responses related to the exposure level of the respondent, and information on additional variables (see Table I). Here, such a combination is called a case. The total number of different respondents is lower than 49 701, because in some studies the respondents evaluated two noise sources. More information about the archive can be found in Miedema and Vos (1998), where exposure-response relationships are presented.

Demographic variables in the data archive are sex, age, education level, occupational status, size of household, homeownership, dependency on the noise source, and use of the noise source. Available attitudinal variables are noise sensitivity and fear of the noise source. Table I gives an overview of the information available per study.

Cases with  $DNL < 45\text{dB}$  or  $DNL > 75\text{dB}$  are excluded from the analyses of effects of personal and situational variables, because linear regression from annoyance to DNL is used and nonlinearity is expected outside the range 45–75 dB.

## II. METHOD OF ANALYSIS

Because noise-annoyance relationships depend on the study and the mode of transportation (Miedema and Vos, 1998), a variable “dataset” that discriminates between studies, and within studies between the modes of transportation investigated, is included in the analyses as an independent variable. Effects of secondary (i.e., demographic and attitudinal)

<sup>a)</sup>Electronic mail: hme.miedema@pg.tno.nl

TABLE I. Studies included in analyses in this paper. The study codes in the first column refer to Fields' (1994) catalog, the second column gives a key reference. Cases are counted in columns 3–5 if valid noise exposure and annoyance data are available. In general, the number of cases in analyses involving a demographic or attitudinal variable is less due to missing values for these variables. A "\*" in columns 6–12 indicates that a variable was included in a study. Age and sex are determined in all studies.

Study	Air	Road	Rail	Education	Occupation	Household size	Home ownership	Depend-ency	Use of source	Sensitivity	Fear
AUL-210	R. B. Bullen <i>et al.</i> , 1986	3212			*	*	*				*
CAN-120	J. S. Bradley, 1976		1112		*	*	*	*			*
CAN-121	F. L. Hall <i>et al.</i> , 1977		1147		*	*	*				
CAN-168	F. L. Hall <i>et al.</i> , 1981	635	574		*	*	*	*	*	*	
FRA-092	M. Vallet <i>et al.</i> , 1978		879		*	*	*		*	*	
FRA-239	M. Vallet <i>et al.</i> , 1986	570	570		*	*	*	*			*
FRA-364	M. Vallet <i>et al.</i> , 1996		848		*	*	*			*	
GER-192	V. Knall <i>et al.</i> , 1983		1579	1569	*	*	*	*	*		
GER-372	J. Kastka, no publication		551		*	*	*				
GER-373	J. Kastka, no publication		421		*	*	*			*	
NET-106	J. E. F. van Dongen, 1981		420			*	*			*	
NET-153	R. G. de Jong, 1979			602	*	*	*	*	*	*	
NET-240	I. D. Diamond <i>et al.</i> , 1986	573	474		*	*	*	*			
NET-258	J. E. F. van Dongen, 1981		304			*	*			*	
NET-276	H. M. E. Miedema <i>et al.</i> , 1988		697	265	*	*	*			*	
NET-361	R. G. de Jong <i>et al.</i> , 1994		788	69	*	*					
NET-362	W. J. Ericsz <i>et al.</i> , 1986		293		*	*	*			*	
NOR-311	T. T. Gjestland <i>et al.</i> , 1990	1396			*	*	*	*	*		*
NOR-328	T. T. Gjestland <i>et al.</i> , 1993a	687			*			*			
NOR-366	T. T. Gjestland <i>et al.</i> , 1991b	322			*			*			
SWE-365	E. Öhrström <i>et al.</i> , 1996			2802		*				*	
SWI-173	J. Nemecek <i>et al.</i> , 1981		1219		*	*	*		*		
UKD-24	A. E. Knowler, 1971	3845			*	*	*	*	*		*
UKD-071	F. J. Langdon, 1976a,b		2067		*	*	*			*	
UKD-072	D. G. Harland, 1977		904		*	*	*		*	*	*
UKD-116	J. M. Fields, 1977			1127	*	*	*	*	*	*	*
UKD-157	I. D. Griffiths <i>et al.</i> , 1980		302		*	*	*			*	
UKD-238	I. D. Diamond <i>et al.</i> , 1986	607	562		*	*	*	*			*
UKD-242	P. Brooker <i>et al.</i> , 1985	1993	410		*					*	
USA-022	W. R. Hazard, 1971	2247			*	*	*	*	*		*
USA-032	W. R. Hazard, 1971	1491			*	*	*	*	*		*
USA-044	H. P. Patterson <i>et al.</i> , 1973	1612			*	*	*	*	*		*
USA-082	S. Fidell <i>et al.</i> , 1975	374			*						*
USA-203	G. J. Raw <i>et al.</i> , 1985	629									
Total		20 193	16 121	6434							

dinal) variables are investigated by fitting the coefficients in the following linear models through multiple regression analysis:

$$A = a + bL + (c_1S_1 + \dots + c_mS_m + (d_1S_1L + \dots + d_mS_mL), \quad (1)$$

$$A = a + bL + (c_1S_1 + \dots + c_mS_m) + (d_1S_1L + \dots + d_mS_mL) + (e_1V_1 + \dots + e_nV_n), \quad (2a)$$

$$A = a + bL + (c_1S_1 + \dots + c_mS_m) + (d_1S_1L + \dots + d_mS_mL) + (f_1V_1L + \dots + f_nV_nL), \quad (2b)$$

$$A = a + bL + (c_1S_1 + \dots + c_mS_m) + (d_1S_1L + \dots + d_mS_mL) + (e_1V_1 + \dots + e_nV_n) + (f_1V_1L + \dots + f_nV_nL), \quad (3)$$

where  $A$  is the annoyance score;  $L$  is the DNL;  $S_i$  is the indicator variable which is 1 for cases from dataset  $i$  and otherwise 0; and  $V_j$  is the indicator variable which is 1 for

cases in category  $j$  of a secondary variable and otherwise 0.

Model (1) assumes separate linear noise-annoyance relationships for individual datasets. The parameters of these relationships can be derived from the model parameters  $a$ ,  $b$ ,  $c_i$ ,  $d_i$  as follows. If there are  $m + 1$  studies, then  $m$  indicator variables  $S_i$  are used in the above models. Each indicator variable is equal to 1 for the cases in one dataset. This means that these indicators are all equal to 0 for the remaining dataset. Consequently, in the first model the parameters  $a$  and  $b$  (intercept and slope) are determined by the cases in the study for which there is no indicator variable. The parameters  $c_i$  and  $d_i$  are the additive "corrections" to  $a$  and  $b$ , respectively, which give the intercept and the slope for study  $i$ . Thus, separate linear relationships are fitted for the individual datasets. Differences between studies, e.g., related to the use of different annoyance questions, are reflected by differences between the individual relationships. An effect of a secondary variable is described in the following models as a change relative to the individual relationships, and therefore it is not likely that this effect depends on differences



TABLE II. The multiple correlation coefficients,  $r$ , for model (1), and the increase in the square of this coefficient,  $\Delta r^2$ , for model (2a), (2b), and (3). An asterisk (\*) indicates that an increase is significant at the 1% level.

	Model (1) $r$	Model (3)			
		Model (2a) $\Delta r^2$ $re$ Model (1)	Model (2b) $\Delta r^2$ $re$ Model (1)	$\Delta r^2$ $re$ Model (2a)	$\Delta r^2$ $re$ Model (2b)
Sex	0.48	0.0000	0.0000	0.0000	0.0000
Age	0.48	0.0025*	0.0027*	0.0005*	0.0003
Education	0.47	0.0008*	0.0009*	0.0006*	0.0004*
Occupation	0.48	0.0004*	0.0005*	0.0003*	0.0003
Household size	0.50	0.0010*	0.0010*	0.0004*	0.0004*
Homeownership	0.46	0.0012*	0.0012*	0.0000	0.0001
Dependency	0.46	0.0012*	0.0014*	0.0006*	0.0004*
Use	0.47	0.0006*	0.0006*	0.0002	0.0002
Sensitivity	0.53	0.0405*	0.0437*	0.0071*	0.0039*
Fear	0.44	0.1148*	0.1135*	0.0000	0.0014*

between studies in the measurement of exposure and annoyance.

In model (2a), the intercept of the linear noise-annoyance relationship also depends on the level of the secondary variable; in model (2b) the slope depends on that variable, and in model (3) the intercept and the slope depend on the secondary variable. If the secondary variable being investigated has  $n+1$  categories, then  $n$  indicator variables  $V_j$  are used in these models. These indicators are all equal to 0 for the remaining category. Consequently, in models (2a), (2b), and (3), the parameters  $a$  and  $b$  are determined by the cases in the study for which there is no indicator variable and which come in the category of the secondary variable for which there is no indicator variable. The parameters  $e_j$  and  $f_j$  are the additive corrections to  $a$  and  $b$ , respectively, which give the intercept and the slope for category  $j$ .

Model (2a) in which a secondary variable is assumed to have a constant effect on annoyance that does not depend on the exposure level, is called a constant effect model. Model (2a) includes constant effects of the categories of one secondary variable. It can be extended to include constant effects of more than one variable. Sex and age are known in all datasets so that it is possible to fit constant effect models which include these demographic variables in addition to another secondary variable. By including these variables, the effect of a secondary variable on annoyance is found, taking into account possible effects of the demographic variables sex and age on annoyance. Models extended with sex and age will be fitted to the data only if these demographic variables are found to have a significant effect on annoyance.

DNL is used as the exposure metric, and the annoyance score is used as the annoyance variable. DNL is chosen as the exposure metric because it is being used in practice, and it has been the exposure metric in important previous exposure response analyses (see the introduction). Furthermore, it generally has a high correlation with other  $L_{Aeq}$ -based metrics, which are widely used. Effects found for secondary variables will not depend on the selection of the exposure metric from these related metrics.

The annoyance score is obtained by assigning numbers to annoyance categories. A typical noise annoyance question is: ‘‘How would you describe your general feelings about the

aircraft noise in this neighborhood? Would you say you are: (1) not at all annoyed, (2) slightly annoyed, (3) moderately annoyed, (4) considerably annoyed, or (5) highly annoyed?’’ On the basis of the assumption that each category of this five-point category scale occupies an equal portion of the annoyance continuum, the midpoints 10, 30, 50, 70, and 90 of five categories from 0 to 100 are assigned as scores to these categories. Scores are assigned in a similar manner if a different number of categories is used. The general rule is

$$\text{score}_{\text{category } i} = 100(i - 1/2)/m,$$

where  $m$  is the number of categories and  $i=1, \dots, m$  is the rank number of the category. For corrections because of expected effects of specific, unusual verbal category labels used in some studies, see Miedema and Vos (1998). The same procedure is used to translate secondary variables into scales from 0 to 100.

### III. RESULTS

The above models are fitted for each secondary variable. Table II gives the multiple correlation coefficients,  $r$ , for model (1). For model (2a) and (2b), the increase in the square of this coefficient (=increase of proportion of explained variance),  $\Delta r^2$ , relative to model 1 is given. For model 3, the increases relative to model (2a) and (2b) are given. It is indicated whether an increase is significant at the 1% level.

Table II shows that noise annoyance is related to ‘‘noise sensitivity’’ and especially to ‘‘fear associated with the source.’’ Noise annoyance has no relation with sex, and a significant but (very) weak relation with the other variables. Of these other variables, ‘‘age’’ has the strongest relation with noise annoyance.

Table III reproduces  $\Delta r^2$  [relative to model (1)] for model (2a) from Table II. Model (2a) is a constant effect model with only one secondary variable. In addition, a constant effect model with a secondary variable combined with age is investigated [model (2a) and age], and  $\Delta r^2$  relative to the constant effect model which only includes age [model (1) and age] is reported in Table III. It is indicated whether an increase is significant at the 1% level. Table III shows that

TABLE III. The increase in the square of the correlation coefficient,  $\Delta r^2$ , for model (2a) and for model (2a) and age. All  $\Delta r^2$  are significant at the 1% level.

	Model (2a) $\Delta r^2$ re Model (1)	Model (2a) and age $\Delta r^2$ re Model (1) and age
Education	0.0008	0.0005
Occupation	0.0004	0.0003
Household size	0.0010	0.0009
Homeownership	0.0012	0.0011
Dependency	0.0012	0.0012
Use	0.0006	0.0007
Sensitivity	0.0405	0.0387
Fear	0.1148	0.1139

the relation of noise annoyance with all but one of the secondary variables becomes slightly weaker if the effect of age is taken into account. The strongest reduction is found for education. The relation with “use of the source” becomes slightly stronger if the effect of age is taken into account.

The outcomes for the individual secondary variables are discussed in more detail in separate sections. If the increase in  $\Delta r^2$  is significant for model (2a) (all secondary variables except sex), then the estimates of  $e_j$  are presented. Because the increase in  $\Delta r^2$  for model (3) is small compared to the increase for model (2a) and (2b), no estimates of the parameters in model (3) will be given. If it is not obvious how the information about a secondary variable was obtained in the surveys, then this is discussed.

### A. Sex of respondent

Table II shows that men and women react in a similar way to transportation noise. The overrepresentation of women in the studies (57.5% is female) is probably related to the fact that they were at home more. Consequently, the outcome indicates indirectly that annoyance is not higher if more time is spent at home.

### B. Age

Table II shows that age has an effect on annoyance. Model (2b) has a slightly better fit than model (2a), which indicates that the effect of age depends on the noise level. Table IV shows that relatively young and relatively old persons are less annoyed. The contribution of age to the explained variance,  $\Delta r^2$ , would be greater if the proportion of respondents in the extreme age classes, which have the largest effect, were greater.

Fields (1993) did not find clear evidence that noise annoyance is related to age. This can be explained at least

TABLE IV. Extra noise annoyance (scale: 0–100) relative to the category 30–39, i.e., the estimates of parameters  $e_j$  in model (2a) and the distribution of the variable “age of the respondent.”

	Age							N
	10–19	20–29	30–39	40–49	50–59	60–69	70+	
Effect ( $e_j$ )	-5.6	-1.5	0	0.4	-0.7	-2.1	-4.2	
Distribution	3%	18%	20%	19%	17%	15%	8%	42 496

TABLE V. Extra noise annoyance (scale: 0–100) relative to the category “2nd level,” i.e., the estimates of parameters  $e_j$  in model (2a) and model (2a) and age, and the distribution of the variable “education level of the respondent.”

	Education level				N
	1st	2nd	High	University	
Effect ( $e_j$ )	-1.0	0	0.7	2.3	
Effect ( $e_j$ ) with age	-0.8	0	0.6	1.9	
Distribution	35%	45%	11%	9%	32 254

partly by the fact that many studies that were included in his meta-analysis only investigated the existence of a linear relationship between age and annoyance. Because, according to our findings, relatively young as well as relatively old persons report less annoyance, there is no linear relationship, but a relationship that has a curvilinear shape.

### C. Education level

Information about education level was translated into categories based upon the International Standard Classification of Education (ISCED). The ISCED defines levels of education uniformly across all countries participating in the United Nations Educational, Scientific, and Cultural Organization (UNESCO). The original ISCED categories were combined into four broader classes because often, available information was not a sufficient basis for a finer distinction. The definition of education levels in terms of the ISCED categories (4 does not exist) is as follows:

1st	=ISCED 0, 1	=completed primary school;
2nd	=ISCED 2, 3	=completed secondary school, high school;
High	=ISCED 5	=completed higher education that does not lead to a first university degree;
University	=ISCED 6, 7	=(polytechnics) university, first degree and postgraduate level.

In some studies it was not the education level, but the number of years of education that was determined, or the age when education was finished. In the latter case, 6 years is subtracted from this age to obtain an estimate of the length of education, on the basis of the assumption that education starts at the age of 6. If only (an estimate of) education length was available, it was used to estimate the education level. Based upon the joint distribution of education length

TABLE VI. Extra noise annoyance (scale: 0–100) relative to the category “medium,” i.e., the estimates of parameters  $e_j$  in model (2a) and model (2a) and age, and the distribution of the variable “occupational status.”

	Occupational status			N
	Low	Medium	High	
Effect ( $e_j$ )	-0.9	0	0.7	
Effect ( $e_j$ ) with age	-0.8	0	0.6	
Distribution	24%	42%	34%	27 018

TABLE VII. Extra noise annoyance (scale: 0–100) relative to the category “2 person household,” i.e., the estimates of parameters  $e_j$  in model (2a) and model (2a) and age, and the distribution of the variable “size of household.”

	Size of household					<i>N</i>
	1	2	3	4	≥5	
Effect ( $e_j$ )	-2.7	0	0.1	-0.6	-0.5	
Effect ( $e_j$ ) with age	-2.6	0	-0.3	-1.3	-1.3	
Distribution	14%	30%	20%	19%	17%	27 241

and education level in the study NET-240, where both variables were available, 1–8 years of education was coded as “1st level,” 9–12 years as “2nd level,” 13–14 years as “high level,” and 15 and more years of education was coded as “university level.”

Table V shows that a little higher noise annoyance is reported if the education is higher. If age is taken into account, then the effect of education on noise annoyance becomes slightly smaller.

#### D. Occupational status

The available information about occupation and occupational status is very diverse. Categories were combined and sometimes ordered to create classes ordered according to occupational status. For example, in study NOR-311 there was a classification of blue collar workers [(1) unskilled and (3) skilled] and a classification of white collar workers [(2) general office worker, (4) professional worker, and (5) manager]. The numbers indicate the rank order used with this information. Such a rank order is not self-evident and involves a subjective judgment. If occupational status was determined for more than one member of the household, then the highest level was taken. Respondents without occupation (student, homemaker, retired, etc.) were excluded from the analysis, except when there was information about the occupation of another member of the household (breadwinner). The number of categories was not equal for all studies. In order to obtain a comparable indicator for occupational status, the categories were translated into a 0–100 scale in the same manner as described above for annoyance. In the present analysis, this scale is divided into three levels: low (0–32), medium (33–66), and high (67–100).

The results in Tables II and VI indicate that the reaction to transportation noise is not strongly related to occupational status. Education and occupational status are two components of socioeconomic status. In general, the classification of the education will be more accurate and unambiguous

TABLE VIII. Extra noise annoyance (scale: 0–100) relative to the category “renter,” i.e., the estimates of parameter  $e_j$  in model (2a) and model (2a) and age, and the distribution of the variable “homeownership.”

	Type of tenure		<i>N</i>
	Renter	Owner	
Effect ( $e_j$ )	0	2.3	
Effect ( $e_j$ ) with age	0	2.2	
Distribution	44%	56%	33 343

TABLE IX. Extra noise annoyance (scale: 0–100) relative to the category 30–39, i.e., the estimates of parameters  $e_j$  in model (2a) and model (2a) and age, and the percentage of respondents per category of the variable “dependency of the respondent on the noise maker.”

	Dependency		
	Depend not	Depend	<i>N</i>
Effect ( $e_j$ )	0	-3.6	
Effect ( $e_j$ ) with age	0	-3.7	
Distribution	91%	9%	21 516

than the classification of occupation. This may partly explain the smaller effect found for occupational status.

#### E. Number of persons in household

Noise annoyance has a weak relation with the size of the household (Table II). Single persons tend to be less annoyed by noise, and persons in a household with two persons tend to be more annoyed by noise, also if the effect of age is taken into account (Table VII).

#### F. Homeownership

Homeowners are expected to be more concerned about environmental noise. This hypothesis is confirmed (Table II). Table VIII shows that the effect of the type of tenure is small, and is not substantially affected by taking the age of the respondent into account.

#### G. Dependency on the noise source

Respondents who depend economically on the activities that cause the noise are expected to report less annoyance than persons who have no economic relation with that activity. Many aircraft noise surveys and three railway noise surveys included a question about this kind of dependency (see Table I). An example of a question about economic dependency is “Are you or anyone else in the family employed at this time by the airport or by an airline company?” with, for example, the following response categories: “yes” and “no.” Table II shows that noise annoyance is related to economic dependency, and Table IX gives estimates of the effect of dependency.

#### H. Use of the noise source

Respondents who frequently use the mode of transportation that causes the noise are expected to report less annoyance than persons who do not use it much. An example of a

TABLE X. Extra noise annoyance (scale: 0–100) relative to the category “no user,” i.e., the estimates of parameters  $e_j$  in model (2a) and model (2a) and age, and the distribution of the variable “use of the transportation concerned.”

	User				
	Not	Low	Use	High	<i>N</i>
Effect ( $e_j$ )	0	-0.9	-1.3	-2.2	
Effect ( $e_j$ ) with age	0	-1.0	-1.4	-2.4	
Distribution	40%	13%	29%	18%	16 800

TABLE XI. Extra noise annoyance (scale: 0–100) relative to the category “low sensitivity,” i.e., the estimates of parameters  $e_j$  in model (2a) and model (2a) and age, and the percentage of respondents per category of the variable “noise sensitivity reported by the respondent.”

	Noise sensitivity			<i>N</i>
	Low	Medium	High	
Effect ( $e_j$ )	0	6.8	16.0	
Effect ( $e_j$ ) with age	0	6.5	15.7	
Distribution	46%	32%	22%	15 171

question about the use of transportation is “Looking back at the last 12 months, how many times have you flown?” with, for example, the following response categories: “10 or more,” “4–9,” “more seldom,” “never.” The responses are translated into a 0–100 scale, which is divided in the present analysis into four categories: not (0–24), low (25–49), use (50–74), and high (75–100). Table II shows that noise annoyance is related to use of source, and Table X gives estimates of the effect of this variable.

### I. Noise sensitivity

Respondents who report to be sensitive to noise are expected to be more annoyed at the same exposure level. In various studies, a question such as “Are you sensitive to noise?” is included with, for example, the following response categories: “more than average,” “about average,” and “less than average.” In order to obtain a comparable indicator for noise sensitivity, the categories were translated to a 0–100 scale in the same manner as described above for

TABLE XII. Extra noise annoyance (scale: 0–100) relative to the category “low fear,” i.e., the estimates of parameters  $e_j$  in model (2a) and model (2a) and age, and the distribution of the variable “fear.”

	Fear			<i>N</i>
	Low	Medium	High	
Effect ( $e_j$ )	0	16.1 <sup>a</sup>	27.4 <sup>a</sup>	
Effect ( $e_j$ ) with age	0	16.1 <sup>a</sup>	27.3 <sup>a</sup>	
Distribution	62%	21%	17%	17 494

<sup>a</sup>Upper boundaries: see the text.

annoyance. In the present analysis, this scale is divided into three levels: low (0–32), medium (33–66), and high (67–100). The results in Tables II and XI show that noise sensitivity has the expected effect. Respondents who report to be not sensitive to noise report less annoyance than highly sensitive respondents.

### J. Fear

Respondents who express fear associated with the activity that causes the noise are expected to be more annoyed at the same exposure level. Fear with respect to the source is determined with a question such as “When you see or hear planes overhead, how often do you feel there is some danger that they might crash nearby?” with, for example, the following response categories: “never,” “sometimes,” “often,” and “very often.” In order to obtain a comparable indicator, again the categories were translated into a 0–100 scale in the same manner as described above for annoyance. In the present analysis, this scale is divided into three levels: low (0–32), medium (33–66), and high (67–100).

TABLE XIII. Questions on “fear” in the studies from which the data are included in the analyses here.

Study	Question	<i>N</i>
AUL-210	Have you ever thought that there is a danger that a plane might crash in this neighborhood? (yes) (no) Using the opinion thermometer, would you please estimate how much you feel afraid or worried about a possible plane crash in this neighborhood. Note: persons who never thought of a crash skipped the latter question. We assigned them to the lowest level on the opinion thermometer.	2345
CAN-120	How concerned are you that a traffic accident will occur near your home? (not at all)...(very much)	1112
FRA-239	see UKD-238	564
NOR-311	Do you ever consider that a plane crash may occur near the place where you live? (no) (yes) if so, do you think about it? (often) (now and then) (seldom) (never) (don't know)	1389
UKD-024	When you hear the aircraft fly overhead, do you ever feel there is a danger they might crash nearby? (yes)(no) (don't know) (inap) Would you say you feel this (very often) (fairly often) (only occasionally) (don't know)? Note: persons who did not hear aircraft skipped the latter question; we assumed that their answer to the first question would have been (no).	3829
UKD-072	Apart from the danger of crossing roads, do you ever feel in danger from traffic when you are walking on the pavements (or along beside the road) in this area? (yes) (no) (don't know/inap)	896
UKD-116	When you hear trains go by, do you ever feel there is any danger they might crash nearby? (no) (very often) (fairly often) (only occasionally) (don't know) (inap)	1113
UKD-238	When watching an aircraft fly over, have you ever been afraid that it would crash? (yes) (no) Has this happened to you: (several times) (a few times) (once or twice) (don't know)?	600
USA-022	When you see or hear planes overhead, how often do you feel there is some danger that they might crash nearby? (never)...(very often). Note: persons who never noticed aircraft skipped the question. We assigned them to the category “never” of the fear question.	2201
USA-032	see USA-022	1474
USA-044	see USA-022	1601
USA-082	see USA-022	370

TABLE XIV. Extra noise annoyance (scale: 0–100) relative to the category “low fear for the three modes of transportation.”

	Fear			N
	Low	Medium	High	
Aircraft	0 (63%)	16.8 <sup>a</sup> (21%)	30.2 <sup>a</sup> (16%)	14 373
Road traffic	0 (43%)	5.7 (19%)	13.5 (38%)	2008
Railway	0 (86%)	14.4 (13%)	43.1? (1%)	1113

<sup>a</sup>Upper boundaries: see the text.

Tables II and XII show that fear is related to noise annoyance in the expected way. At the same exposure level, persons with fear concerning the transportation that causes the noise are more annoyed than persons without such feelings. Because “fear” is found to be such an important determinant of annoyance and it is not obvious how questions about fear were formulated, the (translated) original questions are given in Table XIII.

The effect of fear on annoyance is probably overestimated because of the routing in the questionnaires used in the aircraft studies AUL-210, UKD-024, USA-022, USA-032, USA-044, and USA-082. Persons who did not notice the aircraft skipped the fear question and were assumed to experience no fear of the aircraft. This appears to be a reasonable assumption. Nonetheless, it will not always hold, and therefore it induces some correlation between fear and noise annoyance scores. Estimates of the effect of fear on annoyance based on studies with the described routing can be conceived as upper boundaries.

The nature of the threat is not the same for different modes of transportation. This may interact with the effect of fear on noise annoyance. Therefore, this effect was determined for the three modes of transportation separately by creating indicator variables for the modes of transportation, and extending model (2a) with the products of these indicator variables and the indicator variables for the fear categories. The increase in the square of the multiple correlation coefficient resulting from the addition of these terms to the model is equal to 0.0054, which is significant at the 1%

TABLE XV. Extra noise annoyance (DNL equivalent), the lower and upper bound of the 95%-confidence interval of the estimated extra annoyance, and the range of the estimates for different categories of the same variable.

Categories	Estimated	Lower bound 95% interval	Upper bound 95% interval	Range of Estimates	
	extra annoyance (DNL equivalent)				
Sex	...	...	...	0	
Age	10–20	–4.0	–5.2	–2.9	5
	20–30	–1.1	–1.6	–0.5	
	30–40	0	0	0	
	40–50	0.3	–0.3	.9	
	50–60	–0.5	–1.1	–0.1	
	60–70	–1.5	–2.1	–0.9	
Education	70+	–3.0	–3.7	–3	2
	1st	–0.5	–1.1	–0.0	
	2nd	0	0	0	
	High University	0.4 1.4	0.3 0.6	1.1 2.2	
Occupation	Low	–0.6	–1.1	–0.0	1
	Medium	0	0	0	
	High	0.4	–0.1	1.0	
Household size	1	–1.9	–2.6	–1.1	2
	2	0	0	0	
	3	–0.2	–0.9	0.4	
	4	–0.9	–1.6	–0.2	
	≥5	–1.0	–1.7	–0.2	
Homeownership	Renter	0	0	0	2
	Owner	1.6	1.1	2.1	
Dependency	Depend	0	0	0	3
	Not depend	–2.6	–3.5	–1.7	
Use	Not	0	0	0	2
	Low	–0.7	–1.7	0.2	
	Use	–1.0	–1.7	–0.3	
Sensitivity	High	–1.7	–2.7	–0.7	11
	Low	0	0	0	
	Medium	4.6	4.0	5.3	
Fear	High	11.2	10.5	12.0	19*
	Low	0	0	0	
	Medium	11.4*	11.1	11.8	
	High	19.5*	19.1	19.9	

\*Upper boundaries: see text.

level. The effects for fear estimated in this way per mode of transportation are given in Table XIV.

The results for road traffic and railway noise need to be interpreted with caution because they are based on only three studies (two on road traffic, one on railway noise), and the number of cases is limited. The results indicate that few persons experience fear due to railway transportation. The estimate of the extra annoyance associated with high fear due to trains is unreliable because it is based on very few respondents. The overall conclusion on the basis of Table XIV is that fear is an important determinant of noise annoyance for all modes of transportation.

Table XV gives an overview of the results. The rule used to translate an estimate of extra annoyance (scale 0–100) into the equivalent change in DNL (1.4 annoyance points  $\sim$ 1 dB) is based on a linear regression analysis from the annoyance score (scale 1–100) to DNL. This gave a slope for DNL of 1.44 if no distinction was made between the modes of transportation.

#### IV. CONCLUSION AND DISCUSSION

Fear has a very large impact on annoyance. Persons who experience fear related to the transportation that causes the noise report higher annoyance compared to persons who do not experience such fear. With three categories, the difference between the lowest and the highest fear level is equivalent to a DNL difference of (at most) 19 dB. The effect of fear is found for all three modes of transportation, but it appears that only few persons associate high fear with railway traffic. It is not clear whether the relation of annoyance with fear depends on the actual experience of fear, or is due to a common predisposition to noise annoyance and fear. This point is of practical interest. If actual fear is a condition for the increase in annoyance, then communication that reduces the feelings of danger may also reduce noise annoyance. However, there will be no effect of a reduction in the actual fear if a predisposition to experiencing fear is the cause of higher noise annoyance.

Another important factor is noise sensitivity. With three categories, the difference in annoyance between the nonsensitive and highly sensitive persons is equivalent to a DNL difference of circa 11 dB. The effect of noise sensitivity on annoyance is reduced very little if age is also taken into account.

Since fear and sensitivity appear to be such important factors, further attention to the measurement of these factors is worthwhile. The independence of the measurement of, on the other hand, fear and sensitivity, and, on the other hand, noise annoyance is a point of particular interest. Complications arise because these variables are rated by the same person at the same moment. Having rated fear and sensitivity may affect the rating of noise annoyance, or vice versa. Moreover, the rating of all these variables may be affected by a response tendency, e.g., the tendency to use a certain part of rating scales. These phenomena will introduce some correlation between the variables, and may account for a part of the effect of fear and sensitivity on noise annoyance that has been found. Even though it is expected that the effect of

these phenomena is limited, quantitative information on the strength of these effects would be valuable. In addition, it may be possible to find physiological correlates of noise sensitivity and a predisposition to fear, and to study the relation of these correlates with noise annoyance.

Demographic factors are much less important than fear and noise sensitivity. Noise annoyance is not related to gender, but age has an effect. The largest difference in annoyance between (seven) age classes is equivalent to a DNL difference of circa 5 dB. At the same noise-exposure level, relatively young and relatively old persons are less annoyed than the ages in between. The same general age-related pattern has been also found for annoyance caused by environmental odor from factories and agricultural activities (Miedema, 1992). This indicates that the pattern is not specific for environmental noise. There may be a general relationship between the importance of the environmental quality in the neighborhood and the phase in life. The lower annoyance found for older persons may also be caused by deterioration of the senses.

The findings are consistent with the hypotheses that higher annoyance is reported if the education is higher, the occupational status is higher, the dwelling is owned instead of rented, if a person does not depend on the noise source, and if the use of the transportation that causes the noise is low. In addition, a person in a household consisting of two persons reports more annoyance than a person in a household of another size, also if the age of the person is taken into account. However, the effects of these factors on noise annoyance are small, i.e., the equivalent DNL difference is equal to 1–2 dB, and, in the case of dependency, 3 dB.

- Bitter, C. (1979). "Perception and Experience of Traffic Noise in a Residential District along a State Highway," *Urb. Ecol.* **4**, 161–177, Elsevier, Amsterdam.
- Bradley, J. S. (1976). "A Second Generation Noise Survey," *Proceedings of Inter-Noise 1976*, Institute of Noise Control Engineering, Washington, DC, pp. 377–380.
- Brooker P., and Richmond, C. (1985). "The United Kingdom Aircraft Noise Index Study: Part I—Main Results," *Proceedings of the Institute of Acoustics 1985* (Springer, New York, April 1985), pp. 323–329.
- Bullen, R. B., Hede, A. J., and Kyriacos, E. (1986). "Reaction to Aircraft Noise in Residential Areas Around Australian Airports," *J. Sound Vib.* **108**, 199–225.
- Diamond, I. D., and Walker, J. G. (1986). "CEC Joint Research Project: Community Reactions to Aircraft Noise: Final Report," ISVR, University of Southampton.
- Dongen, J. E. F. van (1981). "Evaluation of Noise Abatement Measures," *Proceedings of Inter-Noise 81*, Nederlands Akoestisch Senootschap, Delft, pp. 813–816.
- Ericsz, W. J., Noordam, A., and Schoonderbeek, W. (1986). "Trollicering van buslijn 9 in Arnhem," *Onderzoek naar de effecten van geluidhinder*, Ministerie van VROM: Geluidreeks GA-HR-12-1, Leidschendam.
- Fidell, S., and Jones, G. (1975). "Effects of Cessation of Late-Night Flights on an Airport Community," *J. Sound Vib.* **42**, 441–427.
- Fidell, S., Barber, D. S., and Schultz, Th. J. (1991). "Updating a Dosage-Effect Relationship for the Prevalence of Annoyance due to General Transportation Noise," *J. Acoust. Soc. Am.* **89**, 221–233.
- Fields, J. M. (1977). "Railway Noise Annoyance in Residential Areas: Current Findings and Suggestions For Future Research," *J. Sound Vib.* **51**, 343–351.
- Fields, J. M. (1992). "Effect of Personal and Situational Variables on Noise Annoyance: With Special Reference to Implications for En Route Noise," National Aeronautics and Space Administration, Georgia Institute of Technology, Atlanta, GA.

- Fields, J. M. (1993). "Effect of Personal and Situational Variables on Noise Annoyance in Residential Areas," *J. Acoust. Soc. Am.* **93**, 2753–63.
- Gjestland, T. T., Liasjø, K. H., Granøien, I. L. N., and Fields, J. M. (1990). "Response to Noise Around Oslo Airport Fornebu," Report No. STF40 A90189, ELAB-RUNIT, Trondheim, Norway.
- Gjestland, T. T., Granøien, I. L. N., and Liasjø, K. H. (1993a). "Aircraft Noise Annoyance. Community Reactions to Noise around Bodo Airport. Part 1: Simple Dose-Response Relationships," SINTEF DELAB, Trondheim, Norway.
- Gjestland, T. T., Granøien, I. L. N., and Liasjø, K. H. (1993b). "Aircraft Noise Annoyance. Community Reactions to Noise around Trondheim Airport Vaernes. Part 1: Simple Dose-Response Relationships. SINTEF DELAB, Trondheim, Norway.
- Griffiths, I. D., Langdon, F. J., and Swan, M. A. (1980). "Subjective Effects of Traffic Noise Exposure: Reliability and Seasonal Effects," *J. Sound Vib.* **71**, 227–240.
- Hall, F. L., Birnie, S. E., Taylor, S. M., and Palmer, J. E. (1981). "Direct Comparison of Community Response to Road Traffic Noise and to Aircraft Noise," *J. Acoust. Soc. Am.* **70**, 1690–1698.
- Hall, F. L., and Taylor, S. M. (1977). "Predicting Community Response to Road Traffic Noise," *J. Sound Vib.* **52**, 387–399.
- Harland, D. G. (1977). "Exposure and Response to Traffic Noise," Proceedings of Inter-Noise 1977, Institute of Noise Control Engineering, Poughkeepsie, pp. A-115–A-121.
- Hazard, W. R. (1971). "Predictions of Noise Disturbance near Large Airports," *J. Sound Vib.* **15**, 425–445.
- de Jong, R. G. (1979). "A Dutch Study on Railway Traffic Noise," *J. Sound Vib.* **66**, 497–502.
- R. G. de Opmeer, C. H. J. M., and Miedema, M. E. (1994). "Hinder door milieuverontreiniging in Nederland," Report 94.056, TNO, Leiden, The Netherlands.
- Knall, V., and Schümer, R. (1983). "The Differing Annoyance Levels of Rail and Road Traffic Noise," *J. Sound Vib.* **87**, 321–326.
- Knowler, A. E. (1971). "The Second Noise and Social Survey Around Heathrow, London Airport," 7th International Congress on Acoustics, Budapest, Vol. 2, 18–26 August 1971, pp. 525–528.
- Langdon, F. J. (1976a). "Noise Nuisance Caused by Road Traffic in Residential Areas: Parts I, II," *J. Sound Vib.* **47**, 243–282.
- Langdon, F. J. (1976b). "Noise Nuisance Caused by Road Traffic in Residential Areas: Part II," *J. Sound Vib.* **49**, 241–256.
- Miedema, H. M. E., and van den Berg, R. (1988). "Community Response to Tramway Noise," *J. Sound Vib.* **120**, 341–346.
- 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. (1992). "Response Functions for Environmental Odour in Residential Areas," TNO-PG Report, Leiden, The Netherlands.
- Nemecek, J., Wehrli, B., and Turrian, V. (1981). "Effects of Street Traffic in Switzerland, a Review of Four Surveys," *J. Sound Vib.* **78**, 223–234.
- Öhrström, E., and Skånberg, A.-B. (1996). "A Field Survey on Effects of Exposure to Noise and Vibration from Railway Traffic, Part 1: Annoyance and Activity Disturbance Effects," *J. Sound Vib.* **193**, 39–47.
- Patterson, H. P., and Connor, W. K. (1973). "Community Responses to Aircraft Noise in Large and Small Cities in the U.S.A.," Proceedings of the International Congress on Noise as a Public Health Problem, Dubrovnik, Yugoslavia, 13–18 May, pp. 707–720, U.S. Environmental Protection Agency, Washington D. C.
- Raw, G. J., and Griffiths, I. D. (1985). "The Effect of Changes in Aircraft Noise Exposure," *J. Sound Vib.* **101**, 273–275.
- Schultz, T. H. J. (1978). "Synthesis of Social Surveys on Noise Annoyance," *J. Acoust. Soc. Am.* **64**, 377–405.
- Vallet, M. G., Pachiaudi, G., Bruyere, J., Signolles, C., Tanguy, Y., Depitre, A., Fischl, M., Francois, J., and Abramowitch, J. (1986). "Réactions de La Communaute Au Bruit Des Avions," Inrets, Bron, France
- Vallet, M., Maurin, M., Page, M. A., Favre, B., and Pachiaudi, G. (1978). "Annoyance from and Habituation to Road Traffic Noise from Urban Expressways," *J. Sound Vib.* **60**, 423–440.
- Vallet, M., Vernet, I., Champelovier, P., and Maurin, M. (1996). "A Road Traffic Noise Index for the Night Time," Proceedings of Inter-Noise 1996, Institute of Noise Control Engineering, Poughkeepsie, pp. 2345–2350.

# Speech intelligibility model including room and loudspeaker influences

L. Faiget<sup>a)</sup>

L.A.U.T.M, Université de Toulouse-Le Mirail, 5 allées Antonio Machado, 31058 Toulouse Cedex 1, France

R. Ruiz<sup>b)</sup>

L.A.R.A, Université de Toulouse-Le Mirail, 5 allées Antonio Machado, 31058 Toulouse Cedex 1, France

(Received 19 May 1998; accepted for publication 25 January 1999)

Loudspeaker directivity and frequency response are of great importance for speech intelligibility estimation. In this work their respective influence is introduced in a new predictor. Properties of the model are in good agreement with expected variations of scores when radiation and frequency response are modified. An experiment shows the accuracy of the predicted scores. Limitations of the model are discussed and future research perspectives are presented. © 1999 Acoustical Society of America. [S0001-4966(99)03405-0]

PACS numbers: 43.55.Hy, 43.71.Gv [JDQ]

## INTRODUCTION

The three main predictors of speech intelligibility in a room are the energy ratios related measures,<sup>1,2</sup> the Speech Transmission Index (S.T.I.),<sup>3</sup> and the Articulation loss of consonants (Alcons).<sup>4</sup> They can be directly computed from the impulse response of the loudspeaker-room-microphone system. With the two first predictors, intelligibility score prediction is very satisfactory. Correlation coefficients between predicted and measured scores are greater than 0.9 with small standard deviations.

Articulation loss of consonants depends on the distance, the volume of the room, and the reverberation time. A modified formulation introduces the directivity factor  $Q$  of the source.<sup>5</sup> Measured scores in different rooms with loudspeakers of high, medium, and low  $Q$  have been used to show an inaccurate prediction of Alcons method and better results for the other techniques.<sup>6,7</sup>

Alcons predictor seems to be the least accurate probably because the loudspeaker influence is limited to its directivity factor  $Q$ . The other techniques include *in situ* radiation and frequency response effects in the measured impulse response.

The object of this paper is to introduce loudspeaker contribution to intelligibility in a model based on an impulse response estimation.<sup>8</sup> Energy ratio based predictors are chosen to separate room and loudspeaker influences on intelligibility scores.

The basis of the work is a model derived from the Lochner and Burger signal-to-noise ratio<sup>2</sup> and from the useful-to-detrimental energy ratio of Bradley.<sup>1</sup> It is modified in order to introduce the loudspeaker directivity and frequency response related parameters named  $R_{dir}$  and  $R_{rf}$ . It is tested in highly reverberant conditions. Results of the prediction give a correlation coefficient of 0.96 with a standard deviation of

6%. Simulation of  $R_{dir}$  and  $R_{rf}$  variations leads to modifications of the predicted scores in good agreement with the current knowledge of the influence of directivity and frequency response on speech intelligibility.

## I. INITIAL MODEL

### A. Initial predictor

The concept of useful and detrimental sound energy related to speech intelligibility has been introduced by Lochner and Burger<sup>2</sup> and developed by Bradley.<sup>1</sup> The ratio  $U_\tau$  of useful-to-detrimental energies can be expressed as follows:

$$U_\tau = 10 \log \left[ \frac{R_\tau}{(1 - R_\tau) + 10^{(-S/N)/10}} \right]. \quad (1)$$

$\tau$  is the time limit between early  $E_e$  and late  $E_l$  energy;  $S/N$  is the signal-to-noise ratio in dB(A) (i.e., the difference between speech and noise sound levels);  $R_\tau$  is the ratio between early and total energy:  $R_\tau = E_e / (E_e + E_l)$ .

In Eq. (1) the numerator is the useful energy and the denominator the detrimental energy. The predictor  $U_\tau$  leads to the following third order polynomial equation between speech intelligibility scores SI(%) (using a Fairbanks rhyme test) and  $U_{80}$  in the 1-kHz octave band (Fig. 1):<sup>1</sup>

$$SI(\%) = 1.219 \cdot U_{80} - 0.02466 U_{80}^2 + 0.00295 U_{80}^3 + 95.65. \quad (2)$$

Intelligibility scores have been measured in rooms where reverberation time values vary from 0.8 to 3.8 s.<sup>1</sup>

### B. Modified predictor

A third order polynomial equation is chosen to simply represent the nonlinear variation of speech intelligibility scores as a function of S.T.I.<sup>3</sup> or  $U_{80}$  [Eq. (2)]. But Fletcher for the Articulation Index<sup>9</sup> and later Lochner and Burger<sup>2</sup> have shown that intelligibility variations are described by a “S” form (Fig. 2) adequate with the lowest scores.

In order to follow this “S” curve, predicted scores  $\hat{I}$ (%) are computed by Eq. (3) based on Fletcher and Galt<sup>10</sup> and Dirks *et al.*<sup>11</sup> formulations:

<sup>a)</sup>Present address: Laboratoire d'Acoustique Appliquée, 5 rue Mchal. Joffre, 06400 Cannes; Electronic mail: Laurent.Faiget@wanadoo.fr

<sup>b)</sup>Author to whom correspondence should be addressed; Electronic mail: ruiz@univ-tlse2.fr



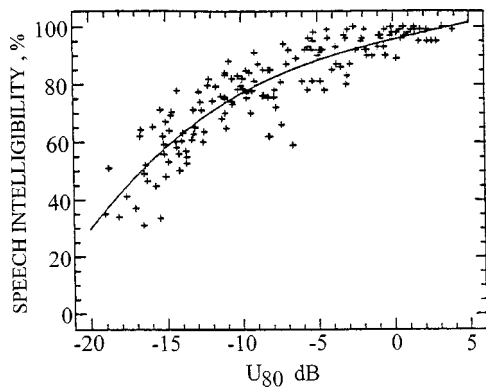


FIG. 1. Measured speech intelligibility scores versus 1-kHz  $U_{80}$  values and best-fit third order polynomial (from Ref. 1).

$$\hat{I}(\%) = 100(1 - 10^{-[(S/N)_{eq} + 40]/60q})^n. \quad (3)$$

$(S/N)_{eq}$  is called the equivalent signal-to-noise ratio and is expressed as follows:

$$(S/N)_{eq} = 10 \log \left( \frac{R_\tau^\alpha}{(1 - R_\tau^\alpha) + 10^{(-S/N)/10}} \right)$$

$$\text{with } R_\tau^\alpha = \frac{\int_0^\tau \alpha(t) h^2(t) dt}{\int_0^\tau h^2(t) dt}. \quad (4)$$

$S/N$  is the signal-to-noise ratio in dB(A);  $\alpha$  is the fraction of the energy of an individual reflection integrated in the useful energy sum (cf. Sec. IC 1),<sup>2,1</sup>  $R_\tau^\alpha$  is the early-to-total energy ratio;  $n$  and  $q$  are the two regression coefficients (instead of four normally used in a third order polynomial regression).  $h(t)$  is the impulse response.

## C. Results

### 1. Rooms and acoustic measurements

Measurements and speech tests have been performed in a reverberant room of 1100 m<sup>3</sup> and in a church of 40 000 m<sup>3</sup>

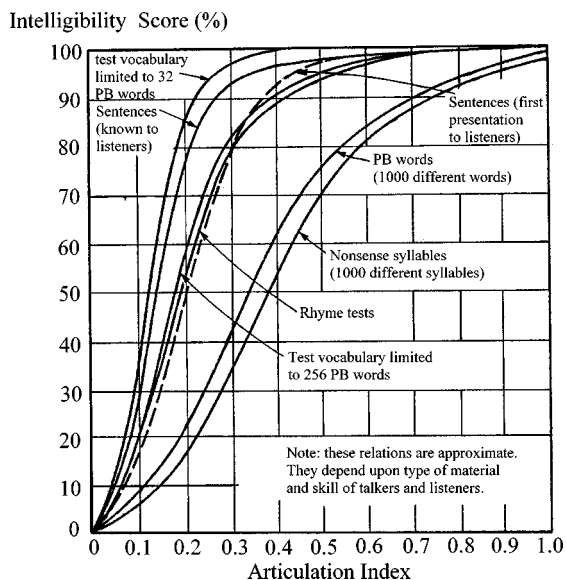


FIG. 2. "S" curves between intelligibility scores and Articulation Index (from Ref. 9).

TABLE I. Mean reverberation time ( $RT_{60}$ ) of the reverberant room and the church A.

$f$ (Hz)	125	250	500	1000	2000	4000	8000
$RT_{60}$ room (s)	10,0	8,8	8,8	8,1	6,9	3,6	1,4
$RT_{60}$ church (s)	14,5	13,4	12,0	10,0	7,3	4,3	1,8

(church A). Reverberation times are reported in Table I. They were measured from the smooth decay curve of the Schroöder integrated impulse response.<sup>12</sup>

Two different loudspeakers have been tested: RB33 and RB90. Their octave band directivity factors  $Q$  are in the Appendix.

Twelve on-axis source–receiver combinations are studied, six in each room. All acoustic measurements are obtained from a Maximum Length Sequence method (M.L.S.) for estimating the Impulse Response (IR). A procedure to determine the real duration of the impulse response  $h(t)$  is applied.<sup>8,13</sup> Indeed, it has been shown that significant errors on energy ratios are obtained if the total time of acquisition is greater than the real duration  $T$  of the impulse response of the system when the measurement is corrupted by extraneous noise. The ratio  $R_\tau^\alpha$  is computed on echogram  $h^2(t)$  where high sound level individual reflections are identified to apply the  $\alpha$  weighting [Eq. (4)]. Lochner and Burger curves (Fig. 10 in Ref. 2) are approximated by the following rule: when the sound level difference between direct sound and individual reflections is greater than 2.5 dB, the 5-dB curve is applied; when the difference is between  $-2.5$  dB and  $+2.5$  dB, the 0-dB curve is used and when the difference is less than  $-2.5$  dB, the  $-5$ -dB curve is applied. This weighting is done for all the samples of the echogram between 0 and 50 ms.

Various sound levels of white noise are emitted by another loudspeaker in order to create signal-to-noise ratios  $S/N$  varying from  $-10$  dB(A) to  $+10$  dB(A) in 5-dB(A) steps at the listeners' positions in the room. These positions are also those of the corresponding impulse response measurements. The experiment leads to 61 acoustical combinations of distances and signal-to-noise ratios.

The choice of the time limit  $\tau$  is based on another set of tests in other large reverberant rooms where 99 different conditions of reverberation, noise, and loudspeakers have been considered.<sup>8</sup> The best correlation coefficient  $r=0.99$  between measured and predicted [Eq. (3)] intelligibility scores and the smallest standard deviation  $\sigma=9.3\%$  are obtained for  $\tau=50$  ms. The variation of  $\tau$  has been sequenced from 10 to 100 ms by steps of 5 ms. As the purpose of this study is also to predict scores in highly reverberant halls, the value of 50 ms has been selected both as time limit of the  $\alpha$  weighting and of the useful energy.  $R_{50}$  becomes the  $D_{50}$  of Thiele;<sup>14</sup> corrections derived from Lochner and Burger curves are applied to individual reflections between 0 and 50 ms according to the preceding rule. Predictor  $(S/N)_{eq}$  is obtained by Eq. (5):

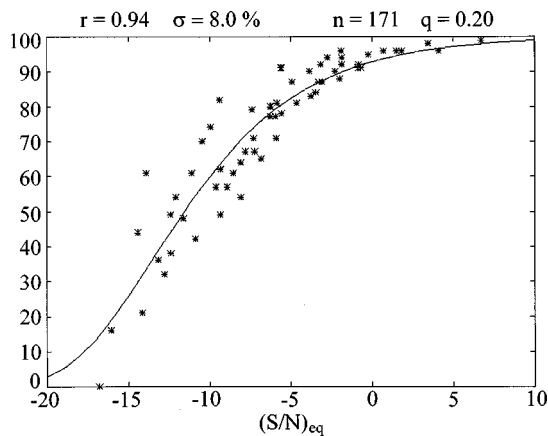


FIG. 3. Measured speech intelligibility scores versus  $(S/N)_{eq}$  corresponding values [Eq. (5)] and best least-squares fit based on the model of Eq. (3).

$$(S/N)_{eq} = 10 \log \left( \frac{D_{50}^\alpha}{(1 - D_{50}^\alpha) + 10^{-(S/N)/10}} \right)$$

$$\text{with } D_{50}^\alpha = \frac{\int_0^{50} \alpha(t) h^2(t) dt}{\int_0^T h^2(t) dt}. \quad (5)$$

## 2. Intelligibility test

The intelligibility test uses 10 phonetically balanced lists of 34 triphonemic French words (mono or dissyllabic). The speaking rate is about nine phonemes per second. Every word is preceded by a sentence without any semantic relation with the word to be recognized. A trial list is proposed. The subjects have to complete a form indicating the vowel(s), consonant(s), syllable(s), or word heard at the end of the sentence. They were approximately 25-year-old students without any auditory problem. The obtained score is the percentage of correctly recognized phonemes.

## 3. Accuracy of the model

Figure 3 plots the results of the 61 speech intelligibility scores versus  $(S/N)_{eq}$  corresponding values. The curve is a regression of the form of Eq. (3). The least-squares fit is obtained for  $n=171$  and  $q=0.20$ . The correlation coefficient  $r$  is then equal to 0.94 and the standard deviation is 8.0%. Other forms of regressions have been tested such as sigmoidal functions. The accuracy of this model has not been improved in terms of correlation coefficient and standard deviation.

## II. LIMITS OF THE MODEL

The aim of this paper is to identify and quantify separately the room and loudspeaker influence on intelligibility scores. In order to prove the necessity of a predictor depending on loudspeaker features and to evaluate the limits of the preceding global model of intelligibility, some experimental results have to be considered (Tables II and III).

First, Table II shows that measured scores  $I\%$  are better predicted from the model for impulse responses obtained with the RB33 source. The higher the ratio  $D_{50}^\alpha$  is, the more directive the source will be. In very noisy conditions, the

TABLE II. Comparison between measured  $I(\%)$  and predicted  $\hat{I}(\%)$  scores for the RB33 and the RB90 loudspeakers in the reverberant room at the distance 4 m for three signal-to-noise ratios in dB(A) (values in brackets are associated standard deviations).

Reverberant room	$(S/N)_{dB(A)}$	$(S/N)_{eq}$	$I(\%)$	$\hat{I}(\%)$
RB33 ( $\alpha \cdot D_{50} = 22,7\%$ )	-10	-16	16,2 (11,6)	17,9
	-5	-11,6	48 (4,8)	47,9
	0	-8,1	63,7 (11,7)	69,7
RB90 ( $\alpha \cdot D_{50} = 43,1\%$ )	-10	-13,9	61,0 (2,7)	31,8
	-5	-9,4	81,6 (4,8)	61,7
	0	-5,6	90,7 (2,6)	79,2

RB90 leads to real scores greater than those obtained with the RB33 in the same source-receiver position. The same is true for predicted scores but with less accuracy. The increase of definition  $D_{50}^\alpha$  improves both measured and estimated scores but in a different way. Therefore, directivity influence must be better accounted for by the model.

Second, measured scores differ in two source-receiver positions where loudspeaker, signal-to-noise ratio, and definition  $D_{50}^\alpha$  remain the same (Table III). The difference is greater for low S/N ratios. But the predicted corresponding scores will be the same in the two rooms for the same S/N ratio. Indeed, the  $(S/N)_{eq}$  predictor only depends on these parameters [Eq. (5)]. This result implies that both  $D_{50}^\alpha$  and S/N ratios are not sufficient to predict intelligibility scores. Source-receiver position in the hall has an effect on measured scores, whereas parameters of the model do not act in the same way.

Finally, room and loudspeaker influence should be considered separately. A loudspeaker parameter related to its radiation in the room should be introduced to improve prediction particularly for high reverberation time values and low signal-to-noise ratios.

## III. ROOM AND LOUDSPEAKER INFLUENCES

The general form of regression is the same as in Eq. (3). The objective is to introduce room and loudspeaker influence in the  $(S/N)_{eq}$  predictor of Eq. (5). It is necessary to identify and separate their respective contribution in the echogram.

The prediction is based on a measurement of the impulse response by an M.L.S. technique in a given source-microphone position in the room. This impulse response  $h(t)$  is defined by the following convolution equation:

TABLE III. Comparison between measured intelligibility scores  $I(\%)$  for various signal-to-noise ratios but for the same value of Definition  $D_{50}^\alpha$  (values into brackets are associated standard deviations).

$(S/N)_{dB(A)}$	$I(\%)$ Reverberant room 4 m	$I(\%)$ Church A 16 m
-5	81,6 (4,8)	62,0 (4,0)
0	90,7 (2,6)	78,2 (3,3)
5	91,4 (2,2)	86,8 (5,9)
10	96,1 (2,1)	91,7 (1,3)

$$h(t) = h_{hp}(t) * h_s(t) + n(t). \quad (6)$$

$h_{hp}(t)$  is the impulse response of the loudspeaker;  $h_s(t)$  is the impulse response of the room;  $n(t)$  is the sum of the acoustic and electrical (or computation) noises.

A procedure of deconvolution is applied to obtain  $h_s(t)$ . It consists in finding the inverse filter  $f(t)$  which verifies  $f(t) * h(t) = h_s(t)$  by inverting the module and taking the opposite of the phase of  $H_{hp}(f)$ , Fourier transform of  $h_{hp}(t)$ .<sup>15</sup> The deconvolution procedure uses the axial impulse response of the loudspeaker. It is measured beforehand in an anechoic chamber or by the near-field/far-field technique.<sup>16</sup> The complex multiplication  $F(f) \cdot H(f)$  is equal to  $H_s(f)$  and an inverse Fourier transform leads to  $h_s(t)$ .

But the main assumption of the method is the absence of noise. It must be removed before deconvolution by averaging impulse responses after each Maximum Length Sequence emission.

When this deconvolution procedure is achieved, the room and loudspeaker influence on the speech intelligibility predictor can be studied from  $h_s(t)$ ,  $h(t)$ , and  $h_{hp}(t)$ .

### A. Room influence: $D_{50}^s$

The concept of useful and detrimental energies is applied to  $h_s(t)$  to define room influence. The proposed ratio  $D_{50}^s$  is the same as definition  $D_{50}$  but is computed on the deconvolved impulse response  $h_s(t)$  and not on the global one  $h(t)$ :

$$D_{50}^s = \frac{\int_0^{50} h_s^2(t) dt}{\int_0^T h_s^2(t) dt}. \quad (7)$$

$T$  is the total time of acquisition and 0 is the time of direct sound arrival measured on  $h(t)$ . Integration onto the total time  $T$  does not introduce errors because impulse response is noiseless.<sup>13</sup> It has been demonstrated that, without noise, the values of the energy ratios are close to equal even if the total duration of acquisition changes.<sup>8</sup>

If  $E_s$  is the energy of speech, useful  $E_u$  and detrimental  $E_d$  ones are, respectively:

$$E_u = D_{50}^s E_s, \quad (8)$$

$$E_d = (1 - D_{50}^s) E_s. \quad (9)$$

The ratio  $D_{50}^s$  represents the energetic contribution of reflections in the first 50 ms after direct sound arrival to the entire energy of reflections in the room. The concept is the same as Bradley but applied to the deconvolved impulse response.

### B. Loudspeaker influence

Loudspeaker influence on speech intelligibility is studied from its on-axis impulse response measurement. The assumption is that all speech energy emitted by the loudspeaker is useful to intelligibility contrary to the room influence in which reverberant energy can act as a noise to degrade speech perception. Therefore, the corresponding features will be in the numerator of the  $(S/N)_{eq}$  predictor.

TABLE IV. Values of  $D_{50}$ ,  $D_{50}^s$ , and  $R_{dir}$  for RB33 and RB90 loudspeakers at the same positions in the two halls.

		Distance (m)	$D_{50}$ (%)	$D_{50}^s$ (%)	$R_{dir}$
Reverberant room	RB33	2	51	29	1,8
		4	27	19	1,4
		6	23	18	1,3
	RB90	2	59	9	6,6
		4	43	6	7,2
		6	38	5	7,6
Church A	RB33	2	79	8	9,9
		8	41	14	2,9
	RB90	16	21	17	1,2
		2	80	7	11,4
		8	64	8	8,0
		16	44	9	4,9

### 1. $R_{dir}$ factor

The  $R_{dir}$  factor is defined by the following ratio.

$$R_{dir} = \frac{D_{50}}{D_{50}^s} = \frac{\int_0^{50} h^2(t) dt / \int_0^T h^2(t) dt}{\int_0^{50} h_s^2(t) dt / \int_0^T h_s^2(t) dt}. \quad (10)$$

$R_{dir}$  is linked to the loudspeaker radiation in the room. In an ideal reflecting room with an ideal omnidirectional source, the amplitude of the first reflections is nearly the same as that of the direct signal and  $R_{dir}$  reaches 1. If a very directive source is used whose the main axis of radiation is in the direction of the microphone,  $D_{50}^s$  becomes lower than  $D_{50}$  which tends to one, and  $R_{dir}$  is greater. In an ideal anechoic room,  $D_{50}$  is equal to one because the duration of  $h(t)$  is less than 50 ms [ $h(t) = h_{hp}(t)$ ].  $D_{50}^s$  is also equal to 1 because  $h_s(t)$  is equivalent to a Dirac impulsion; therefore  $R_{dir} = 1$ .

The higher the directivity of a loudspeaker, the higher  $R_{dir}$  for a measurement of an impulse response at the same source–receiver location in the same room but for different loudspeakers.

By examining the results of Table IV, for the same loudspeaker but for the two halls, it can be noted that  $R_{dir}$  decreases when the source–receiver distance increases except for the reverberant room with the RB90 loudspeaker. This remark shows that  $R_{dir}$  is not a characteristic specific to the loudspeaker directivity: it is related to its radiation in the room. Both in the two halls, the values of  $D_{50}$  decrease with distance, thus indicating a higher energy in the latter part of the impulse response and/or a lower one in the earlier part.

The reverberant sound field becomes more important and the direct sound level diminishes. For  $D_{50}^s$  the variation

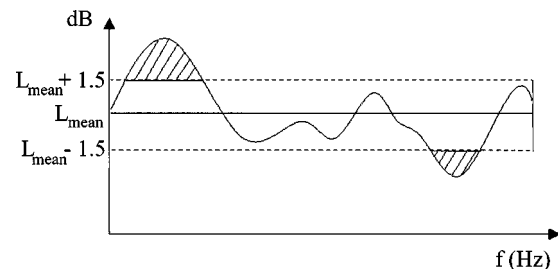


FIG. 4. Simulation of a loudspeaker frequency response with up and down limits for the computation of  $R_{dir}$  ( $\pm 1,5$  dB from mean sound level  $L_{mean}$ ).

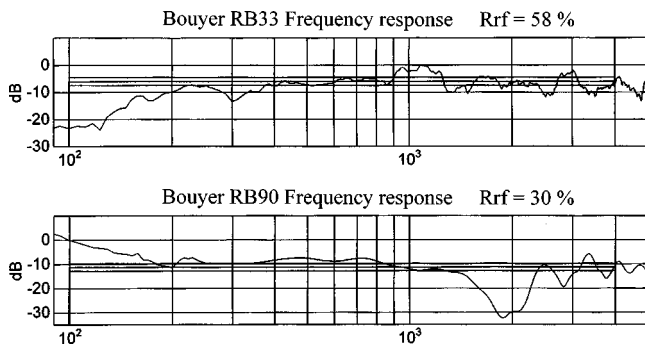


FIG. 5. Frequency responses of the two loudspeakers used in the experiments with up and down limits for the computation of  $R_{rf}$  ( $\pm 1.5$  dB from mean sound level).

is different in the two rooms. The deconvolution of loudspeaker impulse response shows that  $D_{50}^s$  increases with distance in the church for the two sources tested. This is explained by the number of reflections in the first 50 ms which grows with distance in such a long room. The result is not observed in the reverberant room because its volume is smaller. The sense of variation of  $D_{50}$  and  $D_{50}^s$  with distance can be different which leads also to a variable one for  $R_{dir}$ .

## 2. $R_{rf}$ factor

The influence of bandwidth on intelligibility has been studied in telephony applications. Experiments with high- and low-pass filtered speech and masking white noise have shown that the band 300–4000 Hz is sufficient to ensure intelligibility syllable scores greater than 90%.<sup>17</sup> In other similar experiments with nonsense CVC word lists, the score was greater than 97% for the band 100–4000 Hz.<sup>10</sup> Considering that this band is sufficient for a good intelligibility, it is necessary to ensure a flat frequency response of the system in order to reproduce speech without any alteration. Bucklein has measured resonance and antiresonance influence on intelligibility scores.<sup>18</sup> The influence exists, for example, a 25-dB amplitude resonance in the band 1000–2000 Hz leads

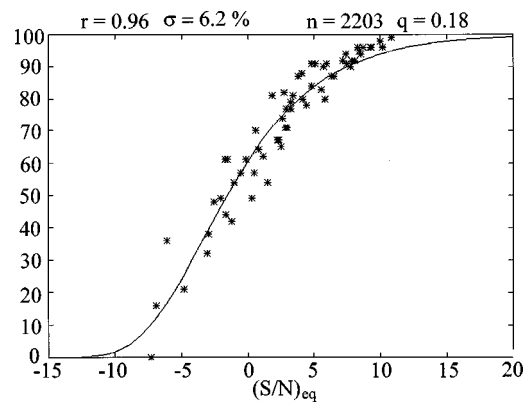


FIG. 6. Measured speech intelligibility scores versus  $(S/N)_{eq}$  predictor corresponding values [Eq. (3)] and best least-squares fit based on the model of Eq. (3).

to a 4% decrease of the score. An equivalent antiresonance only leads to a 1% decrease. He concludes that resonances have a greater detrimental influence than antiresonances and that the frequency response of a system can accept irregularities but with as few wide bandwidth resonances of high level as possible.

Here the purpose is to choose an amplitude tolerance in the band 100–4000 Hz and to define a feature to quantify the frequency response shape on intelligibility scores. The tolerance is fixed at  $\pm 1.5$  dB in order to be more restrictive than the usual  $\pm 3$ -dB audio tolerance (but on a larger bandwidth), and than the imperceptible 2-dB resonances on frequency response.<sup>19</sup>

The criteria  $R_{rf}$  which takes into account frequency response fluctuations in the 100–4000 Hz band is defined by the following relation:

$$R_{rf} = \frac{E_{hp} - E_{n, hp}}{E_{hp}} = 1 - \frac{E_{n, hp}}{E_{hp}}. \quad (11)$$

Quantities  $E_{hp}$ ,  $E_{n, hp}$  are obtained when the tolerance  $\pm 1.5$  dB is applied on the loudspeaker frequency response (Fig. 4).  $E_{hp}$  is the energy of the frequency response in the band

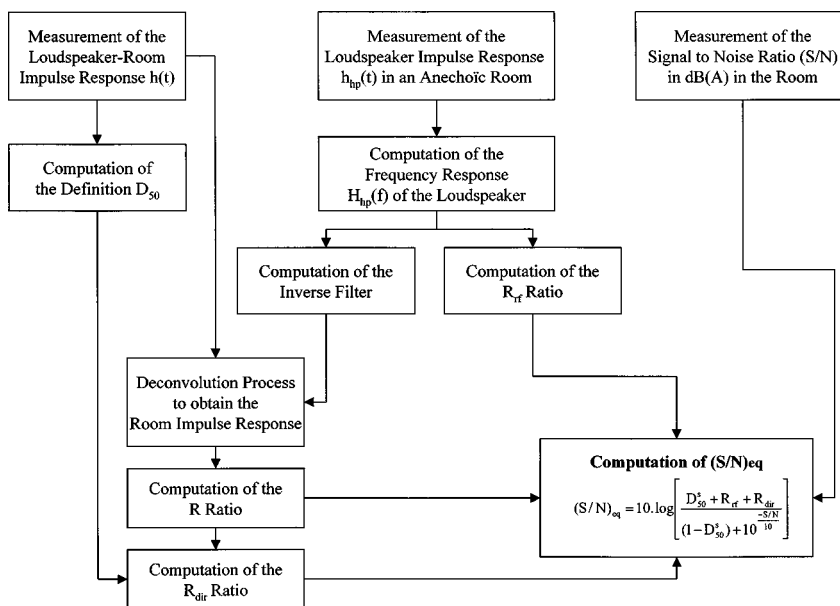


FIG. 7. Method of  $(S/N)_{eq}$  computation.

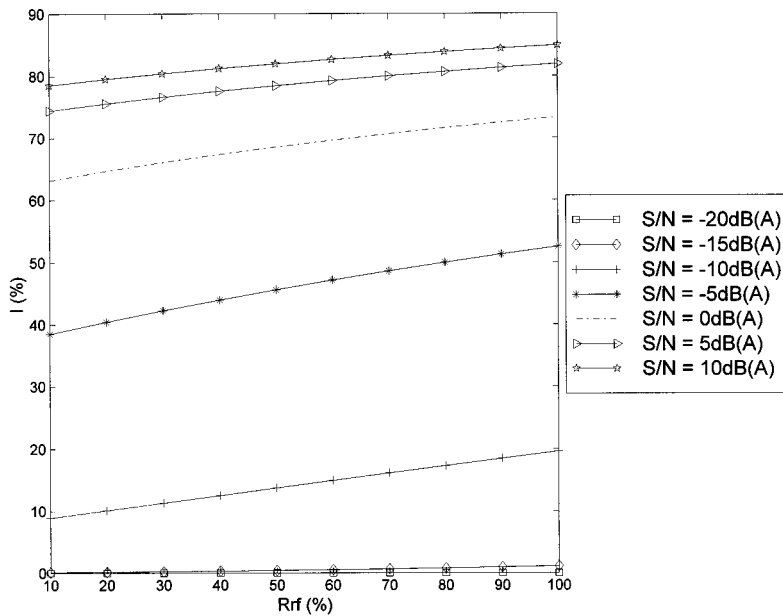


FIG. 8. Predicted intelligibility score versus  $R_{rf}$  for different signal-to-noise ratios (S/N). Values are computed from the model [Eqs. (13) and (14)] with  $R_{dir} = 1.4$  and  $D_{50}^s = 20\%$ .

100–4000 Hz. All of the energy above and under limits of the tolerance are summed up to give  $E_{n,hp}$ . In the case of antiresonances, even if energy is not present under the lower limit, this missing quantity is added to  $E_{n,hp}$  like energy of the resonances above the upper limit. Therefore, if  $E_p$  is the speech energy, the useful energy transmitted by the loudspeaker is the product  $R_{rf}E_p$ . When the frequency response is within the tolerance,  $E_{n,hp}$  is null and 100% of the energy is useful. Figure 5 shows two loudspeaker frequency responses with their respective  $R_{rf}$  values.

#### IV. MODEL INCLUDING SEPARATED INFLUENCE OF LOUDSPEAKER AND ROOM

The new model equation is given by Eq. (3). The equivalent signal-to-noise ratio  $(S/N)_{eq}$  is:

$$(S/N)_{eq} = 10 \log \left( \frac{D_{50}^s E_s + R_{rf} E_s + R_{dir} E_s}{(1 - D_{50}^s) E_s + E_n} \right). \quad (12)$$

The numerator sums the useful parts of speech energy  $E_s$  and the denominator the detrimental ones ( $E_n$  is the noise energy). After simplification, Eq. (12) becomes:

$$(S/N)_{eq} = 10 \log \left( \frac{D_{50}^s + R_{rf} + R_{dir}}{(1 - D_{50}^s) + 10^{-(S/N)/10}} \right). \quad (13)$$

$S/N$  is the signal-to-noise ratio in dB(A).

Figure 6 plots the least-squares regression line between the 61 measured scores (cf. Sec. I C 1) and the predicted ones by the form of Eq. (3) with  $(S/N)_{eq}$  given by Eq. (13). Parameters  $n$  and  $q$  are, respectively, equal to 2203 and 0.18.

The equation of the model is the following:

$$I(\%) = 100(1 - 10^{-[(S/N)_{eq} + 40]/(60 \times 0.18)})^{2203}. \quad (14)$$

The coefficient of determination  $r^2$  is equal to 0.92, which means that 92% of the total variance is explained by the regression giving a correlation coefficient  $r$  of 0.96.<sup>20</sup> Standard deviation is equal to 6.2%.

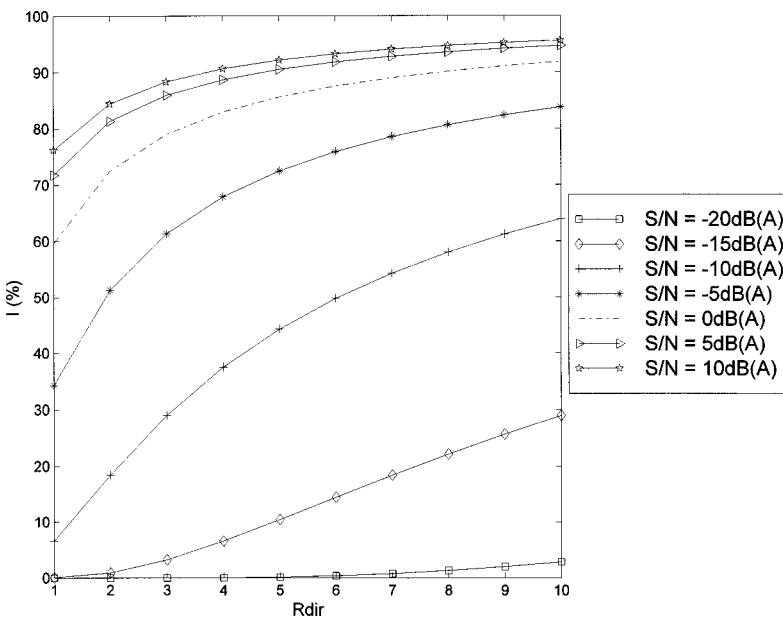


FIG. 9. Predicted intelligibility score versus  $R_{dir}$  for different signal-to-noise ratios (S/N). Values are computed from the model [Eqs. (13) and (14)] with  $R_{rf} = 30\%$  and  $D_{50}^s = 20\%$ .

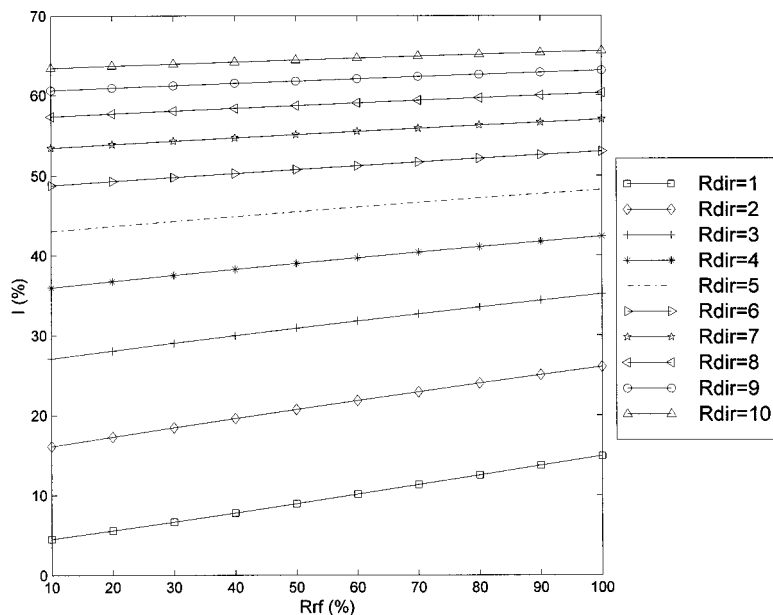


FIG. 10. Predicted intelligibility score versus  $R_{rf}$  for different  $R_{dir}$ . Values are computed from the model [Eqs. (13) and (14)] with  $(S/N) = -10$  dB(A) and  $D_{50}^s = 20\%$ .

It appears that the separation of room and loudspeaker influence on intelligibility scores estimation is obtained without a decrease of the quality of the estimation. The accuracy of the model is slightly better comparing with the results of the first predictor (Fig. 3).

### A. Computation of the equivalent signal-to-noise ratio $(S/N)_{eq}$

Figure 7 illustrates the procedure of  $(S/N)_{eq}$  computation. Three measurements are necessary. The first one is the estimation of loudspeaker impulse response on its main axis of radiation in an anechoic room or by a near-field/far-field technique. The room-loudspeaker impulse response  $h(t)$  and the signal-to-noise ratio depend on the room tested.

### B. Properties of the model

Properties of the model are obtained from Eq. (14) with  $(S/N)_{eq}$  given by Eq. (13). To show the respective influence of  $R_{rf}$  and  $R_{dir}$ , the variations of predicted intelligibility scores are plotted for constant values of the other parameters ( $D_{50}^s$ ,  $S/N$ , and, respectively,  $R_{dir}$  and  $R_{rf}$ ). Charts of Figs. 8–11 are then obtained. Values of  $S/N$  vary from a noisy situation  $[-20$  dB(A)] to a comfortable one  $[+10$  dB(A)] by steps of 5 dB(A).  $R_{rf}$  goes from 10% to 100%, i.e., from a very irregular loudspeaker frequency response to a perfectly flat one. Variations of  $R_{dir}$  are chosen from 1 to 10. One is representative of an omnidirectional sound source in a reverberant room and ten is representative of a more directive one in a more absorbing enclosure. Figures 8 and 9 show that the improvement of scores does not vary linearly versus  $S/N$ . For a given value of  $R_{rf}$  or  $R_{dir}$ , a  $S/N$  variation of 5 dB(A)

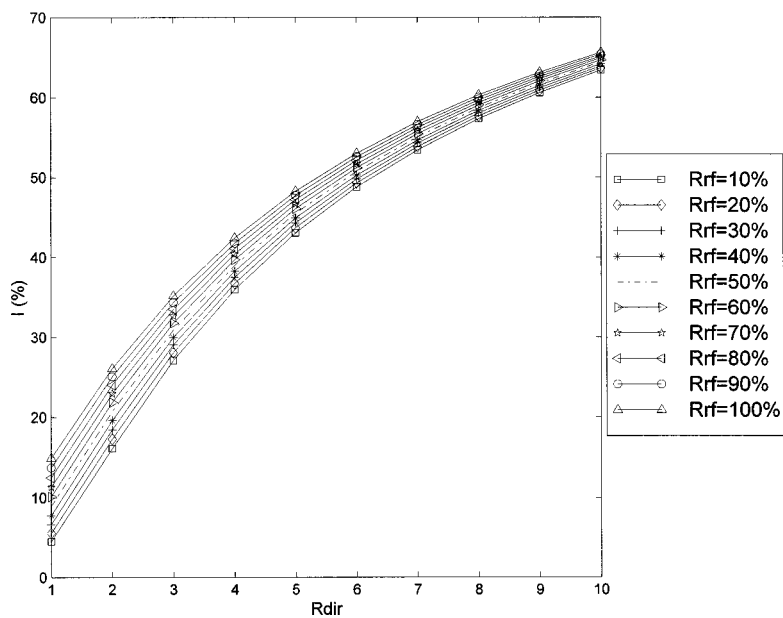


FIG. 11. Predicted intelligibility score versus  $R_{dir}$  for different  $R_{rf}$ . Values are computed from the model [Eqs. (13) and (14)] with  $(S/N) = -10$  dB(A) and  $D_{50}^s = 20\%$ .

TABLE V. Mean reverberation time ( $RT_{60}$ ) of church B in which accuracy of the model has been tested.

$f$ (Hz)	125	250	500	1000	2000	4000
$RT_{60}$ (s)	5,8	6,7	7,1	6,2	5,3	3,8

between two negative values has a greater influence than between two positive ones. For example, if S/N is modified from  $-10$  dB(A) to  $-5$  dB(A), the gain is about 30% (Fig. 8), but if this modification occurs from 5 to 10 dB(A) the gain is less than 5%. Examination of Fig. 8 reveals that the greater improvement of score is approximately 10% for  $S/N = -10$  dB(A) when  $R_{rf}$  varies from 10% to 100%. The effect of a regular frequency response is important when S/N is low.

For  $S/N = +10$  dB(A) improvement is about 6% between  $R_{rf} = 10\%$  and 100%. Such a result is identical to Bucklein one. He has shown that a great degradation of frequency response can induce a 4% decrease of scores (with phoneme lists) in situations where masking noise has no influence.<sup>20</sup>

Influence of  $R_{dir}$  is shown in Fig. 9. The effect of varying  $R_{dir}$  is more important than the  $R_{rf}$  one. The gain is about 56% when  $R_{dir}$  varies from 1 to 10 for an  $S/N = -10$  dB(A). This improvement is smaller when signal-to-noise ratio is greater [20% for  $S/N = +10$  dB(A)].

When signal-to-noise ratio is too low [ $-15$  or  $-20$  dB(A)], influence of  $R_{dir}$  or  $R_{rf}$  is not as important as for other negative S/N. The model shows that when noise sound level becomes too high, scores cannot be improved even with directive and/or flat frequency response loudspeakers. A substantial gain of score is obtained when S/N becomes greater than or equal to  $-10$  dB(A).

As shown in Fig. 10, for a given negative value of S/N ratio, a significant enhancement of scores can be obtained by increasing directivity whatever the value of  $R_{rf}$ . Figure 11 makes it clear that the improvement of scores is smaller when the frequency response is flattened for the same negative S/N ratio and whatever  $R_{dir}$  is.

When noise is predominant, the directivity influence is greater. In a noisy ambience, the use of a highly directive source is recommended for a good intelligibility. But in a hall with a positive S/N ratio, the use of a directive loud-

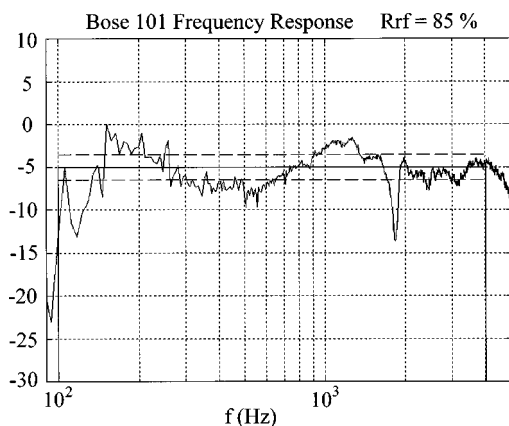


FIG. 12. Bose 101 loudspeaker frequency response with limits for the computation of  $R_{rf}$ .

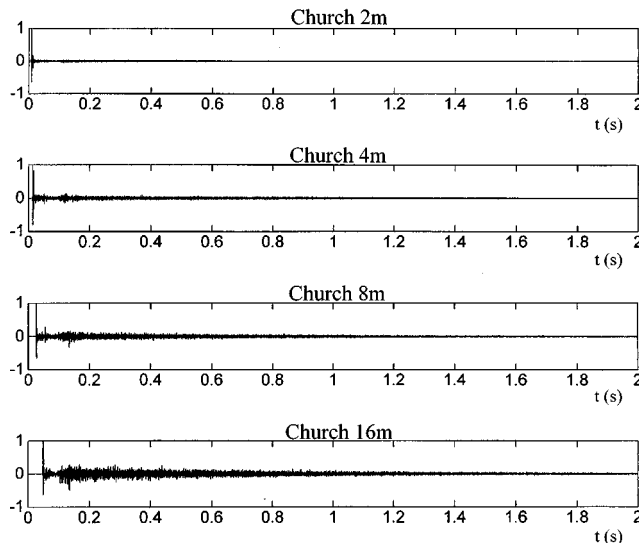


FIG. 13. Impulse responses in the church at 2, 4, 8, and 16 m from the Bose 101 loudspeaker.

speaker is less necessary. Such conclusions are identical to those of Jacob in his experiments with sources of different directivity factors.<sup>6</sup>

All of the plots have been obtained for a  $D_{50}^s = 20\%$ . When other values of  $D_{50}^s$  are chosen, the conclusions about  $R_{dir}$  and  $R_{rf}$  effects on scores are the same.

Properties of the model are in good agreement with known loudspeaker and room influence on intelligibility scores.

### C. Model accuracy

An experiment is done to test the accuracy of the model. The hall is another empty church of  $12\,000\text{ m}^3$  ( $40\text{ m} \times 15\text{ m} \times 20\text{ m}$ ) (church B). A new loudspeaker (Bose 101) is placed on the altar and four points on axis are chosen at distances of 2 m, 4 m, 8 m, and 16 m from the source. Beforehand, reverberation time has been measured and averaged at this points (Table V).

Measurements of on-axis frequency response of the Bose 101 loudspeaker in an anechoic room have given a value of 85% for  $R_{rf}$  (Fig. 12). Figure 13 shows impulse responses at the chosen points in church B before deconvolution procedure. Results of the computation of  $D_{50}^s$  and  $R_{dir}$  are reported in Table VI. With these values  $(S/N)_{eq}$  is computed [Eq. (13)], and predicted scores [Eq. (14)] are indicated in Table VII for the corresponding signal-to-noise ratios.

TABLE VI. Values of  $D_{50}$ ,  $D_{50}^s$ ,  $R_{dir}$  in church B for the loudspeaker used ( $R_{rf} = 85\%$ ).

Distance (m)	$D_{50}(\%)$	$D_{50}^s$	$R_{dir}$	$R_{rf}(\%)$
2	79	52	1,5	85
4	56	31	1,8	85
8	34	22	1,5	85
16	13	11	1,2	85

TABLE VII. Predicted  $\hat{I}(\%)$  and measured  $I(\%)$  scores for different signal-to-noise ratios (S/N) at 2, 4, 8, and 16 m from the source in church B.

Distance (m)	S/N (dB(A))	$\hat{I}(\%)$	$I(\%)$
2	3	83	84
	0	76	69
	-5	54	48
4	0	74	85
	-2	68	72
	-6	47	49
8	0	70	78
	-3	59	63
	-8	29	36
16	0	64	58
	-3	52	37
	-8	22	21

Intelligibility tests (cf. Sec. IC 2) have been performed at the same points as impulse response measurements with 12 subjects divided in 4 groups of 3, each group at one of the four distances. White masking noise is emitted by another loudspeaker in the same vertical plane as the one used for the lists. The sound level of emission is chosen in order to satisfy the required S/N ratios. The sound pressure level of the speech lists is 70 dB(A).

Scores are averaged for each group of listeners and Table VII is a comparison between the predicted  $\hat{I}(\%)$  and measured  $I(\%)$  scores. The mean absolute difference is 6%. Prediction is in good agreement with the measurements.

## V. CONCLUSION

All of the speech tests carried out in order to build the score database have been performed with listeners in the main radiating axis of the loudspeaker. The procedure of deconvolution to obtain the room impulse response uses an on-axis measurement of the loudspeaker impulse response. Therefore, impulse response measurements in the hall are always done in this axis. Scores are predicted by the model in this particular but essential direction of propagation. It is the main actual limitation of the model. The prediction in other directions would require loudspeaker impulse response measurements in these angles of radiation. It would be necessary to build a new database of speech intelligibility tests for various directions, signal-to-noise ratios, and reverberation situations.

A second limitation of the model results in the use of a single loudspeaker for the database and the prediction. The model is not adapted to a multi-loudspeaker sound reinforcement system because in the first 50 ms of the echogram nearby loudspeakers can have a detrimental influence. Deconvolution becomes more complex. When loudspeakers are distant from each other more than 17 m, on-axis prediction acts as if close loudspeaker contributions belong to the detrimental part of the sound field. In order to examine useful or detrimental role of direct sound issued from close loudspeakers, complementary experiments are needed.

Type and spatial position of noise in the hall play an important role in speech intelligibility. All of the experiments have been performed with a wide-band white noise

source (20–20 000 Hz) in the same vertical plane as the speech loudspeaker. Weighting coefficients depending on detrimental influence of position(s) and spectral (or time) properties of masking noise(s) could be introduced in the predictor.

Simulations of the  $R_{dir}$  and  $R_{rf}$  influence on score have shown that when the signal-to-noise ratio is too small, the enhancement is difficult. But in real cases, an increase of high frequency sound levels can improve intelligibility. Such a kind of modification of the source is not taken into account by the model and could also be included after complementary studies.

Another limitation concerns voice quality. In related security applications where recorded messages are not used but a “natural voice” must speak, quality of the voice is very important not only for message recognition but also for the emotional content transmitted. Indeed, listener reaction can depend on its perception of speaker emotion. Some acoustic modifications of vowels and consonants are measured when a speaker is under stress,<sup>21</sup> but they are not taken into account in the speech intelligibility models.

Finally, even if these points limit the application of the prediction for the moment, this model is a basis for future development and now includes a separation of room, loudspeaker and masking noise influence.

## APPENDIX

TABLE AI. Octave band directivity factors  $Q$  for the three loudspeakers used (RB33, RB90, Bose 101).

$f$ (Hz)	125	250	500	1000	2000	4000
$Q_{RB33}$	1,1	1,2	1,6	2,1	2,7	4,1
$Q_{RB90}$	1,7	1,7	2,5	2,5	2,8	4,9
$Q_{Bose101}$	1,2	1,3	2,0	3,2	3,8	6,4

<sup>1</sup>S. Bradley, “Predictors of speech intelligibility in rooms,” J. Acoust. Soc. Am. **80**, 837–845 (1986).

<sup>2</sup>J. P. A. Lochner and J. F. Burger, “The influence of reflections on auditorium acoustics,” J. Sound Vib. **1**, 426–454 (1964).

<sup>3</sup>H. J. M. Steeneken and T. Houtgast, “A physical method for measuring speech-transmission quality,” J. Acoust. Soc. Am. **67**, 318–326 (1980).

<sup>4</sup>V. M. A. Peutz, “Articulation loss of consonants as a criterion for speech-transmission in a room,” J. Audio Eng. Soc. **19**, 915–919 (1971).

<sup>5</sup>W. Klein, “Articulation loss of consonants as a basis for the design and judgment of sound reinforcement systems,” J. Audio Eng. Soc. **19**, 920–922 (1971).

<sup>6</sup>K. D. Jacob, “Subjective and predictive measures of speech intelligibility—the role of loudspeaker directivity,” J. Audio Eng. Soc. **33**, 950–955 (1985).

<sup>7</sup>K. D. Jacob, “Correlation of speech intelligibility tests in reverberant rooms with three predictive algorithms,” J. Audio Eng. Soc. **37**, 1020–1029 (1989).

<sup>8</sup>L. Faiget, “Séparation de l’Influence du Local et de l’Enceinte pour la Prédiction de l’Intelligibilité dans des Conditions d’Écoute Difficiles,” Thèse de Doctorat de l’Université Toulouse III, Toulouse, France (1997).

<sup>9</sup>ANSI S315-1969, “American National Standard methods for the computation of the Articulation Index” (ANSI, New York, 1969).

<sup>10</sup>H. Fletcher and R. H. Galt, “The perception of speech and its relation to telephony,” J. Acoust. Soc. Am. **22**, 89–151 (1950).

<sup>11</sup>D. Dirks, T. S. Bell, R. N. Rossman, and G. E. Kincaid, “Articulation index predictions of contextually dependent words,” J. Acoust. Soc. Am. **80**, 82–92 (1986).



- <sup>12</sup>M. R. Schroöder, "New method of measuring reverberation time," J. Acoust. Soc. Am. **37**, 409–412 (1965).
- <sup>13</sup>L. Faiget, C. Legros, and R. Ruiz, "Optimization of the impulse response length: Application to noisy and highly reverberant rooms," J. Audio Eng. Soc. **46**, 741–750 (1998).
- <sup>14</sup>R. Thiele, "Richtungsverteilung und Zeitfolge der Schallruckwürfe in Raumen," *Acustica* **3**, 291–302 (1953).
- <sup>15</sup>J. Berkhout, D. De Vries, and M. M. Boone, "A new method to acquire impulse responses in concert halls," J. Acoust. Soc. Am. **68**, 179–183 (1980).
- <sup>16</sup>C. Struck and S. Temme, "Simulated free field measurements," J. Audio Eng. Soc. **42**, 467–482 (1994).
- <sup>17</sup>I. Pollack, "Effects of high pass and low pass filtering on the intelligibility of speech in noise," J. Acoust. Soc. Am. **20**, 259–266 (1948).
- <sup>18</sup>R. Bucklein, "The audibility of frequency response irregularities" J. Audio Eng. Soc. **29**, 126–131 (1981).
- <sup>19</sup>D. E. L. Shorter, "A survey of performance criteria and design considerations for high-quality monitoring loudspeakers," J. Acoust. Soc. Am. **2**, 56–71 (1958).
- <sup>20</sup>T. H. Wonnacott and R. J. Wonnacott, *Statistique* (Wiley, New York, 1995).
- <sup>21</sup>R. Ruiz, E. Absil, B. Harmegnies, C. Legros, and D. Poch, "Time- and spectrum-related variabilities in stressed speech under laboratory and real conditions," Speech Commun. **20**, 111–129 (1996).

# On direction finding of an emitting source from time delays

Baruch Berdugo

*The Julius Silver Institute of Biomedical Engineering, Technion—IIT, Haifa 32000, Israel*

Miriam A. Doron

*RAFAEL, Dep. 23, Haifa 31021, Israel*

Judith Rosenhouse

*The Department of General Studies, Technion—IIT, Haifa 32000, Israel*

Haim Azhari<sup>a)</sup>

*The Julius Silver Institute of Biomedical Engineering, Technion—IIT, Haifa 32000, Israel*

(Received 28 February 1998; revised 5 March 1999; accepted 18 March 1999)

This paper presents a statistically and computationally efficient algorithm for direction finding of a single far-field source using a multi-sensor array. The algorithm extracts the azimuth and elevation angles directly from the estimated time delays between the array elements. Hence, it is referred to herein as the time delay direction finding (TDDF) algorithm. An asymptotic performance analysis, using a small error assumption, is conducted. For any 1-D and 2-D array configurations, it is shown that the TDDF algorithm achieves the Cramer Rao lower bound (CRLB) for the azimuth and elevation estimates provided that the noise is Gaussian and spatially uncorrected and that the time delay estimator achieves the CRLB as well. Moreover, with the suggested algorithm no constraints on the array geometry are required. For the general 3-D case the algorithm does not achieve the CRLB for a general array. However, it is shown that for array geometries which obey certain constraints the CRLB is achieved as well. The TDDF algorithm offers several advantages over the beamforming approach. First, it is more efficient in terms of computational load. Second, the azimuth estimator does not require the *a priori* knowledge of the wave propagation velocity. Third, the TDDF algorithm is suitable for applications where the arrival time is the only measured input, in contrast to the beamformer, which is not applicable in this case. © 1999 Acoustical Society of America. [S0001-4966(99)05106-1]

PACS numbers: 43.60.Gk, 43.60.Cg [JCB]

## INTRODUCTION

In various applications of array signal processing such as radar, sonar, and seismology, there is a great interest in detection and localization of wideband sources.<sup>1</sup> The problem of estimating the direction of arrival (DOA) of wideband sources using a sensor array has been studied extensively in the literature.<sup>2–14</sup> A common approach<sup>2–7</sup> to this problem, for a single source scenario, is to use the time delay estimation between two sensors to determine the DOA. Many techniques for estimating the travel time delay between *two* receiving sensors have been investigated, see, e.g., Refs. 2–7. For the *single* source and a *multi-sensor* case, Hahn and Tretter<sup>8</sup> introduced the maximum likelihood (ML) delay-vector estimator. The ML DOA estimators for the multi-sensor and multi-source cases have also been studied extensively.<sup>12–14</sup>

It is well known<sup>11</sup> that the ML DOA estimator, for the single-source case with a spatially uncorrelated noise, can be realized as a focused beamformer. In this paper an alternative approach is proposed, in which the DOA is extracted directly from the estimated time delays between the array elements (referred to as the time delay vector). This approach is an extension to the multi-sensor case of the work in Refs. 10

and 11, where the DOA is extracted from the time delay between *two* sensors for the far-field case.

The suggested time delay direction finding (TDDF) algorithm utilizes the linear relationship between the time delay vector and the DOA vector in Cartesian coordinates. This linear relationship allows a closed form estimation of the DOA vector. The transformation to polar coordinates, i.e., azimuth and elevation, is straightforward for 1-D and 2-D array geometries. For arbitrarily chosen 3-D array configurations, the best estimator requires a simple nonlinear least squares minimization. Alternatively, a closed form suboptimal solution for the 3-D configuration is also suggested. Finally, it is shown that the TDDF azimuth and elevation estimator achieves the CRLB provided that the time delay estimators achieves the CRLB as well.

## I. METHODS

### A. The time delay direction finding (TDDF) algorithm

Consider an array of  $M$  identical omnidirectional sensors with a known arbitrary geometry measuring the wave field generated by a single far-field wideband source in the presence of an additive noise. Let  $\vec{\mathbf{r}}_i$  denote the location of the  $i$ th sensor, where  $\vec{\mathbf{r}}_i = [x_i, y_i, z_i]$  for the 3-D array,  $\vec{\mathbf{r}}_i = [x_i, y_i]$  for the 2-D case, and  $\vec{\mathbf{r}}_i = [x_i]$  for the 1-D case, and let  $\phi$  and  $\theta$  denote the azimuth and elevation angles of the radiating source, respectively (see Fig. 1).

<sup>a)</sup>Electronic mail: Haim@Biomed.Technion.ac.il

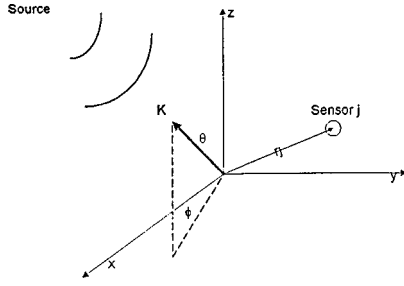


FIG. 1. Schematic representation of the model and the coordinate system used here.

Let us now define the differential delay vector,

$$\vec{\tau} = [\tau_{12}, \tau_{13}, \dots, \tau_{1M}]^T; \quad \tau_{1j} \equiv \tau_j - \tau_1, \quad (1)$$

where the first sensor serves as a reference. The signal DOA vector for the far-field case is given by

$$\vec{\mathbf{k}} = \begin{bmatrix} k_x \\ k_y \\ k_z \end{bmatrix} = \begin{bmatrix} \sin(\theta)\cos(\phi) \\ \sin(\theta)\sin(\phi) \\ \cos(\theta) \end{bmatrix}. \quad (2)$$

The time delay between any two sensors is equal to the projection of the distance vector between them along the  $\vec{\mathbf{k}}$  vector divided by the sound velocity. Consequently, the delay vector can be expressed as follows:

$$\vec{\tau} = -\frac{\mathbf{R}\vec{\mathbf{k}}}{c}; \quad \mathbf{R} \equiv \begin{bmatrix} \vec{r}_2 - \vec{r}_1 \\ \vdots \\ \vec{r}_M - \vec{r}_1 \end{bmatrix}, \quad (3)$$

where  $c$  is the wave velocity and the matrix  $\mathbf{R}$  is composed of the distance vectors between all the sensors and the reference sensor.

The objective is to estimate  $\vec{\mathbf{k}}$  from the measured time delay vector  $\vec{\tau}$ . Studying Eq. (3), it is evident that the problem is overdetermined. Thus, it is suggested to apply the least squares (LS) method to obtain the estimation. Defining the error as the difference between the measured time difference vector and the evaluated time vector (calculated from the assumed  $\vec{\mathbf{k}}$  vector), the error vector is given by

$$\vec{\varepsilon} = \left( \frac{\mathbf{R}\vec{\mathbf{k}}}{c} + \hat{\vec{\tau}} \right). \quad (4)$$

In the general case, the measurement errors of the time delay vector need not be uncorrelated. Hence, the solution depends on the covariance matrix  $\Lambda_\tau$  of the delays measurements which is defined by

$$\Lambda_\tau = E\{\hat{\vec{\tau}}\hat{\vec{\tau}}^T\} - E\{\hat{\vec{\tau}}\}E\{\hat{\vec{\tau}}\}^T = \text{COV}\{\hat{\vec{\tau}}\}, \quad (5)$$

where  $E\{\}$  denotes the expected value operator. The problem is ‘‘overdetermined’’ for  $M > 3$ . The LS solution for  $\vec{\mathbf{k}}$ , the DOA vector, in this case is given by<sup>15</sup>

$$\begin{aligned} \hat{\vec{\mathbf{k}}} &= \underset{\vec{\mathbf{k}}}{\text{Arg Min}} \left\{ \left( \frac{\mathbf{R}\vec{\mathbf{k}}}{c} + \hat{\vec{\tau}} \right)^T \Lambda_\tau^{-1} \left( \frac{\mathbf{R}\vec{\mathbf{k}}}{c} + \hat{\vec{\tau}} \right) \right\} \\ &= -c(\mathbf{R}^T \Lambda_\tau^{-1} \mathbf{R})^{-1} \mathbf{R}^T \Lambda_\tau^{-1} \hat{\vec{\tau}} \equiv -c\mathbf{B}\hat{\vec{\tau}}. \end{aligned} \quad (6)$$

Thus, estimating the DOA vector becomes a simple multiplication between the measured time delay vector  $\vec{\tau}$  and a data-independent matrix  $\mathbf{B}$ . The matrix  $\mathbf{B}$  depends on the array geometry (through  $\mathbf{R}$ ) and the time delay covariance matrix which under the assumption of spatially uncorrelated noise is known *a priori* up to a multiplicative factor which cancels out in this equation. Consequently, it can be calculated off-line.

In order to express the DOA vector in terms of azimuth and elevation, one has to write the vector  $\vec{\mathbf{k}}$  in a polar coordinate representation. For a 1-D array configuration only  $\hat{k}_x$  can be estimated. Hence, assuming horizontal elevation, the azimuth angle is given by

$$\hat{\phi} = \cos^{-1}(\hat{k}_x). \quad (7)$$

With a 2-D array, both the azimuth and elevation angles can be calculated by

$$\hat{\phi} = \tan^{-1}(\hat{k}_y / \hat{k}_x), \quad (8a)$$

$$\hat{\theta} = \cos^{-1}(\hat{k}_z) = \cos^{-1}((1 - (\hat{k}_x^2 + \hat{k}_y^2))^{1/2}). \quad (8b)$$

For the case of a 3-D array, Eq. (2) yields three nonlinear equations with two unknowns ( $\phi, \theta$ ). Again the problem is overdetermined. Thus, the azimuth and elevation angles ( $\phi, \theta$ ) can be evaluated as the nonlinear least square estimator solving Eq. (2):

$$(\hat{\phi}, \hat{\theta}) = \underset{\phi, \theta}{\text{Arg Min}} \{ (\vec{\mathbf{k}} - \vec{\mathbf{k}}(\hat{\phi}, \hat{\theta}))^T \Lambda_k^{-1} (\vec{\mathbf{k}} - \vec{\mathbf{k}}(\hat{\phi}, \hat{\theta})) \}, \quad (9)$$

where  $\Lambda_k$  is the covariance matrix of vector  $\vec{\mathbf{k}}$ , which is given in Appendix A [Eq. (A4)].

An alternative simplified close form suboptimal estimate is proposed by

$$\hat{\phi} = \tan^{-1}(\hat{k}_y / \hat{k}_x), \quad (10a)$$

$$\hat{\theta} = \tan^{-1}((\hat{k}_x^2 + \hat{k}_y^2)^{1/2} / \hat{k}_z). \quad (10b)$$

In Appendix A the performance of the TDDF algorithm is analyzed, and it is shown that it is asymptotically efficient. Furthermore, it is shown that under certain geometrical constraints for the sensors arrangement even the closed form 3-D solution achieves the CRLB.

Importantly, it should be noted that the azimuth estimates given above for the 2-D and 3-D array configurations are independent of the wave velocity  $c$  (stems from the fact that the solutions are given in terms of the ratio between  $\hat{k}_y$  and  $\hat{k}_x$ , and both are linear functions of  $c$ ). Therefore, errors in the assumed sound speed will not induce errors in the azimuth angle.

## B. Performance analysis

In Appendix A the covariance matrix for  $(\hat{\phi}, \hat{\theta})$  is calculated. The performance of the TDDF estimator is compared to the theoretical CRLB. It is shown that for the 1-D and 2-D cases the estimator is asymptotically efficient since it achieves the bound. For the 3-D case the closed form es-

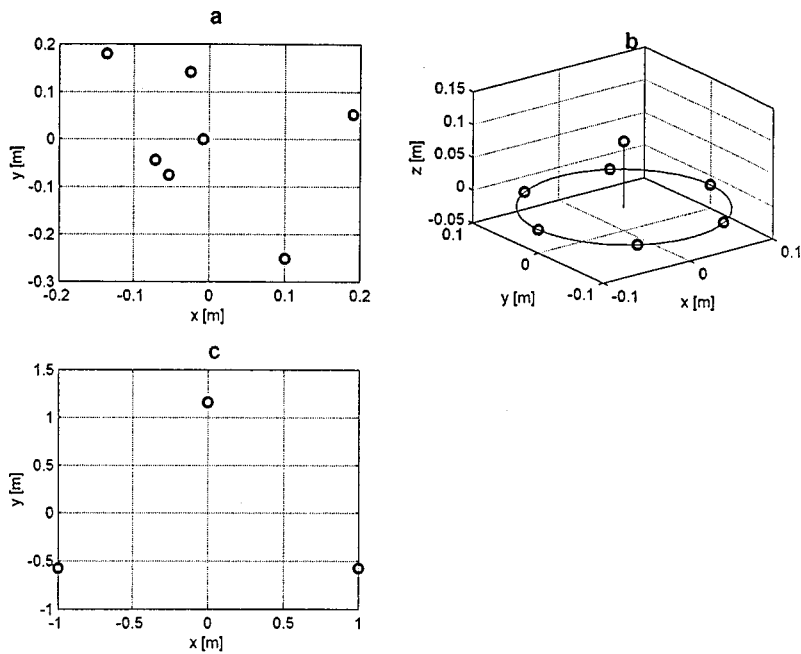


FIG. 2. (a) The 2-D array geometry consisting of seven randomly located microphones, which was used in the first two numerical simulations. (b) The 3-D array geometry, which was used in the third numerical simulation and the experimental measurements. (c) The 2-D array consisting of three microphones which was used in the last numerical simulation.

timator given in Eq. (10) is not always efficient. However, we derived constraints on the array geometry in which the CRLB is also achieved.

## II. RESULTS

In this section, the performance of the TDDF algorithm is demonstrated via numerical simulations and by experimental results.

### A. Numerical simulations

Simulations were conducted for 2-D and 3-D arrays. The 2-D array was comprised of randomly located seven sensors, as shown in Fig. 2(a). In the first set of simulations the source was positioned at a fixed location with an azimuth angle of  $60^\circ$  and an elevation of  $30^\circ$ . The SNR was scanned

in the range of  $-10$  to  $+10$  dB, the integration time was 50 ms, and the frequency bandwidth was 500–1500 Hz. The noisy estimates of the time delay vectors were generated as Gaussian random vectors with a covariance matrix  $\Lambda_7$  given in Eq. (A2). The propagation velocity was taken to be 340 m/s.

Five hundred Monte Carlo runs were performed for each SNR value. The azimuth angle was calculated with the TDDF algorithm and the corresponding errors were computed. The standard deviation of the localization errors was then estimated. For comparison, the CRLB was also calculated as explained in Appendix A. The results are depicted in Fig. 3. For clarity of presentation, only 25 azimuth estimation errors are plotted (as small dots) at each SNR level. The standard deviations of the TDDF estimator are depicted as circles and the corresponding CRLB is depicted as a solid line

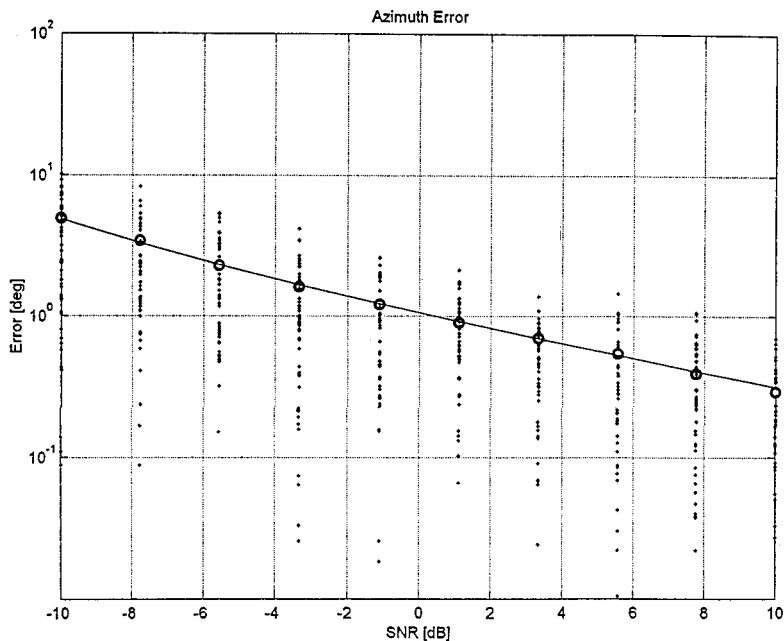


FIG. 3. Azimuth estimation errors of the simulated 2-D array, as a function of the SNR. The source is positioned at an azimuth of  $60^\circ$  and an elevation angle of  $30^\circ$ . The solid line is the CRLB. The dots depict the magnitude of the errors of the first 25 individual runs (out of the 500 used). The circles depict the standard deviation of the TDDF estimator.

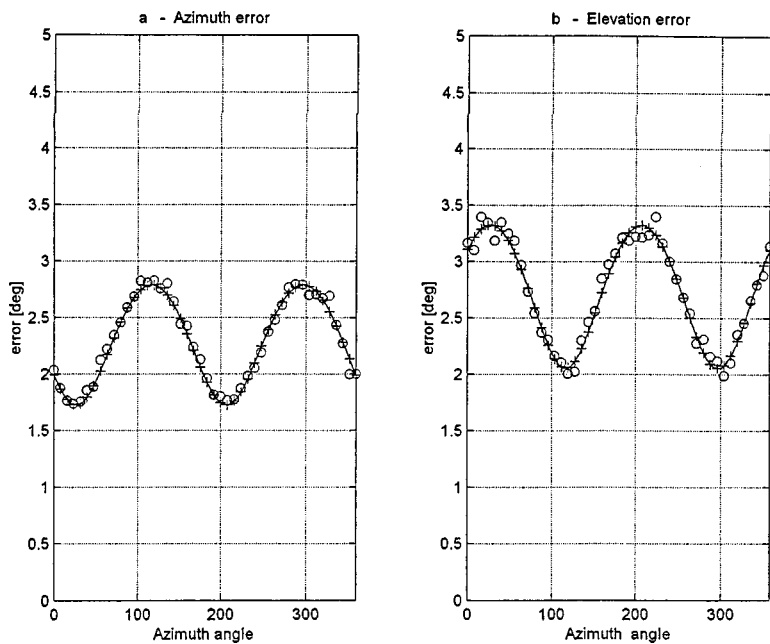


FIG. 4. The errors of the TDDF algorithm for the azimuth (a) and elevation (b) as a function of the azimuth, for the 2-D array shown in Fig. 2(a). The SNR is  $-6$  dB, the elevation angle is  $50^\circ$ . The solid line is the CRLB, and the circles depict the standard deviation of the TDDF estimator.

line for the entire range. It can be seen that the standard deviations of the TDDF estimator are effectively located on the CRLB line (in accordance with the mathematical derivation given in Appendix A).

In a second set of simulations, the same 2-D array was used. The SNR was held fixed at  $-6$  dB. The elevation angle was set to  $50^\circ$ , and the azimuth angle was scanned in the range of  $0^\circ$ – $360^\circ$ . The corresponding CRLB line was calculated and the standard deviation of the TDDF estimator was evaluated for each angle. Figure 4(a) and (b) displays the errors of the TDDF algorithm for the azimuth and elevation, respectively. From these curves it is observed that the TDDF estimator achieves the CRLB, both in azimuth and elevation, for a 2-D array with an arbitrary geometry.

In the third set of simulations the 3-D array shown in Fig. 2(b) was used. This array has six sensors equally dis-

tributed on a circle with a radius of 0.1 m and the seventh sensor is located 0.1 m above the center of the array. The SNR was again held fixed at  $-6$  dB. The azimuth was set to  $20^\circ$ , and the elevation angle was scanned in the range  $0^\circ$ – $180^\circ$ . The following quantities were calculated this time: the CRLB, the standard deviation of the TDDF estimator, and the theoretical standard deviation calculated from Eq. (A15). Figure 5(a) and (b) plots these quantities as a function of the elevation angle.

The 3-D array used here obeys the condition given in Eq. (A17). Consequently, the CRLB is achieved for the azimuth TDDF estimates [Fig. 5(a)]. For the elevation angle, however, the obtained estimate errors are larger than the CRLB. This observation is consistent with the fact that the array geometry does not comply with condition given in Eq. (A19). Nevertheless, the degradation is moderate for this ar-

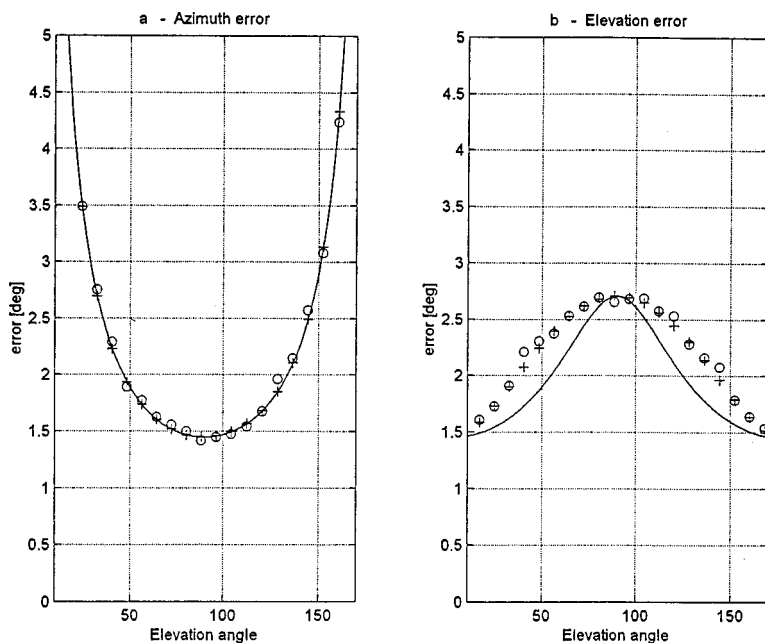


FIG. 5. The errors of the TDDF algorithm for the azimuth (a) and elevation (b) as a function of the elevation angle, for the 3-D array shown in Fig. 2(b). The SNR is  $-6$  dB, the azimuth is  $20^\circ$ . The solid line is the CRLB. The circles depict the standard deviation of the TDDF estimator, and the analytic variance expressions [Eq. (A13)] are depicted as '+.'.

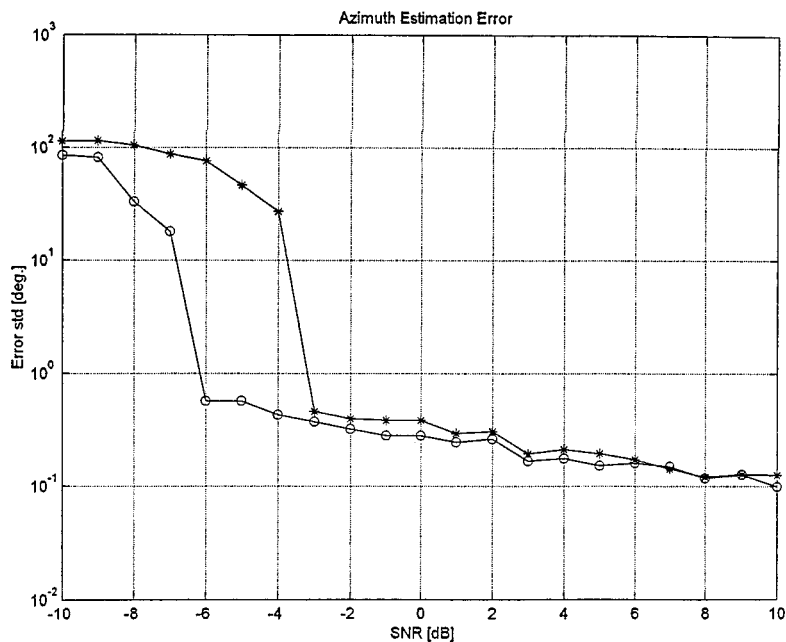


FIG. 6. Azimuth estimation errors of the three-microphone 2-D array used in the last simulation, as a function of the SNR. The source is positioned at an azimuth of  $30^\circ$ . The standard deviation of the TDDF estimate are depicted by “\*” and the standard deviation of the Beamformer is plotted by “○.”

ray configuration. This implies that the closed form estimation given by equation set (10) is sufficiently accurate for practical purposes. Finally, it can be observed that the estimated errors match the theoretical standard deviations predicted by Eq. (A15). It should also be noted that bias was also estimated in all the above simulations and was found negligible (two orders of magnitude smaller than the variance contribution to the total error).

In practical applications it is usually not easy to obtain the optimal estimate for the time delay vector  $\vec{\tau}$ . In the last numerical example we demonstrate the performance of the TDDF algorithm when using a suboptimal estimator for the time delay vector as presented in Appendix C. The time delay vector was estimated via a cross correlation between the reference sensor (no. 1) and the other sensors. The performance of the TDDF algorithm is compared to that of a beamformer. This simulation uses the 2-D array consisting of three microphones which is shown in Fig. 2(c). The source direction was set at an azimuth of  $30^\circ$  and an elevation angle of  $90^\circ$ . The SNR was scanned in the range of  $-10$  to  $+10$  dB. Here, the simulation generates the time record of the sensor data assuming spatially uncorrected noise. Both the signal and the noise were random Gaussian variables with a bandwidth of 100–3000 Hz. In order to perform the beam steering required in the beamformer the data were interpolated by a factor of 10. The standard deviation error of both estimators was estimated by 100 Monte Carlo runs. The standard deviations of the TDDF estimate are depicted by asterisks in Fig. 6 while the standard deviation of the beamformer is denoted by circles. As can be seen for most of the studied SNR range the performance of the TDDF is the same as that of the beamformer. However, the threshold point for the TDDF appears at  $\text{SNR} = -3$  dB, which is higher than the threshold observed for the beamformer ( $-6$  dB). This result is not surprising since the TDDF is not an ML estimator as the beamformer. Potentially there are two factors that can cause the performance of the TDDF to collapse. The first is the time delay’s vector estimation process, and the second is

the nonlinear operation for estimating  $\phi$ . In all the cases we have tested, the time delay estimate was the first one to diverge. Practically it was observed that the cross-correlation functions have generated spurious peaks at low SNR, and this is probably the main cause for the performance diverging at low SNR.

## B. Experimental results

A 3-D array consisting of seven microphones (Audio-Technica MT350B) arranged in the same configuration depicted in Fig. 2(b) (radius=0.1 m, height=0.1 m) was used. Two experiments were conducted. In the first experiment the array was located in an anechoic chamber (internal dimension of  $1.7 \times 1.7 \times 1.7$  m). In the second experiment the array was placed in an ordinary room. The sound source was a recorded male voice (Richard Burton) reading a 20-s long sentence. The signal was played via a loudspeaker located 1.5 m from the array. The outputs of the array were recorded using an 8-channel tape recorder (Sony-pc208A). The time delay between the sensors and the central microphone was estimated by filtering the data by a bandpass filter of 500–1500 Hz, and performing a cross-correlation process. The integration time was 40 ms, yielding about 500 independent measurements to estimate the system performance. After completion of each set of measurements, the array was rotated by  $30^\circ$  and the procedure was repeated.

The azimuth angles corresponding to each set of measurements was estimated using the TDDF algorithm. The standard deviation of the errors for the TDDF estimates was then evaluated. The results are outlined in Fig. 7. The data from the anechoic chamber is denoted by “○” and the data from the regular room is presented by “\*”. As can be observed, the average TDDF error for the anechoic chamber experiment was about  $1.5^\circ$ . The second experiment was held in the regular room, and the average error was about  $5^\circ$ . This degradation is attributed to the room reverberations and the background noise. We have measured the reverberation time

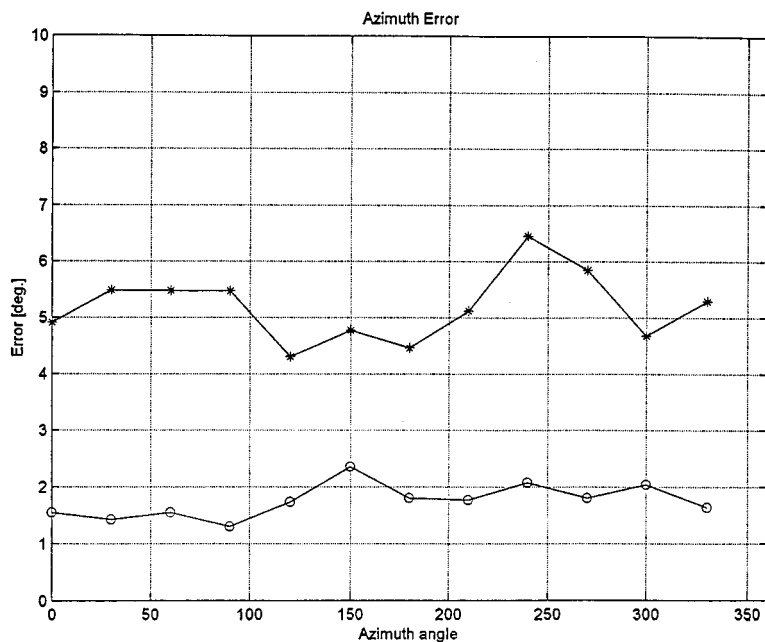


FIG. 7. Experimentally measured azimuth errors of the TDDF algorithm as a function of the azimuth using the 3-D array. The source was a speech signal played from a loud speaker 1.5 m away from the array. The circles “○” denotes the results measured in an unechoic chamber, and the “\*” indicates the results measured in an ordinary room.

in both rooms. In the unechoic chamber the reverberation time was about 10 ms, while in the regular room the reverberation time was about 250 ms. Thus, we believe that the room reverberation was the major cause for the degradation in the accuracy of the direction estimates.

### III. DISCUSSION

This paper presents and analyzes the time delay direction finding (TDDF) algorithm for a single emitting source using a multi-sensor array. The algorithm extracts the azimuth and elevation angles directly from the estimated time delays between the array elements. The algorithm offers computational simplicity as it utilizes the linear relationship between the time delay vector and the DOA vector in Cartesian coordinates. This linear relationship allows a closed form estimation of the DOA vector.

An asymptotic performance analysis of the TDDF algorithm, using a small error assumption, is performed. For the 1-D and 2-D array configurations it is shown that the TDDF algorithm achieves the Cramer Rao lower bound (CRLB) provided that the time delay vector estimator achieves the CRLB as well. This was proven mathematically in Appendix A and was demonstrated by numerical simulations. For a 3-D array configuration a suboptimal closed form estimator is presented [Eq. (10)]. Nevertheless, it is shown that when using array geometries that obey certain constraints the closed form solution also achieves the CRLB. If the array obeys the condition given by Eq. (A17), then the azimuth estimation is statistically efficient. Furthermore, if the array also obeys the constraints that are given in Eq. (A19), the estimator is efficient for the elevation angle as well.

Numerical and experimental results were given to demonstrate the performance of the TDDF algorithm. The experimental results with a seven-microphone array have shown that in an unechoic chamber the average TDDF azimuth error was about 1.5°, while in a regular room the av-

erage error was about 5°. These results indicate that the TDDF can serve as a practical tool for passive localization of a single radiating source.

The proposed TDDF algorithm offers several advantages over the popular beamforming approach.<sup>11</sup> First, the TDDF algorithm is considerably more efficient in terms of computational load. It calculates the azimuth and the elevation angle directly from the estimated time delays, and does not involve a two-dimensional search over the array manifold as the beamformer.

Second, for the 2-D and 3-D array configurations, the TDDF algorithm does not require the *a priori* knowledge of the propagation velocity to estimate the azimuth [see Eqs. (8a) and (10a), respectively]. This property of the TDDF is very useful in acoustic applications where uncertainty in the propagation velocity occurs due to wind and temperature variations.<sup>16</sup> This is contrary to the beamformer, which uses the wave propagation velocity as input. In principle, the beamforming process could scan the velocity as an additional unknown parameter. However, this would substantially increase the computational load as an additional parameter would have to be scanned.

The third advantage of the TDDF method arises in applications where the signal is a very short transient and the arrival time of the pulse is directly measured by the system hardware. Since only the time delays are available in this case, the beamforming is not applicable. For the TDDF algorithm, on the other hand, this information is sufficient.

Finally, in certain acoustical and geophysical applications, loss of spatial coherence of the signal received at the sensors may occur if the distance between the sensors is large,<sup>1,16</sup> thus precluding the use of the beamforming approach. In such cases, nevertheless, time delay between sensors can still be estimated via incoherent processing means, such as time of arrival difference, and the TDDF algorithm is still applicable.

The TDDF has one major disadvantage. It is limited to a

single source scenario. The beamformer algorithm, on the other hand, can localize more than one source, provided that the angular separation between the sensors is more than the beam width.

## ACKNOWLEDGMENTS

The authors thank the ‘‘Lamar Signal Processing Israel’’ for providing the facilities and for their assistance in performing the experiments. The support of the Irving and Adeele Rosenberg Foundation Inc. (HA) is also acknowledged.

## APPENDIX A: PERFORMANCE ANALYSIS OF THE TDDF ALGORITHM

In order to evaluate the performance of the TDDF algorithm, analytic expressions for the accuracy of the azimuth and elevation estimations for the 1-D, 2-D, and 3-D cases are first derived. These expressions are then compared to the expression of the CRLB as derived by Nielsen,<sup>17</sup> and cited in Appendix B.

For uniformity and simplicity of notations let us define  $\gamma = \phi$  for the 1-D case, and  $\tilde{\gamma} = (\phi, \theta)^T$  for the 2-D and 3-D cases. Under the assumption of small errors the covariance matrix of  $\gamma$  can be expressed as

$$\Lambda_\gamma = \bar{\nabla}_k \gamma \cdot \Lambda_k \cdot \bar{\nabla}_k^T \gamma = \bar{\nabla}_k \gamma \cdot \bar{\nabla}_\tau k \cdot \Lambda_\tau \cdot \bar{\nabla}_\tau^T k \cdot \bar{\nabla}_k^T \gamma, \quad (\text{A1})$$

where  $\Lambda_x$  is the covariance matrix of the vector  $x$ , and  $\bar{\nabla}_{x,y}$  is the gradient (Jacobian) of  $\bar{y}$  with respect to  $\bar{x}$ . Again, for a small error assumption, it can be verified that the DOA estimates are asymptotically unbiased, thus the covariance matrix represents the total error of the estimator.

Clearly, the performance of the TDDF algorithm depends on the covariance matrix of the time delay vector. To demonstrate the performance of the TDDF algorithm we shall assume that the time delay vector estimator achieves the CRLB. An efficient algorithm for estimation of the time delay vector, assuming that both the signal and the noise are zero mean uncorrelated Gaussian processes and the noise is spatially uncorrelated, has been proposed and studied by Hahn and Tretter.<sup>8</sup> Their work presents an estimator for the time delay vector which achieves the CRLB, and does not require the beamformer process. Their scheme is based on estimating  $M(M-1)$  individual time delays via a prefiltered correlators. The vector  $\tilde{\tau}$  is obtained by a linear combination of the individual time delays. The covariance matrix for the time delay vector for this estimator, assuming that the SNR is the same for all the sensors, is given by

$$\Lambda_\tau = \text{CRLB}_\tau = \frac{\sigma_\tau^2}{2} [\mathbf{I}_{M-1} + \tilde{\mathbf{I}}_{M-1} \tilde{\mathbf{I}}_{M-1}^T], \quad (\text{A2})$$

where  $\mathbf{I}_M$  is the  $M \times M$  identity matrix,  $\tilde{\mathbf{I}}_M$  is an  $M$ -dimensional vector of ones,

$$\sigma_\tau^2 = \left( \sum_{l=1}^{L_{\max}} (\omega_0 l)^2 \rho(l) \frac{M \rho(l)}{1 + M \rho(l)} \right)^{-1},$$

and  $\rho(l) \equiv S(l)/N(l)$  denotes the SNR at the frequency  $(\omega_0 l)$ .

In the following it shall be assumed that  $\Lambda_\tau$  is given by Eq. (A2), i.e., efficient estimate of  $\tilde{\tau}$ ,

In Appendix B we derive  $\Lambda_\tau$  for suboptimal time delay estimator via only  $(M-1)$  correlators, using one sensor as a reference sensor, i.e., an efficient estimate is only obtained for the separate pairwise delays. It is shown that for sufficiently high SNR this estimator also achieves the CRLB.

First, the covariance matrix of the direction vector  $\bar{k}$  is calculated. From Eq. (6),

$$\bar{\nabla}_\tau^T \bar{k} = c \mathbf{B}, \quad (\text{A3})$$

and therefore

$$\Lambda_k = E\{\bar{k} \cdot \bar{k}^T\} = \mathbf{B} \Lambda_\tau \mathbf{B}^T c^2. \quad (\text{A4})$$

Using the definition of  $\mathbf{B}$  in Eq. (6) and applying some algebraic simplifications yields.

$$\Lambda_k = (\mathbf{R} \Lambda_\tau^{-1} \mathbf{R}^T)^{-1} c^2. \quad (\text{A5})$$

Applying the matrix inversion lemma to Eq. (A2) it can be written that

$$\Lambda_\tau^{-1} = \frac{2}{\sigma_\tau^2} \left[ \mathbf{I}_{M-1} - \frac{\tilde{\mathbf{I}}_{M-1} \tilde{\mathbf{I}}_{M-1}^T}{M} \right]. \quad (\text{A6})$$

From the definitions of  $\mathbf{R}$  in Eq. (3) it follows that

$$\mathbf{R} = [-\tilde{\mathbf{I}}_{M-1} \quad \mathbf{I}_{M-1}] \mathbf{P}, \quad (\text{A7})$$

where  $\mathbf{P}$  is the sensor position matrix defined by

$$\mathbf{P} = [\tilde{\mathbf{x}}, \tilde{\mathbf{y}}, \tilde{\mathbf{z}}] = \begin{bmatrix} x_1 & y_1 & z_1 \\ \vdots & \vdots & \vdots \\ x_M & y_M & z_M \end{bmatrix}. \quad (\text{A8})$$

Substituting Eqs. (A6) and (A7) into Eq. (A5) yields, after some algebraic manipulations,

$$\Lambda_k = \left( \mathbf{P}^T \mathbf{P} - \frac{(\mathbf{P}^T \tilde{\mathbf{I}}_M)(\mathbf{P}^T \tilde{\mathbf{I}}_M)^T}{M} \right)^{-1} \frac{\sigma_\tau^2}{2} c^2. \quad (\text{A9})$$

Assuming without any loss of generality that the coordinate origin is in the center of gravity, i.e.,  $\mathbf{P}^T \tilde{\mathbf{I}}_M = \tilde{\mathbf{0}}$ , we finally get the simple expression for the DOA vector covariance matrix,

$$\Lambda_k = (\mathbf{P}^T \mathbf{P})^{-1} \frac{\sigma_\tau^2}{2} c^2. \quad (\text{A10})$$

In the following, the expressions for the accuracy of the azimuth and elevation TDDF estimations are derived and compared to the CRLB which is cited in Appendix B.

### 1. A linear (1-D) array configuration

For the 1-D case,  $\gamma = \phi$  and  $\bar{\mathbf{k}} = k_x = \cos(\phi)$ . Thus  $\bar{\nabla}_k \gamma = -1/\sin(\phi)$  and

$$\Lambda_k = \frac{1}{\sum x_i^2} \frac{\sigma_\tau^2}{2} c^2$$

inserted into Eq. (11) gives



$$\sigma_\phi^2 = \frac{1}{\sin^2(\phi) \sum x_i^2} \frac{\sigma_\tau^2}{2} c^2. \quad (\text{A11})$$

The CRLB for the 1-D case is given by  $\text{CRLB}(\phi) = 1/J_{\phi\phi}$  (see Appendix B). Substituting  $y_i = z_i = 0$  in this expressions yields

$$\text{CRLB}(\phi) = \frac{1}{\sin^2(\phi) \sum x_i^2} \frac{\sigma_\tau^2}{2} c^2.$$

As can be observed this expression is identical to the right-hand side of Eq. (A11), indicating that the TDDF estimate achieves the CRLB in this case.

## 2. A planar (2-D) array configuration

From Eqs. (2) and (8) the Jacobian  $\bar{\nabla}_k \gamma$  is given by

$$\begin{aligned} \bar{\nabla}_k \gamma &= \frac{1}{-\sin(\theta)\cos(\theta)} \begin{bmatrix} -\sin(\phi)\cos(\theta) & \cos(\phi)\cos(\theta) \\ -\cos(\phi)\sin(\theta) & -\sin(\phi)\sin(\theta) \end{bmatrix}. \end{aligned} \quad (\text{A12})$$

Inserting Eqs. (A12) and (A10) into Eq. (A1) yields

$$\begin{aligned} \sigma_\phi^2 &= \frac{1}{\sin^2(\theta)} \\ &\times \frac{\cos^2(\phi) \sum x_i^2 + 2 \cos(\phi) \sin(\phi) \sum x_i y_i + \sin^2(\phi) \sum y_i^2}{\sum x_i^2 \sum y_i^2 - (\sum x_i y_i)^2} \\ &\times \frac{\sigma_\tau^2}{2} c^2 \end{aligned} \quad (\text{A13a})$$

and

$$\begin{aligned} \sigma_\theta^2 &= \frac{1}{\cos^2(\theta)} \\ &\times \frac{\sin^2(\phi) \sum x_i^2 - 2 \cos(\phi) \sin(\phi) \sum x_i y_i + \cos^2(\phi) \sum y_i^2}{\sum x_i^2 \sum y_i^2 - (\sum x_i y_i)^2} \\ &\times \frac{\sigma_\tau^2}{2} c^2. \end{aligned} \quad (\text{A13b})$$

Applying a lengthy but straightforward evaluation of the expressions for the CRLB Eq. (A1) for both the azimuth and the elevation angles for the 2-D arrays, i.e.,  $z_i = 0$ , shows that they are identical to Eq. (A13). Thus, it is concluded that the TDDF algorithm is a statistically efficient estimator for 2-D array, which reaches the CRLB. It is important to note that no constraints on the array geometry were applied.

## 3. A spatial (3-D) array configuration

The estimate of  $\gamma = (\phi, \theta)$  for the 3-D array case involves a nonlinear LS minimization Eq. (9). An alternate closed form suboptimal estimator was suggested in Eq. (10). Here we calculate the performance of the suboptimal estimator and derive the conditions on the array geometry that guarantee statistical efficiency (achieves the CRLB).

From Eqs. (2) and (10) the Jacobian  $\bar{\nabla}_k \gamma$  is given by

$$\bar{\nabla}_k \gamma = \begin{bmatrix} -\frac{\sin(\phi)}{\sin(\theta)} & \frac{\cos(\phi)}{\sin(\theta)} & 0 \\ \cos(\theta)\cos(\phi) & \cos(\theta)\sin(\phi) & -\sin(\theta) \end{bmatrix}. \quad (\text{A14})$$

Using Eqs. (A1) and (A5),

$$\begin{aligned} \sigma_\phi^2 &= \frac{\sigma_\tau^2 c^2}{2 \sin^2(\theta)} [\sin(\phi) \quad \cos(\phi) \quad 0] \\ &\times (\mathbf{P}^T \mathbf{P})^{-1} [\sin(\phi) \quad \cos(\phi) \quad 0]^T \end{aligned} \quad (\text{A15})$$

and

$$\begin{aligned} \sigma_\theta^2 &= \frac{\sigma_\tau^2 c^2}{2} [\cos(\theta)\cos(\phi) \quad \cos(\theta)\sin(\phi) \quad -\sin(\theta)] \\ &\times (\mathbf{P}^T \mathbf{P})^{-1} \cdot [\cos(\theta)\cos(\phi) \quad \cos(\theta)\sin(\phi) - \sin(\theta)]^T. \end{aligned} \quad (\text{A16})$$

In general this estimator is not efficient. However, if the array obeys the following geometrical conditions,

$$\sum x_i y_i = \sum x_i z_i = \sum y_i z_i = 0, \quad (\text{A17})$$

$$\sum x_i^2 = \sum y_i^2,$$

then the variance of the azimuth angle estimation is given by

$$\sigma_\phi^2 = \frac{1}{\sin^2(\theta)} \frac{1}{\sum x_i^2} \frac{\sigma_\tau^2}{2} c^2. \quad (\text{A18})$$

Evaluating the CRLB for the azimuth estimate under the same condition yields an identical expression. Thus, under the conditions outlined in Eq. (A17), the TDDF algorithm is also an efficient estimator for the azimuth angle. When studying the conditions for uncoupled estimates of azimuth and elevation angles, Nielsen<sup>17</sup> has also reached the same conditions given in Eq. (A17), and gave a few examples of 3-D arrays obeying these constraints.

If we further constrain the array geometry and require a fully balanced array configuration, which obeys the following condition,

$$\sum x_i^2 = \sum z_i^2. \quad (\text{A19})$$

In addition to the conditions outlined in Eq. (A17), it can be shown that for the elevation estimate,

$$\sigma_\theta^2 = \frac{1}{\sum x_i^2} \frac{\sigma_\tau^2}{2} c^2 = \text{CRLB}_\theta. \quad (\text{A20})$$

Thus, under the above conditions the TDDF algorithm achieves the CRLB for both the azimuth and the elevation angles, and is therefore an asymptotically efficient estimator.

## APPENDIX B: CRLB FOR THE AZIMUTH AND ELEVATION ANGLES

Nielsen<sup>17</sup> derived analytic expressions for the Cramer Rao lower bound for the estimation errors of the azimuth angle  $\phi$  and the elevation angle  $\theta$ , using 3-D arrays:

$$\begin{aligned} \text{CRLB}(\theta) &= J_{\phi\phi} / (J_{\phi\phi} J_{\theta\theta} - J_{\theta\phi}^2), \\ \text{CRLB}(\phi) &= J_{\theta\theta} / (J_{\phi\phi} J_{\theta\theta} - J_{\theta\phi}^2), \end{aligned} \quad (\text{B1})$$

where

$$\begin{aligned} J_{\phi\phi} &= G \sin^2(\theta) \sum_{i=1}^M [x_i \sin(\phi) - y_i \cos(\phi)]^2, \\ J_{\theta\theta} &= G \sum_{i=1}^M [x_i \cos(\phi) \cos(\theta) + y_i \sin(\phi) \cos(\theta) \\ &\quad - z_i \sin(\theta)]^2, \\ J_{\phi\theta} &= G \sin(\theta) \sum_{i=1}^M [x_i \sin(\phi) - y_i \cos(\phi)] \\ &\quad \times [x_i \cos(\phi) \cos(\theta) + y_i \sin(\phi) \cos(\theta) - z_i \sin(\theta)], \\ G &= \sum_{l=1}^{L_{\max}} (\omega_0 l) \frac{M \rho^2(l)}{1 + M \rho(l)}. \end{aligned}$$

In these expressions, the coordinates origin is in the center of gravity of the array, i.e.,

$$\sum x_i = \sum y_i = \sum z_i = 0,$$

and the coordinate system is given in Fig. 1

### APPENDIX C: SUBOPTIMAL ESTIMATION FOR THE TIME DELAYS VECTOR

In this appendix we consider a suboptimal estimation of the time delays vector which estimates only  $M-1$  time delays between the first sensor relative to all the other sensors in the array. Each time delay estimation is based on the data of these two sensors only, and ignores the fact it is part of an  $M$ -sensor array. Efficient estimate for the time delay between two sensors can be obtained by maximizing the generalized cross correlation.<sup>2</sup>

Based on the derivation in Ref. 8, the covariance matrix of this estimator is given by

$$\begin{aligned} \Lambda_{\tau} &= \frac{1}{\left\{ \sum_{k=1}^{K_{\max}} 2(\omega_0 l)^2 \rho^2 / (1 + 2\rho) \right\}^2} \sum_{k=1}^{k_{\max}} 2 \frac{(\omega_0 l)^2 \rho^2}{(1 + 2\rho)^2} \\ &\quad \times \begin{bmatrix} 1 + 2\rho & \cdot & \rho & \cdot \\ \cdot & 1 + 2\rho & \cdot & \cdot \\ \rho & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & 1 + 2\rho \end{bmatrix}. \end{aligned} \quad (\text{C1})$$

In general this estimator does not achieve the CRLB, however for high SNR case  $2\rho \gg 1$  it can be seen that

$$\Lambda_{\tau} \approx \frac{1}{2 \cdot \sum_{k=1}^{K_{\max}} (\omega_0 l)^2 \rho} [\mathbf{I}_{M-1} + \tilde{\mathbf{I}}_{M-1} \tilde{\mathbf{I}}_{M-1}^T].$$

Evaluating the  $\text{CRLB}_{\tau}$  as given in Eq. (A2) for the high SNR case yields the same expression, i.e., for a good SNR case the suboptimal time delays estimator achieves the  $\text{CRLB}_{\tau}$ .

<sup>1</sup>S. Haykin, editor, *Array Signal Processing* (Prentice-Hall, Englewood Cliffs, NJ, 1985).

<sup>2</sup>G. C. Carter, "Time delay estimation for passive signal processing," *IEEE Trans. Acoust., Speech, Signal Process.* **ASSP-29**, 463–470 (1981).

<sup>3</sup>A. H. Quazi, "An overview on time delay estimation in active and passive system for target localization," *IEEE Trans. Acoust., Speech, Signal Process.* **ASSP-29**, 527–533 (1981).

<sup>4</sup>A. G. Piersol, "Time delay estimation using phase data," *IEEE Trans. Acoust., Speech, Signal Process.* **ASSP-29**, 471–477 (1981).

<sup>5</sup>C. H. Knapp and G. C. Carter, "The generalized correlation method for estimation of time delay," *IEEE Trans. Acoust., Speech, Signal Process.* **ASSP-24**, 320–327 (1976).

<sup>6</sup>C. Hassab and R. E. Boucher, "Optimum estimation of time-delay by a generalized correlator," *IEEE Trans. Acoust., Speech, Signal Process.* **ASSP-27**, 373–380 (1979).

<sup>7</sup>J. A. Stuller and Nancy Hubing, "New perspective for maximum likelihood time-delay estimation," *IEEE Trans. Signal Process.* **SP-45**, 513–525 (1997).

<sup>8</sup>W. R. Hahn and S. A. Tretter, "Optimum processing for delay-vector estimation in passive signal arrays," *IEEE Trans. Inf. Theory* **IT-19**, 608–614 (1973).

<sup>9</sup>L. C. Ng and Y. Bar-Shalom, "Multisensor multitarget time delay vector estimation," *IEEE Trans. Acoust., Speech, Signal Process.* **ASSP-34**, 669–677 (1986).

<sup>10</sup>W. R. Hann, "Optimum signal processing for passive sonar range and bearing estimation," *J. Acoust. Soc. Am.* **58**, 201–207 (1975).

<sup>11</sup>V. H. MacDonald and P. M. Schultheiss, "Optimum passive bearing estimation in spatially incoherent noise environment," *J. Acoust. Soc. Am.* **46**, 37–43 (1969).

<sup>12</sup>P. E. Stoica and A. Nehorai, "Performance study of conditional and unconditional direction-of-arrival estimation," *IEEE Trans. Signal Process.* **SP-38**, 1783–1795 (1990).

<sup>13</sup>M. A. Doron, H. Messer, and A. J. Weiss, "Maximum likelihood direction finding of wideband sources," *IEEE Trans. Signal Process.* **SP-41**, 411–414 (1993).

<sup>14</sup>M. A. Doron, "Direction Finding of Narrowband and Wideband Sources," Ph.D. thesis, Tel-Aviv University, 1992.

<sup>15</sup>A. Gelb, *Applied Optimal Estimation* (MIT, Cambridge, 1974), p. 103.

<sup>16</sup>NASA Conference Publication, "Joint acoustic propagation experiment (JAPE-91) workshop."

<sup>17</sup>R. O. Nielsen, "Azimuth and elevation angle estimation with a three-dimensional array," *IEEE Trans. Signal Process.* **SP-19**, 84–86 (1994).

# Model-based dispersive wave processing: A recursive Bayesian solution

J. V. Candy and D. H. Chambers

University of California, Lawrence Livermore National Laboratory, P.O. Box 808,  
L-156 Livermore, California 94551

(Received 22 April 1996; accepted for publication 18 February 1999)

Wave propagation through dispersive media represents a significant problem in many acoustic applications, especially in ocean acoustics, seismology, and nondestructive evaluation. In this paper we propose a propagation model that can easily represent many classes of dispersive waves and proceed to develop the model-based solution to the wave processing problem. It is shown that the underlying wave system is nonlinear and time-variable requiring a recursive processor. Thus the general solution to the model-based dispersive wave enhancement problem is developed using a Bayesian *maximum a posteriori* (MAP) approach and shown to lead to the recursive, nonlinear extended Kalman filter (EKF) processor. The problem of internal wave estimation is cast within this framework. The specific processor is developed and applied to data synthesized by a sophisticated simulator demonstrating the feasibility of this approach. © 1999 Acoustical Society of America. [S0001-4966(99)00106-X]

PACS numbers: 43.60.Gk, 43.40.Ph, 43.20.Jr [JCB]

## INTRODUCTION

Dispersive waves in acoustics occur in many varied applications, most notably in seismic exploration when the properties of the borehole are under investigation during logging;<sup>1-3</sup> in ultrasonics, when the condition of material parts are being interrogated for flaws during nondestructive evaluation;<sup>4,5</sup> and in ocean acoustics, when a target is to be detected/localized.<sup>6,7</sup> The need to develop techniques to process waves propagating in a dispersive medium is very important in acoustic applications.

In the seismic borehole logging problem, the dispersion relation is represented linearly in some cases leading to slant stacking and semblance processing<sup>8,9</sup> to extract the desired wave information (velocity, slowness, etc.) and enabling the characterization of the borehole. More current seismic techniques emphasize nonparametric representations of the media to characterize the borehole by its (spatial) wave number frequency (temporal) representation using sophisticated estimators to extract the required information.<sup>10,11</sup> The basic approach is to estimate the wave number frequency spectrum at each narrow-band temporal frequency by finding a set of spatial poles or equivalently complex exponentials. The nonlinear dispersion relation is not modeled explicitly, but captured nonparametrically at each temporal frequency. Note that if an underlying theoretical model of this relation existed, then it could be used to estimate the dispersion directly from the data or spectra characterizing the propagation explicitly.

From the signal processing perspective, some work on wave estimation has been accomplished in estimating wave functions from measured array data. The usual array signal processing approach is based on plane wave assumptions in a homogeneous medium<sup>12</sup> where the received data are decomposed into narrow frequency bands using fast Fourier transforms, the spectral covariance matrix is estimated, and sophisticated subspace processing algorithms are employed to

extract the desired direction-of-arrival information. Alternatively parametric methods have also been developed to solve this problem and have been analyzed thoroughly.<sup>13,14</sup> In Ref. 13 a generic model capturing various pertinent representations related to wave propagation was developed, but the results are implicitly limited to stationary or quasi-stationary processes because of the need to estimate the sample covariance matrix or its equivalent smoothed version. This work establishes the maximum likelihood solution for "fitting" exponential models in terms of prediction error polynomials which can be applied in the dispersive wave case to the nonparametric seismic approach discussed previously. Perhaps the most related work on parametric dispersive wave processing is discussed in Ref. 15 where the physical parameters of dispersive waves are to be estimated from a linear array of measurements. Here the approach is to estimate complex slowness (inverse phase velocity) in temporal frequency bands under additive Gaussian noise assumptions. The wide-band solution is then achieved as an incoherent integration over the set of temporal frequencies under the assumption of uncorrelated noise in each frequency band. This approach explicitly incorporates a dispersion model into the processor and estimates various critical parameters from the measured data.

In ocean acoustics, related problems have also been addressed in a wide variety of applications. The inclusion of a propagation model into any signal processing scheme provides a means of introducing environmental information in a self-consistent manner. Recent work in ocean acoustics<sup>16,17</sup> has shown that a propagation model can be embedded into a signal processing algorithm to solve various enhancement, localization, and detection problems. There a normal mode propagator is selected and parameters such as ocean depth, sound velocity profile, and the ocean bottom conditions are introduced along with acoustic measurements from a vertical array. When the model parameters are unknown and/or

changing, the approach can be extended to be parametrically adaptive enabling the processor to perform well even under these hostile conditions.<sup>18</sup> For instance, the adaptive processor can provide sufficient information to localize the acoustic target directly from the array measurements.<sup>19</sup>

Thus dispersive wave propagation through various media is a significant problem in many acoustic applications. From the scientific viewpoint, many wave-type phenomena must propagate through a hostile, noisy environment that changes rapidly in short periods of time causing great distortions in signal content and leading to false estimates and conclusions. The characteristic common to each of these applications is the frequency dependence of the wave speed. The dispersive nature of a wave system may result from either an intrinsic frequency dependence of bulk properties of the medium or from the geometrical or mechanical properties of the system such as ocean surface waves, internal gravity waves, and wave propagation in a waveguide. Unfortunately explicit dispersion and envelope relations are highly nonlinear; therefore, we must develop an approach that somehow incorporates this *a priori* information, when it is available.<sup>20</sup>

In this paper we approach the dispersive wave processing problem from a different perspective. We assume that we do have an explicit expression for both the dispersion relation and envelope function and we would like to “extract” the propagating wave signal from a set of noisy array measurements. This type of problem can be posed in the model-based signal processing framework<sup>20</sup> in the sense that we have explicit physical information describing the dispersion relation, envelope function, and noise and we would like to use it to extract the desired signal (propagating wave) for detection or imaging. Thus the problem we address in this paper is to find an “optimal” (in a mean-squared error sense) dispersive wave processor using the *a priori* information available. We will show that a model-based signal processing scheme applicable to dispersive wave systems can be developed from the basic properties of wave propagation in a dispersive medium.

Our approach is to develop a state-space description of a dispersive wave measured by a sensor array. The wave pulse is assumed to be generated by an impulsive source at a known position and a known time. We consider the source pulse a superposition of wave components of many frequencies over a broadband. Since the system is dispersive, each component propagates at a different speed, resulting in a spreading or dispersing of the pulse over space and time as it propagates. For broadband pulses this spreading can be quite severe. This spreading is described by the dispersion relation of the system which relates the frequency of each component to its wave number or equivalently in the temporal domain it is represented as a time-varying function. Whitham<sup>21</sup> shows how the dispersion relation can be used to construct a space-time representation of pulse spreading. We utilize his method to formulate a complete state-space representation of the wave in the time domain from the dispersion relation combined with an envelope function. The dispersion relation completely describes the propagation properties of the dispersive system, while the envelope is related to its initial

conditions. Once specified, it is then possible to develop a model-based processing scheme for dispersive waves. This processor evolves directly from the modified plane wave and internal wave techniques developed using an approximation of the dispersion relation reported on previously.<sup>22</sup>

In Sec. I, we first motivate Whitham’s<sup>21</sup> dispersive wave model and then develop the corresponding model-based processor (MBP) using the Bayesian *maximum a posteriori* (MAP) approach to solve the wave estimation problem. Next, we develop it for a specific application—the problem of internal wave estimation in Sec. II. We apply the resulting MBP to synthesized data in Sec. III demonstrating its feasibility, and summarize our results in the final section.

## I. MODEL-BASED DISPERSIVE WAVE PROCESSOR

In this section we develop a nonlinear, model-based processor based on a dispersive wave model developed by Whitham.<sup>21</sup> It is shown that once the envelope and dispersion relation along with the required initial conditions are specified, the solution to the wave estimation problem can be accomplished. Here we develop and motivate the underlying dispersive wave model and then cast it into state-space form. Once this is accomplished, a Bayesian MAP solution is developed for the resulting time-variable wave system and the processor is shown to be the recursive, extended Kalman filter (EKF) solution.<sup>20</sup>

### A. Dispersive state-space propagator

First, we develop the state-space representation of a general dispersive wave system obtained from a simple physical characterization of a dispersive wave measured by a sensor array. Following Whitham,<sup>21</sup> we define a generic *dispersive wave* as any system which admits solutions of the general form

$$u(x,t) = \alpha(x,t) \sin[\theta(x,t)], \quad (1)$$

where  $u$  is the measured field and  $\alpha(x,t)$ ,  $\theta(x,t)$  are the respective instantaneous envelope and phase functions. The phase is assumed to be monotonic in  $x$  and  $t$ , and the envelope is assumed to be slowly varying compared to the phase. This general form incorporates all of the features which are normally associated with wave motion. The phase function describes the oscillatory characteristic and propagation of a wave, while the slowly varying envelope allows modulation and attenuation of the wave without destroying its wavelike character. The local values of instantaneous *wave number* and *frequency* can be defined as

$$\kappa(x,t) \equiv \frac{\partial \theta}{\partial x}, \quad \omega(x,t) \equiv -\frac{\partial \theta}{\partial t}. \quad (2)$$

These functions are also assumed to be slowly varying compared to the local wavelength and period and describe the frequency modulation of a dispersive wave train. By slowly varying we mean that we can approximate the instantaneous *phase function* as

$$\theta(x,t) \approx \kappa(x,t)x - \omega(x,t)t. \quad (3)$$

The combination of Eqs. (1) and (3) can be considered an asymptotic solution to some dispersive wave system. Asymptotic forms of this kind can be formally derived for simple linear dispersive systems using the method of stationary phase to evaluate the exact Fourier integral solution (see Whitham<sup>21</sup> for details).

The instantaneous wave number  $\kappa(x,t)$  and frequency  $\omega(x,t)$  are local generalizations of their usual definitions, which are especially suited for describing dispersive waves. Physically, they describe the modulation of wave number and frequency which a sensor experiences as a dispersive wave train sweeps past. In this sense the instantaneous,  $\kappa(x,t)$  and  $\omega(x,t)$  are natural local extensions of wave number and frequency. The assumption that these are slowly varying does *not* restrict this representation to narrow bandwidth pulses. It does assume that the wave has propagated far enough so that the change in  $\kappa$  and  $\omega$  over a wavelength or period is small. For a detailed discussion of the mathematical basis and range of applicability of this representation we refer the interested reader to Whitham's text.<sup>21</sup>

Finally to complete the specification of a dispersive wave system, we define the *dispersion relation*,  $\omega \equiv \omega(\kappa, x, t)$ . This can be an algebraic function of  $\kappa(x,t)$  as in our subsequent application or it can also depend on  $x$  and  $t$  separately to represent an instantaneous or equivalently *time-varying*, nonuniform wave systems. Here we will write  $\omega = \omega(\kappa)$  where the  $x$  and  $t$  dependence is through the wave number function  $\kappa(x,t)$  and any system nonuniformity is implied. This phase function and the envelope are the only parts of the description which are unique to the particular type of wave system under investigation. The choice of dispersion relation enables the differentiation between acoustic radiation, ocean surface waves, internal gravity waves, or any other wave type. Thus the dispersion relation is equivalent to the governing equations for a particular wave system.<sup>21</sup> Our only restriction on it in this paper is that it is independent of the envelope,  $\alpha(x,t)$ . This restricts our formulation to *linear dispersive waves*. Note also that it can easily be extended to include doppler frequency shifts as well (see Sec. III for details).

From Eqs. (2) and (3) it can be shown that the phase fronts of any wave travel at the *phase speed* defined by

$$c_p(\kappa) \equiv \frac{\omega}{\kappa}, \quad (4)$$

while the points of constant wave number  $\kappa$  travel at the *group velocity* defined by

$$c_g(\kappa) \equiv \frac{\partial \omega}{\partial \kappa}. \quad (5)$$

These two speeds are not the same in general and are functions of wave number  $\kappa$ . The group velocity has the additional significance of being the *energy* propagation speed for the wave system, that is, the energy in the wave packet is carried at this velocity. As such it plays a central role in the state-space formulation of a general dispersive wave system.

Now consider the problem where an impulse occurs at time  $t=0$  at the spatial origin,  $x=0$ . The impulse can be represented by the superposition of wave components with

instantaneous wave numbers. A broadband wave train is generated by the impulse and propagates away from the origin. Each wave number component in the train propagates with group velocity given by Eq. (5). If a sensor is placed at a distance  $x$  away from the origin, then the *local* instantaneous wave number,  $\kappa(x,t)$ , observed at time  $t>0$  is related to position  $x$  by the group velocity

$$x = c_g(\kappa(x,t))t. \quad (6)$$

This relation is simply a restatement of the definition of the instantaneous group velocity as the speed at which a given wave number  $\kappa(x,t)$  propagates in the wave train. Note also that the wavelength is also instantaneous using the Whitham representation, since the whole range of wave numbers is still present, that is,

$$\lambda(x,t) = \frac{2\pi}{\kappa(x,t)}. \quad (7)$$

The actual sensor measurement  $u(t)$  at  $x$  is given by combining Eqs. (1) and (3), that is,

$$u(t) = \alpha(x,t;\kappa) \sin[\kappa(t)x - \omega(\kappa)t], \quad (8)$$

where we have suppressed the dependence of  $\kappa$  on  $x$  and allowed the envelope to be a function of  $\kappa$ . We will choose the wave number  $\kappa(t)$  at  $x$  as our state variable and develop a dynamical equation for its temporal evolution by differentiating Eq. (6) using the chain rule

$$\frac{dc_g(\kappa)}{dt} = \frac{d\kappa}{dt} \times \frac{dc_g(\kappa)}{d\kappa}, \quad (9)$$

to obtain

$$0 = \frac{d\kappa}{dt} \left[ \frac{dc_g(\kappa)}{d\kappa} t \right] + c_g(\kappa). \quad (10)$$

Now solving for  $d\kappa/dt$  and substituting the expression for group velocity of Eq. (5) in terms of our original dispersion relation, we obtain

$$\frac{d\kappa}{dt} = -\frac{1}{t} \left[ \frac{d\omega(\kappa)}{d\kappa} \right] \left[ \frac{d^2\omega(\kappa)}{d\kappa^2} \right]^{-1}; \quad t>0, \quad (11)$$

which shows how the local wave number evolves instantaneously (dynamically) as a function of the underlying dispersion relation  $\omega(\kappa)$ . If we couple this expression back to the original dispersive wave solution, then we have a general continuous-time, spatio-temporal, dispersive wave, state-space representation with state defined by  $\kappa(t)$ .

The initial condition for the state is also obtained from Eq. (6). The state equation actually has no meaning until the leading edge of the wave train passes sensor location  $x$ , which we call the *onset* of the wave train. This occurs at a time  $t_0>0$  specified by the fastest wave component,

$$x = c_g(\kappa(t_0))t_0, \quad (12)$$

where  $c_g(\kappa(t_0))$  is the maximum group velocity for  $\kappa = \kappa(t_0)$ . The initial value of the state is therefore determined by solving this implicit equation for  $\kappa(t_0)$ . This completes our state-space description for a dispersive wave system measured by a single sensor.

Suppose we sample this wave with an array of  $L$ -sensors oriented in the direction of propagation, that is,  $x \rightarrow x_l$ ,  $l = 1, \dots, L$ , giving  $L$  wave numbers and  $L$  initial conditions. If the entire state space is to be initialized at the same time, care must be taken to select the initialization time to be after the leading edge of the wave has passed through the entire array. Let  $t_0$  be the time the leading edge passes the sensor  $L$ , the sensor farthest from the origin, then

$$x_L = c_g(\kappa_L(t_0))t_0, \quad (13)$$

where  $\kappa_L(t_0)$  is the local wave number when the group velocity is maximum. The initial conditions for the other sensors in the array are obtained by solving  $x_l = c_g(\kappa_l(t_0))t_0$  for each  $l$ . Thus the continuous time, ‘‘spatially’’ sampled, dispersive wave, state-space model is given by:

$$\begin{aligned} \frac{d\kappa_l}{dt} &= -\frac{1}{t} \left[ \frac{d\omega(\kappa_l)}{d\kappa_l} \right] \left[ \frac{d^2\omega(\kappa_l)}{d\kappa_l^2} \right]^{-1}, \quad t \geq t_0; \\ u_l(t) &= \alpha(t; \kappa_l) \sin[\kappa_l x_l - \omega(\kappa_l)t], \\ \kappa_l(t_0), \quad &l = 1, \dots, L. \end{aligned} \quad (14)$$

We can further discretize this model temporally by sampling  $t \rightarrow t_k$ , and also by replacing the derivatives with their first difference approximations. Since we know that the dispersive medium in which the wave propagates is uncertain, then we can also characterize uncertainties with statistical models, one of which is the well-known Gauss–Markov model.<sup>20</sup> Performing these operations we achieve our desired result, the discrete, spatio-temporal, *dispersive wave state-space* Gauss–Markov model:

$$\begin{aligned} \kappa_l(t_{k+1}) &= \kappa_l(t_k) - \frac{\Delta t_k}{t_k} \left[ \frac{d\omega(\kappa_l)}{d\kappa_l} \right] \left[ \frac{d^2\omega(\kappa_l)}{d\kappa_l^2} \right]^{-1} \\ &\quad + \Delta t_k w_l(t_k), \quad t_k \geq t_0; \\ u_l(t_k) &= \alpha(t_k; \kappa_l) \sin[\kappa_l(t_k)x_l - \omega(\kappa_l)t_k] + v_l(t_k), \\ \kappa_l(t_0); \quad &l = 1, \dots, L, \end{aligned} \quad (15)$$

where  $w_l(t_k)$  and  $v_l(t_k)$  are assumed zero mean, Gaussian noise sources with respective covariances,  $R_{ww}(t_k)$ ,  $R_{vv}(t_k)$ . The general vector Gauss–Markov form can be found in Ref. 20 and is simply given by the set of nonlinear, time-variable propagation equations as

$$\begin{aligned} \kappa(t_{k+1}) &= \mathbf{a}[\kappa, t_k] + \Delta t_k \mathbf{w}(t_k), \\ \mathbf{u}(t_k) &= \mathbf{c}[\kappa, t_k] + \mathbf{v}(t_k), \end{aligned} \quad (16)$$

where  $\mathbf{a}[\cdot]$ ,  $\mathbf{c}[\cdot]$  are the respective nonlinear vector system and measurement functions with the corresponding state and measurement covariances defined by:

$$\begin{aligned} \mathbf{P}(t_{k+1}) &= \mathbf{A}(\kappa, t_k) \mathbf{P}(t_k) \mathbf{A}^T(\kappa, t_k) + R_{ww}(t_k), \\ R_{uu}(t_k) &= \mathbf{C}(\kappa, t_k) \mathbf{P}(t_k) \mathbf{C}^T(\kappa, t_k) + R_{vv}(t_k), \end{aligned} \quad (17)$$

with the system and measurement jacobians,  $\mathbf{A}[\kappa] \equiv \partial \mathbf{a} / \partial \kappa$  and  $\mathbf{C}[\kappa] \equiv \partial \mathbf{c} / \partial \kappa$ . The subsequent development of our processor will rely on this statistical formulation for both simulation and estimation.

## B. Dispersive model-based processor

Next we develop the model-based processor (MBP) based on the vector representation of the wave numbers and wave field, that is, we define the vectors,  $\mathbf{u}(t_k) \equiv [u_1(t_k), \dots, u_L(t_k)]^T$  and  $\kappa \equiv [\kappa_1(t_k), \dots, \kappa_L(t_k)]^T$ . Once the dispersive wave is characterized by the underlying Gauss–Markov representation, then the *dispersive wave estimation problem* can be specified by:

GIVEN the Gauss–Markov model [Eq. (15)] characterized by the dispersive wave state-space model and a set of noisy measurements,  $\{\mathbf{u}(t_k)\}$ , FIND the best (minimum error variance) estimate of the wave, that is, find  $\hat{\mathbf{u}}(t_k)$ .

The minimum variance solution to this problem can be obtained by the maximizing *a posteriori* density as follows. Define the set of wave field measurements as:  $U_N := \{\mathbf{u}(t_N), \dots, \mathbf{u}(t_0)\}$ ;  $\mathbf{u} \in C^{L \times 1}$ , then the Bayesian maximum *a posteriori* (MAP) estimator of the wave number functions must maximize the posterior density given by

$$\Pr(\kappa(t_{k+1}) | U_{k+1}) = \frac{\Pr(\kappa(t_{k+1}), U_{k+1})}{\Pr(U_{k+1})}. \quad (18)$$

From Bayes’ rule we have that

$$\begin{aligned} \Pr(\kappa(t_{k+1}) | U_{k+1}) &= \frac{\Pr(\mathbf{u}(t_{k+1}) | \kappa(t_{k+1}), U_k) \times \Pr(\kappa(t_{k+1}) | U_k)}{\Pr(\mathbf{u}(t_{k+1}) | U_k)}. \end{aligned} \quad (19)$$

Under the approximate Gauss–Markov assumptions of the previous section, these densities can be specified by

$$\Pr(\mathbf{u}(t_{k+1}) | U_k) \sim \mathcal{N}(\hat{\mathbf{u}}(t_{k+1|k}), \mathbf{R}_{\epsilon\epsilon}(t_k)); \quad (20)$$

$$\Pr(\mathbf{u}(t_{k+1}) | \kappa(t_{k+1}), U_k) \sim \mathcal{N}(\mathbf{c}[\kappa(t_k)], \mathbf{R}_{vv}(t_k)); \quad (21)$$

and

$$\Pr(\kappa(t_{k+1}) | U_k) \sim \mathcal{N}(\hat{\kappa}(t_{k+1|k}), \tilde{\mathbf{P}}(t_{k+1|k})), \quad (22)$$

where the estimation error [Eq. (23)] covariance is defined by  $\tilde{\mathbf{P}}(t_{k+1|k}) \equiv \text{Cov}(\tilde{\kappa})$  and is given by

$$\begin{aligned} \tilde{\mathbf{P}}(t_{k+1|k}) &= \mathbf{A}(\hat{\kappa}(t_{k+1|k})) \tilde{\mathbf{P}}(t_{k+1|k}) \mathbf{A}^T(\hat{\kappa}(t_{k+1|k})) \\ &\quad + \mathbf{R}_{ww}(t_k). \end{aligned} \quad (23)$$

Here the notation  $\hat{\kappa}(t_{k+1|k}) \equiv E\{\kappa(t_{k+1}) | U_k\}$  is the conditional mean, that is, the ‘‘best’’ (minimum variance) estimate at time  $t_{k+1}$  based on the previous measurements up to time  $t_k$ . The symbol  $\sim$  means ‘‘distributed as’’ and  $\mathcal{N}(m, v)$  is a gaussian distribution of mean  $m$  and variance  $v$ .

Now substituting these densities into the *a posteriori* and performing the necessary manipulations (see Ref. 20 for details), we obtain

$$\begin{aligned} \Pr(\kappa(t_{k+1}) | U_{k+1}) &= C_l \exp\left\{ -\frac{1}{2} (\mathbf{v}(t_{k+1}) \mathbf{R}_{vv}^{-1}(t_{k+1}) \mathbf{v}^T(t_{k+1}) \right. \\ &\quad \left. + \tilde{\kappa}^T(t_{k+1|k}) \tilde{\mathbf{P}}^{-1}(t_{k+1|k}) \tilde{\kappa}(t_{k+1|k}) \right. \\ &\quad \left. - \epsilon^T(t_{k+1}) \mathbf{R}_{\epsilon\epsilon}^{-1}(t_{k+1}) \epsilon(t_{k+1}) \right\}, \end{aligned} \quad (24)$$

where the *wave number estimation error* vector is defined by

TABLE I. Dispersive wave: Discrete extended Kalman filter algorithm.

Prediction	
$\hat{\kappa}(t_{k+1} t_k) = \mathbf{a}[\hat{\kappa}(t_{k+1} t_k)]$	(state prediction)
$\tilde{\mathbf{P}}(t_{k+1} t_k) = \mathbf{A}[\mathbf{g}\tilde{\mathbf{R}}(t_{k+1} t_k)]\tilde{\mathbf{P}}(t_k t_k)\mathbf{A}^T[\hat{\kappa}(t_{k+1} t_k)] + \mathbf{R}_{ww}(t_k)$	(covariance prediction)
Innovation	
$\epsilon(t_k) = \mathbf{u}(t_k) - \hat{\mathbf{u}}(t_{k+1} t_k) = \mathbf{u}(t_k) - \mathbf{c}[\hat{\kappa}(t_{k+1} t_k)]$	(innovation)
$\mathbf{R}_{\epsilon\epsilon}(t_k) = \mathbf{C}[\hat{\kappa}(t_{k+1} t_k)]\tilde{\mathbf{P}}(t_{k+1} t_k)\mathbf{C}^T[\hat{\kappa}(t_{k+1} t_k)] + \mathbf{R}_{vv}(t_k)$	(innovation covariance)
Gain	
$\mathbf{K}(t_k) = \tilde{\mathbf{P}}(t_{k+1} t_k)\mathbf{C}^T[\hat{\kappa}(t_{k+1} t_k)]\mathbf{R}_{\epsilon\epsilon}^{-1}(t_k)$	(Kalman gain or weight)
Correction	
$\hat{\kappa}(t_{k+1} t_{k+1}) = \hat{\kappa}(t_{k+1} t_k) + \mathbf{K}(t_{k+1})\epsilon(t_{k+1})$	(state correction)
$\tilde{\mathbf{P}}(t_{k+1} t_{k+1}) = [\mathbf{I} - \mathbf{K}(t_{k+1})\mathbf{C}[\hat{\kappa}(t_{k+1} t_k)]]\tilde{\mathbf{P}}(t_{k+1} t_k)$	(covariance correction)
Initial conditions	
$\hat{\kappa}(0 0) \quad \tilde{\mathbf{P}}(0 0) \quad \mathbf{A}[\hat{\kappa}(t_{k+1} t_k)] \equiv \frac{\partial}{\partial \kappa} \mathbf{a}[\kappa] \Big _{\kappa=\hat{\kappa}(t_{k+1} t_k)} \quad \mathbf{C}[\hat{\kappa}(t_{k+1} t_k)] \equiv \frac{\partial}{\partial \kappa} \mathbf{c}[\kappa] \Big _{\kappa=\hat{\kappa}(t_{k+1} t_k)}$	

$$\tilde{\kappa}(t_{k+1}|k) \equiv \kappa(t_k) - \hat{\kappa}(t_{k+1}|k), \quad (25)$$

and the corresponding vector *innovation* is

$$\epsilon(t_k) \equiv \mathbf{u}(t_k) - \hat{\mathbf{u}}(t_{k+1}|k), \quad (26)$$

with the enhanced (MAP estimated) wave-field measurement given by

$$\begin{aligned} \hat{\mathbf{u}}(t_{k+1}|k) &= \mathbf{c}[\hat{\kappa}(t_{k+1}|k)] \\ &= \alpha[(\hat{\kappa}(t_{k+1}|k), t_k)] \\ &\quad \times \sin[(\hat{\kappa}(t_{k+1}|k)x - \omega(\hat{\kappa}(t_{k+1}|k))t_k], \end{aligned} \quad (27)$$

for the constant  $C_l$  with respective innovations covariance matrix,  $\mathbf{R}_{\epsilon\epsilon}(t_k)$ . By maximizing the *a posteriori* density or equivalently its logarithm, we have the so-called *MAP equation*,

$$\frac{\partial}{\partial \kappa} \ln \Pr(\kappa(t_{k+1})|U_{k+1}) \Big|_{\kappa=\hat{\kappa}_{\text{MAP}}} = \mathbf{0}. \quad (28)$$

Differentiating the posterior density and noting that  $\hat{\kappa}(t_{k+1}|k)$  and  $\epsilon(t_k)$  are both functions of the data set,  $U_k$ , we obtain

$$\begin{aligned} &\frac{\partial}{\partial \kappa} \ln \Pr(\kappa(t_{k+1})|U_{k+1}) \\ &= \mathbf{C}(\hat{\kappa}(t_{k+1}|k))\mathbf{R}_{vv}^{-1}(t_{k+1})(\mathbf{u}(t_{k+1}) \\ &\quad - \mathbf{C}^T(\hat{\kappa}(t_{k+1}|t_k))\kappa(t_{k+1})) - \tilde{\mathbf{P}}^{-1}(t_{k+1}|k)(\kappa(t_{k+1}) \\ &\quad - \hat{\kappa}(t_{k+1}|k)) = \mathbf{0}. \end{aligned} \quad (29)$$

Solving the MAP equation for  $\kappa = \hat{\kappa}_{\text{MAP}}$  yields

$$\begin{aligned} \hat{\kappa}_{\text{MAP}}(t_{k+1}) &= \tilde{\mathbf{P}}(t_{k+1}|k+1)[\tilde{\mathbf{P}}^{-1}(t_{k+1}|k)\hat{\kappa}(t_{k+1}|k) \\ &\quad + \mathbf{C}^T(\hat{\kappa}(t_{k+1}|k))\mathbf{R}_{vv}^{-1}(t_{k+1})\mathbf{u}(t_{k+1})], \end{aligned} \quad (30)$$

where we have used the fact that

$$\begin{aligned} \tilde{\mathbf{P}}(t_{k+1}|k+1) &= [\mathbf{C}(\hat{\kappa}(t_{k+1}|k))\mathbf{R}_{vv}^{-1}(t_{k+1})\mathbf{C}^T(\hat{\kappa}(t_{k+1}|k)) \\ &\quad + \tilde{\mathbf{P}}^{-1}(t_{k+1}|k)]^{-1}, \end{aligned} \quad (31)$$

from the matrix inversion lemma.<sup>23</sup> It has been shown<sup>20</sup> that the gain or weighting matrix can be written as

$$\mathbf{K}(t_{k+1}) = \tilde{\mathbf{P}}(t_{k+1}|k+1)\mathbf{C}^T(\hat{\kappa}(t_{k+1}|k))\mathbf{R}_{vv}^{-1}(t_{k+1}) \quad (32)$$

and therefore performing the indicated multiplication in Eq. (30) and using this gain expression we obtain the MAP solution

$$\begin{aligned} \hat{\kappa}_{\text{MAP}}(t_{k+1}) &= \hat{\kappa}(t_{k+1}|k) - \mathbf{K}(t_{k+1})\mathbf{c}[\hat{\kappa}(t_{k+1}|k)] \\ &\quad + \mathbf{K}(t_{k+1})\mathbf{u}(t_{k+1}). \end{aligned} \quad (33)$$

Thus we see that  $\hat{\kappa}_{\text{MAP}}(t_{k+1}) = \hat{\kappa}(t_{k+1}|k+1)$  evolves as (see Ref. 20, pp. 80–81 for more details)

$$\hat{\kappa}(t_{k+1}|k+1) = \hat{\kappa}(t_{k+1}|k) + \mathbf{K}(t_{k+1})\epsilon(t_{k+1}), \quad (34)$$

where this expression is the corrected estimate [Eq. (37)] and shown in the algorithm of Table I. Thus the model-based solution to this wave enhancement problem can be achieved using the nonlinear extended Kalman filter (EKF) algorithm which is given (simply) as:

*Prediction:*

$$\begin{aligned} \hat{\kappa}_l(t_{k+1}|k) &= \hat{\kappa}_l(t_k|k) - \frac{\Delta t_k}{t_k} \left[ \frac{d\omega(\hat{\kappa}_l)}{d\kappa_l} \right] \left[ \frac{d^2\omega(\hat{\kappa}_l)}{d\kappa_l^2} \right]^{-1}, \\ &\quad l = 1, \dots, L; \end{aligned} \quad (35)$$

*Innovation:*

$$\begin{aligned} \epsilon(t_k) &= \mathbf{u}(t_k) - \hat{\mathbf{u}}(t_{k+1}|k), \\ \hat{\mathbf{u}}_l(t_{k+1}|k) &= \alpha(t_k; \hat{\kappa}_l(t_{k+1}|k)) \sin[\hat{\kappa}_l(t_{k+1}|k)x_l \\ &\quad - \omega(\hat{\kappa}_l)t_k], \quad l = 1, \dots, L; \end{aligned} \quad (36)$$

*Correction:*

$$\hat{\kappa}(t_{k+1|k+1}) = \hat{\kappa}(t_{k+1|k}) + \mathbf{K}(t_k) \boldsymbol{\epsilon}(t_k); \quad (37)$$

Gain:

$$\mathbf{K}(t_k) = \tilde{\mathbf{P}}(t_{k+1|k}) \mathbf{C}^T(t_k) \mathbf{R}_{\epsilon\epsilon}^{-1}(t_k). \quad (38)$$

Here the predicted and corrected covariances are in Table I. From the table we see that in order to construct the optimal dispersive wave model-based processor, we must not only specify the required initial conditions, but also the respective system and measurement Jacobians: ( $da[\bullet]/d\kappa$ ) and ( $dc[\bullet]/d\kappa$ ). For our general solution, we see from Eq. (15) that

$$a[\kappa_l, t_k] = \kappa_l(t_k) - \frac{\Delta t_k}{t_k} \left[ \frac{d\omega(\kappa_l)}{d\kappa_l} \right] \left[ \frac{d^2\omega(\kappa_l)}{d\kappa_l^2} \right]^{-1}, \quad (39)$$

$$c[\kappa_l, t_k] = \alpha(t_k; \kappa_l) \sin[\kappa_l(t_k)x_l - \omega(\kappa_l)t_k],$$

$$l = 1, \dots, L.$$

The Jacobians then follow easily as

$$A[\kappa_l(t_k), t_k] = 1 - \frac{\Delta t_k}{t_k} \left( 1 - \frac{\left[ \frac{d\omega(\kappa_l)}{d\kappa_l} \right] \left[ \frac{d^3\omega(\kappa_l)}{d\kappa_l^3} \right]}{\left[ \frac{d^2\omega(\kappa_l)}{d\kappa_l^2} \right]^2} \right), \quad (40)$$

$$l = 1, \dots, L;$$

$$C[\kappa_l(t_k), t_k] = \frac{d\alpha_l(t_k)}{d\kappa_l} \sin[\kappa_l(t_k)x_l - \omega(\kappa_l)t_k]$$

$$+ \alpha_l(t_k) \cos[\kappa_l(t_k)x_l - \omega(\kappa_l)t_k]$$

$$\times \left( x_l - \frac{d\omega(\kappa_l)}{d\kappa_l} t_k \right).$$

The last factor in the measurement Jacobian expression can be rewritten as  $x_l - c_g(\kappa_l(t_k))t_k$ . This factor is identically zero for a  $\kappa_l(t_k)$  that satisfies Eq. (6). However, in the Gauss–Markov model, the evolution equation was modified by adding a random noise term. Thus the factor in the measurement Jacobian consists of the time integral of the random forcing term,  $w_l(t_k)$ , in addition to the error caused by the time discretization.

Before we complete this section, let us consider a simple example and see how it fits into this framework. Let us assume that we have a simple, *nondispersive*, one-dimensional monochromatic wave model representing an acoustic source in the far field (e.g., direction finding problem), then we have

$$u(x, t) = \alpha \sin[\kappa_0 x - \omega_0 t], \quad (41)$$

where the frequency,  $\omega_0 = \kappa_0 \times c$  and  $c$  the velocity in the medium. Clearly, the phase and group velocities are identical, since

$$c_p(\kappa) = \frac{\omega_0}{\kappa_0} = c, \quad c_g(\kappa_0) = \frac{d\omega_0}{d\kappa} = c. \quad (42)$$

In this example, we use the fact that the wave number is a constant ( $d\kappa/dt = 0$ ), therefore, our propagation model in the *discrete* state-space framework is simply given by

$$\kappa_l(t_{k+1}) = \kappa_l(t_k),$$

$$u_l(t_k) = \alpha \sin[\kappa_l x_l - \omega_0 t_k]; \quad l = 1, \dots, L. \quad (43)$$

Correspondingly, the model-based processor then for this problem takes on the simplified form:

$$\hat{\kappa}(t_{k+1|k+1}) \equiv \hat{\kappa}(t_{k+1|k}) + \mathbf{K}(t_k) [\mathbf{u}(t_k) - \hat{\mathbf{u}}(t_{k+1|k})], \quad (44)$$

where the optimal wave estimate at the  $l$ -th sensor is

$$\hat{u}_l(t_{k+1|k}) = \alpha \sin[\hat{\kappa}_l(t_{k+1|k})x_l - \omega_0 t_k], \quad (45)$$

and the required Jacobians are simply

$$A[\kappa_l, t] = 1, \quad C[\kappa_l, t] = x_l - ct_k. \quad (46)$$

It is interesting to note that if we assume a source at bearing angle  $\theta$ , then the component wave number along the array direction is  $\kappa_l = \kappa_0 \sin \theta_0$  and this processor becomes equivalent to that employed in Ref. 7 for source tracking. This completes the section on dispersive wave estimation; next we consider the application of this processor for internal waves.

## II. INTERNAL WAVE DISPERSIVE PROCESSOR

When operating in a stratified environment with relatively sharp density gradients any excitation that disturbs the pycnocline (density profile) will generate *internal waves*.<sup>24</sup> These waves are volume gravity waves and can be generated from tidal flow against islands, sea mounts, and continental shelf edges or surface/internal wave interactions created displacements in the pycnocline. A ship traveling along the surface of a stratified ocean creates various visible wakes: the turbulent or centerline wake, the Kelvin wake, and surface generated internal waves in some cases. Internal waves have been measured experimentally both in controlled environments as well as the open ocean<sup>25</sup> and observed using synthetic aperture radar processing techniques from satellite imagery.<sup>26,27</sup> From the scientific viewpoint, it is of high interest to understand the effect of internal waves on acoustic propagation in the ocean<sup>6</sup> as well as the ability to measure their effect directly using current sensor technology.

To apply our processor to the internal wave enhancement problem, we first revisit the original dispersive wave system

$$u(x, t) = \alpha(x, t) \sin[\theta(x, t)], \quad (47)$$

and apply this structure to the internal wave dynamics where  $u$  represents the measured *velocity field* and  $\alpha(x, t)$  and  $\theta(x, t)$  are the respective envelope and phase to be specified by the internal wave structure. As before,

$$\theta(x, t) = \kappa(x, t)x - \omega(\kappa)t. \quad (48)$$

To complete the specification of a dispersive wave system, we define the *internal wave dispersion relation* by

$$\omega \equiv \omega_0(\kappa) + \kappa(x, t)v, \quad (49)$$

where we have included the additive velocity term to account for the effects of a Doppler shift created by the ambient current and the ambient current velocity is defined by  $v$ . That is, in the original formulation we have replaced the position



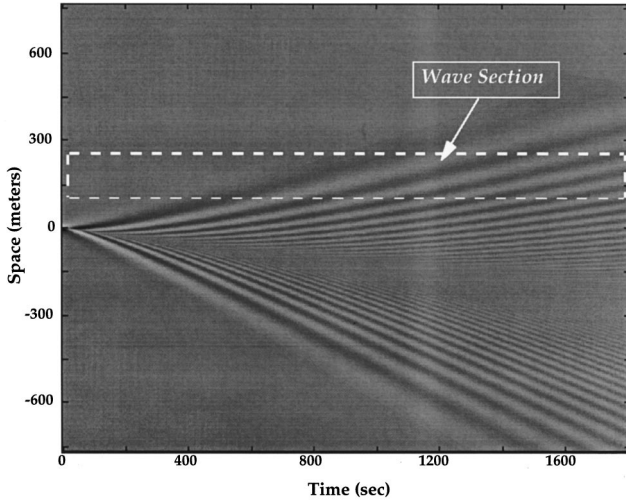


FIG. 1. Synthesized internal wave 1024×361 image ( $C_0=0.3418$  m,  $N_0=0.1369$  r/s,  $L=1024$ ) with processing section displayed.

variable by  $x \rightarrow x - \kappa(x, t)v$  which leads to the above equation. For  $\omega_0(k)$ , we use a dispersion model based on some empirical results, the Barber approximation,<sup>28</sup> for internal wave dispersion and group velocity. Thus we have

$$\omega_0(\kappa) = \frac{C_0 \kappa(t)}{1 + \frac{C_0}{N_0} \kappa(t)}, \quad (50)$$

with  $C_0$  is the initial phase velocity and  $N_0$  is the maximum of the Brunt–Väisälä frequency profile. Using the method of stationary phase to evaluate the Fourier integral expansion for internal waves radiated from a moving source,<sup>29,30</sup> it is possible to derive the following<sup>21</sup> approximation to the amplitude modulation function as

$$\alpha(t_k) = \frac{A}{\sqrt{t}} [c_p(\kappa)]^{3/2} \sin[\omega(\kappa, t_k) T_w], \quad (51)$$

where  $A$  is a constant amplitude governing the overall envelope gain,  $T_w$  is a temporal window width, and  $c_p(\kappa)$  is the phase speed defined below.

Using the simulator,<sup>29</sup> we generate a dense spatio-temporal velocity field. The noise-free synthesis of the field created by Doppler shifted internal wave is shown in Fig. 1. Here the spatial dimension is 1024 samples spaced at  $\Delta x = 1.5$  m an aperture of 1536 m with the corresponding temporal dimensions of 361 samples at  $\Delta t = 5$  s representing a propagation time of approximately  $\frac{1}{2}$  h. For this internal wave simulation, we choose a peak Brunt–Väisälä frequency of  $N_0 = 0.137$  r/s and a long wave ( $\kappa = 0$ ) phase speed of  $C_0 = 0.34$  m/s with the ambient current (Doppler shift) of  $v = -0.1$  cm/s. For our subsequent investigation we select a section of the velocity field annotated on the figure and extracted for display in Fig. 2 to illustrate the section both as an image and the equivalent spatio-temporal representation.

Now in terms of these relations we have the phase speed given in terms of the above dispersion relation

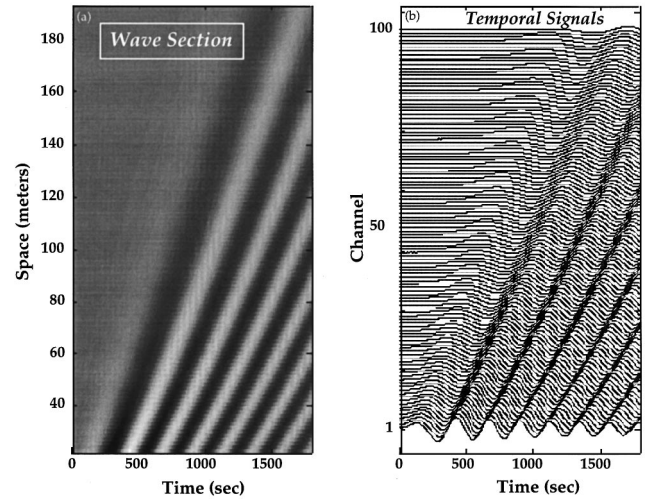


FIG. 2. Internal wave synthesis ( $C_0=0.3418$  m,  $N_0=0.1369$  r/s,  $L=100$ ): (a) 100×361 image section for processing; (b) equivalent spatio-temporal representation.

$$c_p(\kappa) = \frac{C_0}{1 + \frac{C_0}{N_0} \kappa(t)} + v, \quad (52)$$

and the corresponding group velocity defined given by

$$c_g(\kappa) = c_{g_0}(\kappa) + v = \frac{C_0}{\left(1 + \frac{C_0}{N_0} \kappa(t)\right)^2} + v, \quad (53)$$

where we define  $c_{g_0} \equiv (\partial \omega_0 / \partial \kappa)$  and  $v$  is the ambient current velocity. As before, the actual sensor measurement  $u(t)$  at  $x$  is given by

$$u(t) = \alpha(x, t) \sin[\kappa(t)x - (\omega_0(\kappa) + \kappa(t)v)t]. \quad (54)$$

With this information in hand, we can now specify the required dispersive wave state-space model of Eq. (15) as:

$$\begin{aligned} \kappa_l(t_{k+1}) &= \kappa_l(t_k) + \frac{\Delta t_k}{t_k} \frac{N_0}{2} \frac{1 + \frac{v}{c_{g_0}(\kappa_l)}}{\sqrt{C_0 c_{g_0}(\kappa_l)}}, \quad t_k \geq t_0; \\ u_l(t_k) &= \alpha_l(t_k) \sin[\kappa_l(t_k)x_l - (\omega_0(\kappa_l) + \kappa_l(t_k)v)t_k]; \\ l &= 1, \dots, L. \end{aligned} \quad (55)$$

Substituting for  $c_g$  and solving for the wave number at the  $l$ th sensor, we obtain the initial conditions as

$$\kappa_l(t_0) = \frac{N_0}{C_0} \left( \sqrt{\frac{x_L}{x_l \left(1 + \frac{v}{C_0}\right)} - 1} - \frac{v}{C_0} x_L \right). \quad (56)$$

Now all that remains to completely define the MAP processor is the required system and measurement Jacobians. From the internal wave dynamics above we have that:

$$a[\kappa_l, t_k] = \kappa_l(t_k) + \frac{\Delta t_k N_0}{t_k} \frac{1 + \frac{v}{c_{g_0}(\kappa_l)}}{\sqrt{C_0 c_{g_0}(\kappa_l)}}, \quad t_k \geq t_0; \quad (57)$$

$$c[\kappa_l, t_k] = \alpha_l(t_k) \sin[\kappa_l(t_k)x_l - (\omega_0(\kappa_l) + \kappa(t_k)v)t_k].$$

The Jacobians then follow easily from above as:

$$A[\kappa_l, t_k] = 1 - \frac{\Delta t_k}{t_k} \left[ 1 - \frac{3}{2} \frac{c_{g_0}(\kappa_l) + v}{c_{g_0}(\kappa_l)} \right], \quad t_k \geq t_0; \quad (58)$$

$$C[\kappa_l, t_k] = \frac{d\alpha_l(t_k)}{d\kappa_l} \sin \left[ \kappa_l(t_k)x_l - (\omega_0(\kappa_l) + \kappa(t_k)v)t_k + \alpha_l(t_k) \right] \times \cos \left[ \kappa_l(t_k)x_l - (\omega_0(\kappa_l) + \kappa(t_k)v)t_k \right] \times \left( x_l - \left( \frac{d\omega_0(\kappa_l)}{d\kappa_l} + v \right) t_k \right),$$

with the required derivatives available from the internal wave dispersion and envelope functions as:

$$\frac{d\omega(\kappa_l)}{d\kappa_l} = \frac{C_0}{\left(1 + \frac{C_0}{N_0} \kappa_l\right)^2} + v, \quad (59)$$

$$\frac{d\alpha_l(t_k)}{d\kappa_l} = \frac{A}{C_0 \sqrt{t}} \left( \frac{\omega(\kappa_l)}{\kappa_l} \right)^{5/2} \left[ T_w \left( \frac{\omega(\kappa_l)}{\kappa_l} \right) \right] \times \cos(\omega(\kappa_l)T_w) - \frac{3}{2} \frac{C_0}{N_0} \sin(\omega(\kappa_l)T_w).$$

Thus the MAP estimator can now be constructed using the formulation of the previous section and substituting the internal wave model and above noted Jacobians. Next we consider the application of the processor to synthesized internal wave systems and evaluate its performance. Here we evaluate the performance of the model-based processor applied to noisy internal wave dynamics. We appeal to recursive estimation theory<sup>20,23</sup> to assess the overall performance of the processor, that is, we perform basic statistical tests that are used to indicate whether or not the processor is “tuned” and then observe the enhancement achieved. The data used in this study are synthesized initially by a sophisticated internal wave processor<sup>29</sup> which has been used to both design experiments and analyze (post facto) the phenomenology observed.

We use the simulator to synthesize internal wave dynamics corresponding to an internal wave field experiment performed in Loch Linnhe, Scotland in 1994.<sup>31,32</sup> Loch Linnhe is a narrow, salt water estuary located on the west coast of Scotland which possesses favorable stratification conducive for the propagation of internal waves. The internal waves are generated by surface ships which are typically imaged by radar systems employing synthetic aperture radar techniques and real aperture radar imaging of the waves from low-grazing angle marine radars positioned on the surrounding hillsides as well as an array of current meters providing measurements of hydrodynamic currents associated with the

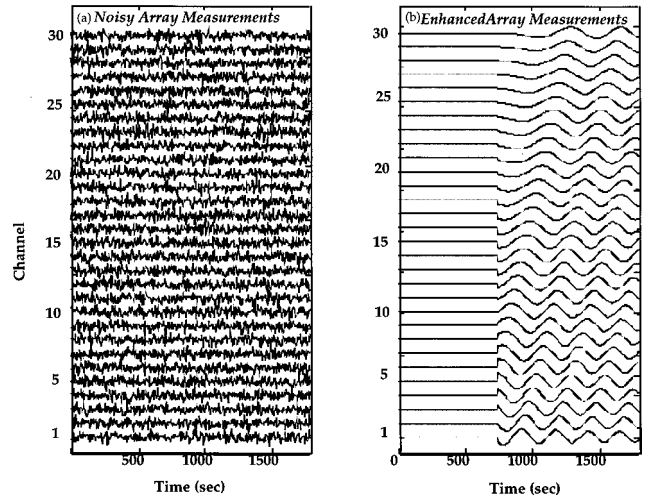


FIG. 3. Equivalent model-based internal wave spatio-temporal enhancement: (a) synthesized wave ( $-23$  dB SNR); (b) enhanced internal wave.

internal waves. The general objective of this experiment was to examine the relationship between modulations observed in radar images and ship generated, internal waves by fusing the radar data with surface currents measured by an array of current meter sensors.

The simulation (shown in Figs. 1 and 2) was performed based on the SNR defined by:

$$\text{SNR} := \sigma^2 / R_{vv},$$

where  $\sigma^2$  is the energy in the true image (scaled to unit variance) and  $R_{vv}$  is the measurement noise variance extracted from experimental data. For this feasibility study, we resample the synthesizer wave section even further primarily motivated by decreasing the overall computational processing time. We decided that the final image or equivalently spatio-temporal velocity field would include the entire range of temporal samples, but spatially we assume a line array of 30 sensors (enough to illustrate the wave structure) spaced at  $\Delta x = 4$  m representing an aperture of 120 m. The data were contaminated with additive Gaussian noise at a  $-23$  dB SNR and shown in Fig. 3(a). Here we observe the effect of additive Gaussian noise in obscuring (visually) the internal wave dynamics of the synthesized velocity field. We see the noisy internal wave spatio-temporal signals synthesized, while the enhanced wave estimates are shown in Fig. 3(b). The results appear quite good, however, they must be analyzed from the statistical perspective to actually assess the overall performance of the processor.

### A. Minimum variance design

Once the Gauss–Markov simulation is complete, the MBP is designed to achieve a minimum variance estimate in the linear model case or a best mean-squared error estimate in the nonlinear model case. If the models used in the processor are “exactly” those used to perform the simulation and synthesize the design data, then the minimum variance or best mean-squared error estimates result. For the nonlinear case, the estimates are deemed approximate minimum variance because the EKF processor uses linearization tech-

niques which approximate the nonlinear functions in order to obtain the required estimates. We use the results of the Gauss–Markov simulations to bound the overall expected performance of the processor on real data. We essentially use simulations at various SNRs to obtain a “feel” for learning how to tune (adjusting noise covariances, etc.) the processor using our particular model sets and parameters.

Once we have completed the simulations and studied the sensitivity of the processor to various parameter values, we are ready to attack the actual data set. With real data, the processor can only perform as well as the dynamical models used represent the underlying phenomenology generating the data. Poor models used in the processor can actually cause the performance to degrade substantially and enhancement may not be possible at all in contrast to the simulation step where we have assured that the model “faithfully” represents the data. In practice, it is never possible to accurately model everything. Thus the goal of minimum variance design is to achieve as close to the optimal design as possible by investigating the consistency of the processor relative to the measured data. This is accomplished by utilizing the theoretical properties of the processor,<sup>20,23</sup> that is, the processor is deemed *optimal*, if and only if, the residual/innovations sequence is zero mean and statistically white or uncorrelated. This approach to performance analysis is much like the results in time series/regression analysis which implies that when the model “explains” or fits the data, no information remains and the residuals are uncorrelated. Therefore, when applying the processor to real data, it is necessary to adjust or “tune” the model parameters (usually the elements of the process noise covariance matrix,  $\mathbf{R}_{ww}$ , see Ref. 20 for details) until the innovations are zero mean/white. If it is not possible to achieve this property, then the models are deemed inadequate and must be improved by incorporating more of the phenomenology. The important point here is that the model-based schemes enable the modeler to assess how well the model is performing on real data and decide where it may be improved. For instance, for the experimental data, we can statistically test the innovations and show that they are white, but when we visually observe the sample correlation function estimate used in the test, it is clear that there still remains some correlation in the innovations. This leads us, as modelers, to believe that we have not captured all of the phenomenology that has generated the data and therefore we must improve the model or explain why the model is inadequate.

Care must be taken when using these statistical tests, because if the models are nonlinear or nonstationary, then the usual whiteness/zero-mean tests, that is, testing that 95% of the sample (normalized) innovation correlations lie within the bounds given by

$$\left[ \hat{c}_{ee}(k) \pm \frac{1.96}{\sqrt{N}} \right], \quad \hat{c}_{ee}(k) = \frac{R_{ee}(k)}{R_{ee}(0)}, \quad (60)$$

and testing for zero mean as

$$\left[ \hat{m}_e(k) < 1.96 \sqrt{\frac{\hat{R}_{ee}(k)}{N}} \right], \quad (61)$$

rely on quasi-stationary assumptions and sample statistics to estimate the required correlations. However, it can be argued heuristically that when the estimator is tuned, the nonstationarities are being tracked by the MBP even in the nonlinear case and therefore, the innovations should be covariance stationary and the tests are valid. If the data are nonstationary then a more reliable statistic to use is the *weighted sum-squared residual* (WSSR) which is a measure of the overall global estimation performance for the MBP processor, determining the “whiteness” of the innovations sequence.<sup>20</sup> It essentially aggregates all of the information available in the innovation vector and tests whiteness by requiring that the decision function lies below the specified threshold to be deemed statistically white. If the WSSR statistic does lie beneath the calculated threshold, then theoretically the estimator is tuned and said to converge. Thus overall performance of the processor can be assessed by analyzing the statistical properties of the innovations which is essentially the approach we take in this feasibility test for the minimum variance design. There are other tests along with these that can be used with real data to check the consistency of the processor and we refer the reader to Ref. 20, Chapter 5 for more details. It should also be noted that the MBP is tuned primarily by adjusting the process noise covariance matrix ( $R_{ww}$ ) and testing the resulting innovations sequence. Thus we limit our design procedure on synthesized data to tuning and then statistically testing the innovations for the zero-mean/whiteness properties.

## B. Model-based wave estimator design

We used SSPACK\_PC, a model-based signal processing toolbox available in MATLAB to design the processor.<sup>33,34</sup> The model-based wave estimator is able to extract the internal wave signatures quite effectively even though the embedded dispersive wave model is just an approximation to the actual wave dynamics. We show the spatio-temporal interpretation (and display) of the noisy and enhanced signals in Fig. 3. To confirm the processor performance that we observed visually, we perform individual whiteness tests on each of the temporal sensor outputs. A set of statistical zero-mean/whiteness tests for this particular simulation are shown in Figs. 4–5, where we have chosen the sensor outputs: 1, 30 to give a feel for the processor performance across the array. In each figure we show the performance at that sensor. First, we observe the wave number (state) estimate (solid line) along with the true wave number (dashed line) from the synthesizer. Clearly, the estimates overlay the true wave number dynamics. Next, we see the estimated velocity field (solid line) at the sensor and that provided by the synthesizer (dashed line). As expected, the estimates are not perfect, lacking enough of the detailed dynamics to generate the initial onset and propagation along with some phase differences; however, the estimates based on the dispersive model of Eq. (47) appear adequate for this application. Next, we observe the corresponding sensor innovations indicating the difference between the noisy and predicted measurement along with the bounds predicted by the processor. Here we note that 95% of the innovations should lie within the bounds indicating a reasonable mini-

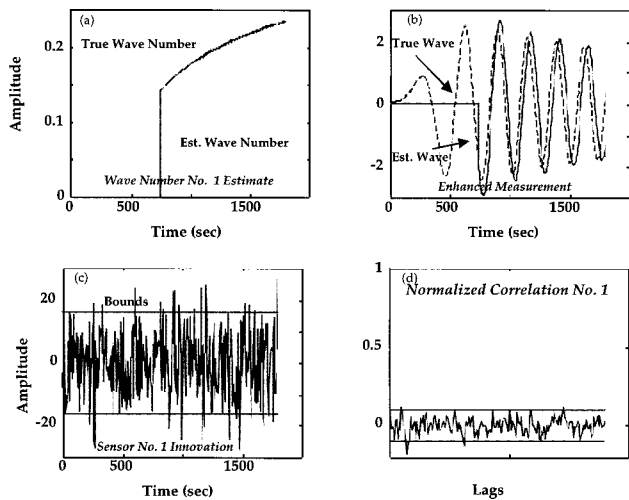


FIG. 4. Model-based internal wave estimation: (a) estimated (solid line) and true (dashed line) wave number no. 1; (b) estimated (solid line) and true (dashed line) internal wave; (c) residual/innovation (8.9% out); (d) zero-mean/whiteness testing ( $0.02 < 1.5$ , 3.9% out).

imum variance design. In this case the samples exceed the bounds, somewhat indicating the processor could probably be tuned even better, but for demonstration purposes these results are satisfactory. Finally the corresponding zero-mean/whiteness tests for the selected sensors are shown. As mentioned these statistical tests are necessary to assess the overall performance of the processor. In these representative realizations each of the sensor innovations are deemed statistically *white*, that is, 95% of the sample correlations lie within or equivalently 5% fall outside the bounds (shown in the figure) and *zero mean*, estimated mean is less than the calculated bound (see figure caption). Performing these tests on all of the sensor innovation outputs revealed that each individually satisfied these statistical tests, indicating a zero-mean/white innovations and a near-optimal processor. However, all that this indicates is that the dispersive wave model is robust enough to capture the dominant internal wave dynamics thereby indicating the potential for more practical use

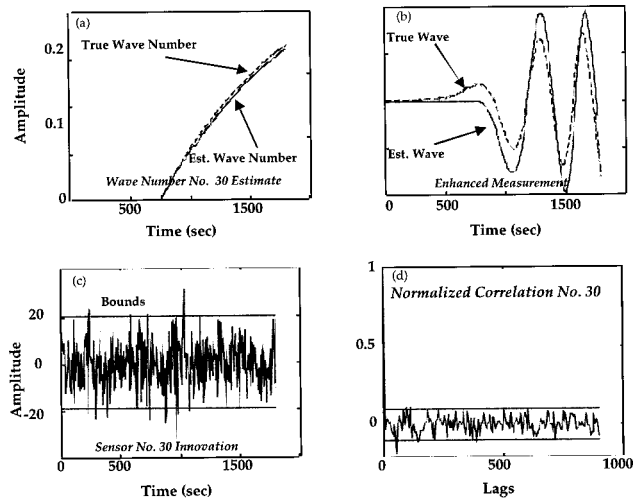


FIG. 5. Model-based internal wave estimation: (a) estimated (solid line) and true (dashed line) wave number no. 30; (b) estimated (solid line) and true (dashed line) internal wave; (c) residual/innovation (3.9% out); (d) zero-mean/whiteness testing ( $0.31 < 1.4$ , 2.8% out).

when applied to noisy measurement data. Thus these simulations show that the Gauss–Markov formulation enables us to capture various uncertainties of internal waves as well as its associated statistics in a completely consistent framework.<sup>20</sup>

To further assess the feasibility of this approach, we ran the estimator on other synthetic data sets at 0 dB, -10 dB, and -13 dB with similar results, that is, the model-based approach enabled a near-optimal Bayesian solution with all sensor innovation sequences statistically testing as zero mean/white. We are currently applying this technique to carefully controlled experimental measurements and the preliminary results have led us to continue to pursue this approach. Next we summarize our findings.

### III. SUMMARY

In this paper we have developed a dispersive wave processor. Starting with Whitham's<sup>21</sup> solution to the propagation of waves in a dispersive medium, we developed the approximate (due to nonlinear systems) maximum *a posteriori* (MAP) solution using a Bayesian formulation of the wave-field estimation problem. The generality of this result enables the specification of a particular wave system by its underlying dispersion relation and envelope function which can be “plugged into” the algorithm to obtain the MAP solution.

Once the dispersive wave solution is developed, we apply it to design a model-based internal wave processor to solve a multichannel internal wave estimation problem. Starting with these general results developed for dispersive wave propagation and an optimal Bayesian solution, the enhancer is designed exclusively for internal waves demonstrating its applicability to a real world problem.

Once formulated, the performance of the MBP was evaluated on data synthesized at a -23-dB signal-to-noise ratio and shown to be effective in enhancing the final wave-field estimate. The simulator used to generate synthetic data is a very sophisticated representation of the internal wave dynamics incorporating both vertical and horizontal propagation to synthesize a meaningful characterization useful in analysis and experimental design.<sup>29</sup> In this case the processor was “tuned” by adjusting the assumed noise covariances until a zero-mean, white residual or innovations sequence could be achieved approaching the optimality of the processor.

There are some interesting conclusions that can be drawn from this work. First and most obvious, we have taken a dispersive wave solution model and developed a recursive maximum *a posteriori* processor to find a solution to the multichannel wave estimation problem. Under the assumption of additive Gaussian noise, this solution can be approximated with an extended Kalman filter. Next we have shown that it is possible to develop and apply this processor to a set of data synthesized by using a sophisticated internal wave model usually called the “truth model” and adding Gaussian noise. These data have essentially been used to demonstrate: (1) the feasibility of designing the dispersive wave processor; (2) the steps required to develop the processor; and (3) the fact that the solution to the wave estimation problem is

“robust” enough to capture the dominant dynamics of the internal wave phenomenology enabling the processor to “track” noisy ( $-23$  dB) data.

## ACKNOWLEDGMENTS

We would like to acknowledge the motivation and support of Dr. R. Twogood, Program Leader LLNL Imaging & Detection Program and James Brase, Project Leader for the Radar Ocean Imaging Project, supported by the Office of the Secretary of Defense and led by Donna Kulla. We also gratefully acknowledge our colleagues for the many clarifications and technical discussions: Dr. D. Holliday, H. Jones, D. Mantrom, and Dr. H. Robey. This work was performed under the auspices of the Department of Energy by the Lawrence Livermore National Laboratory under Contract No. W-7405-Eng-48.

- <sup>1</sup>C. H. Cheng and M. N. Toksoz, “Elastic wave propagation in a fluid-filled borehole and synthetic acoustic logs,” *Geophysics* **46**, 1042–1053 (1981).
- <sup>2</sup>M. Schoenberg, T. L. Marzetta, J. Aron, and R. P. Porter, “Space–time dependence of acoustic waves in a borehole,” *J. Acoust. Soc. Am.* **70**, 1496–1507 (1981).
- <sup>3</sup>A. L. Kurkjian and S. K. Chang, “Acoustic multipole source in fluid-filled boreholes,” *Geophysics* **51**, 148–163 (1986).
- <sup>4</sup>I. A. Viktorov, *Raleigh and Lamb Waves: Physical Theory and Applications* (Plenum, New York, 1967).
- <sup>5</sup>J. L. Rose, J. J. Ditri, A. Pilarski, J. Zhang, F. Carr, and W. McNight, “A guided wave inspection technique for nuclear steam generator tubing,” *Nondestr. Test. (Chicago)* **92**, 191–195 (1992).
- <sup>6</sup>C. S. Clay and H. Medwin, *Acoustical Oceanography* (Wiley, New York, 1977).
- <sup>7</sup>F. B. Jensen, W. A. Kuperman, M. B. Porter, and H. Schmidt, *Computational Ocean Acoustics* (American Institute of Physics, New York, 1994).
- <sup>8</sup>E. A. Robinson, T. S. Durrani, and G. Peardon, *Geophysical Signal Processing* (Prentice-Hall, New Jersey, 1986).
- <sup>9</sup>C. V. Kimball and T. L. Marzetta, “Semblance processing of borehole acoustic array data,” *Geophysics* **49**, 274–281 (1984).
- <sup>10</sup>J. H. McClellan, “Two-dimensional spectrum analysis in sonic logging,” *IEEE Trans. Acoustics, Speech, Signal Proc. Magazine*, pp. 12–17, July 1986.
- <sup>11</sup>S. W. Lang, A. L. Kurkjian, J. H. McClellan, C. F. Morris, and T. W. Parks, “Estimating slowness dispersion from arrays of sonic logging waveforms,” *Geophysics* **52**, 530–544 (1987).
- <sup>12</sup>D. H. Johnson and D. E. Dudgeon, *Array Signal Processing* (Prentice-Hall, New Jersey, 1993).
- <sup>13</sup>Y. Bresler and A. Macovski, “Exact maximum likelihood estimation of superimposed exponential signals in noise,” *IEEE Trans. Acoust., Speech, Signal Process.* **34**, 1081–1089 (1986).
- <sup>14</sup>P. Stoica and A. Nehorai, “Statistical efficiency study of direction of estimation methods, part I: Analysis of MUSIC and a preliminary study of MLM,” in *Advances in Spectrum Analysis and Array Processing: Vol. 2* (Prentice-Hall, Englewood Cliffs, NJ, 1991), pp. 263–306.
- <sup>15</sup>C. V. Kimball, P. Lewicki, and N. I. Wijeyesekera, “Error analysis of maximum likelihood estimates of physical parameters from one or more dispersive waves,” *IEEE Trans. Signal Process.* **43**, 2928–2936 (1995).
- <sup>16</sup>A. Tolstoy, *Matched-Field Processing for Ocean Acoustics* (World Scientific, New Jersey, 1993).
- <sup>17</sup>J. V. Candy and E. J. Sullivan, “Ocean acoustic signal processing: A model-based approach,” *J. Acoust. Soc. Am.* **92**, 3185–3201 (1992).
- <sup>18</sup>J. V. Candy and E. J. Sullivan, “Model-based identification: An adaptive approach to ocean acoustic processing,” *IEEE J. Ocean Eng.* **21**, 273–289 (1996).
- <sup>19</sup>J. V. Candy and E. J. Sullivan, “Passive localization in ocean acoustics: A model-based approach,” *J. Acoust. Soc. Am.* **98**, 1455–1471 (1995).
- <sup>20</sup>J. V. Candy, *Signal Processing: The Model-Based Approach* (McGraw-Hill, New York, 1986).
- <sup>21</sup>G. B. Whitham, *Linear and Nonlinear Waves* (Wiley, New York, 1974).
- <sup>22</sup>J. V. Candy and D. H. Chambers, “Internal wave processing: A model-based approach,” *IEEE J. Ocean Eng.* **21**, 37–52 (1996).
- <sup>23</sup>A. Jazwinski, *Stochastic Processes and Filtering Theory* (Academic, New York, 1970).
- <sup>24</sup>J. R. Apel, *Principles of Ocean Physics* (Academic, New York, 1987).
- <sup>25</sup>C. Garrett and W. Munk, “Internal waves in the ocean,” *Annu. Rev. Fluid Mech.* **11**, 339–369 (1979).
- <sup>26</sup>W. Alpers, “Theory of radar imaging of internal waves,” *Nature (London)* **314**, 245–247 (1985).
- <sup>27</sup>D. R. Thompson and R. F. Gasparovic, “Intensity modulation in SAR images of internal waves,” *Nature (London)* **320**, 345–348 (1986).
- <sup>28</sup>B. C. Barber, “On the dispersion relation for trapped internal waves,” *J. Fluid Mech.* **252**, 31–49 (1993).
- <sup>29</sup>M. Milder, “Internal Waves Radiated by a Moving Source: Analytic Simulation, National Technical Information Service,” Doc. No. AD-782-262, 1974.
- <sup>30</sup>J. W. Miles, “Internal waves generated by a horizontally moving source,” *Geophys. Fluid Dyn.* **2**, 63–87 (1971).
- <sup>31</sup>D. D. Mantrom, “Loch Linnhe ’94: Test Operations Description and On-Site Analysis US Activities,” *Lawr. Liverm. Nat. Lab. Report., UCRL-ID-119197*, 1994.
- <sup>32</sup>H. F. Robey and D. L. Ravizza, “Loch Linnhe Experiment 1994: Background Stratification and Shear Measurements Part 1: Profile Summary and Dispersion Relations,” *Lawr. Liverm. Nat. Lab. Report., UCRL-ID-119352*, 1994.
- <sup>33</sup>J. V. Candy and P. M. Candy, “SSPACK\_PC: A model-based signal processing package on personal computers,” *DSP Applications* **2**, 33–42 (1993).
- <sup>34</sup>MATLAB (The Math Works, Inc., Boston, 1993).

# Detection of phase- or frequency-modulated signals in reverberation noise

Valérie Carmillet, Pierre-Olivier Amblard,<sup>a)</sup> and Geneviève Jourdain  
*Laboratoire des Images et des Signaux-LIS UPRESA CNRS 5083, ENSIEG-BP 46,  
38 402 Saint-Martin d'Hères Cedex, France*

(Received 22 June 1998; revised 21 January 1999; accepted 25 February 1999)

This paper is devoted to the detection of phase- or frequency-modulated signals corrupted by the reverberation noise they have created. When the reverberation is modeled as a time-varying stochastic filtering of the emitted signal, the optimal detector consists of prewhitening the observation before applying a matched filter. This theory is very difficult to implement because of the nonstationarity of the reverberation noise. In this paper, a suboptimal detector proposed by Kay and Salisbury in 1990 is generalized to phase or frequency modulated signals. The reverberation is assumed locally stationary; the signal is cut into blocks and the whitening is performed as follows: An autoregressive (AR) modeling of one block allows the whitening of the following block. After whitening, an appropriate matched filter is applied. This scheme is explained in detail. The local stationarity assumption of the reverberation noise is discussed in light of the Itakura spectral distance. Performance of the proposed approach is evaluated theoretically when all parameters are known, and experimentally using three different real-data sets. © 1999 Acoustical Society of America. [S0001-4966(99)02606-5]

PACS numbers: 43.60.Gk, 43.30.Vk [JCB]

## INTRODUCTION AND PROBLEM STATEMENT

In underwater acoustics, problems due to the presence of reverberation noise in active sonar detection are now well known. They are especially important when a low Doppler target is situated just under the ocean surface, in a convergence zone, or in shallow water.

Detection of a target in reverberation noise is difficult because the reverberation is dependent on the signal (hence a high correlation between the signal and the noise); it is non-stationary, and it is strongly colored.<sup>1,2</sup>

In the case of strongly colored noise, the classical matched filter fails to give good detection, and its performance dramatically decreases as the signal-to-noise ratio decreases. It is well-known that to overcome this difficulty the data must be prewhitened prior to any matched filtering. Of course, this whitening requires the knowledge of the structure of the noise covariance (or spectrum) function. If this information is not known *a priori*, the covariance function can be estimated. But the case of reverberation noise is more difficult since it is not stationary and its covariance function cannot be easily estimated using only one signal. Hence, suboptimal strategies based on some strong hypotheses (like local stationarity of the reverberation) must be considered.

Kay and Salisbury<sup>3</sup> gave a suboptimal solution to the problem of detecting a sinusoid in the reverberation noise created by the sine. Their approach relies on the assumption that the noise is locally stationary, and uses a prewhitener based on an AR (autoregressive) modeling of the noise. This approach has also been studied in radar contexts.<sup>4</sup>

Today, active sonars increasingly use phase- or

frequency-modulated signals, such as chirp signals (linear, hyperbolic frequency-modulated signals), binary phase shift keying (BPSK), etc. The approach proposed by Kay and Salisbury<sup>3</sup> is therefore no longer valid, since it only works for sinusoids. In this paper, we generalize the results for sinusoids<sup>3</sup> to any kind of emitted signals. That general theory is presented in Sec. I. The fundamental assumption of local stationarity is discussed in light of the spectral measure of Itakura. The suboptimal detector is then detailed in Sec. II, and its theoretical performance presented in the case of known parameters. In Sec. III, we present applications of this approach on three real-data sets of reverberation acquired in the Mediterranean Sea. In addition, experimental performances in the case of unknown parameters are studied. We finally discuss some points and present some ideas for future study.

## I. DETECTION OF PHASE- OR FREQUENCY-MODULATED SIGNALS IN NONSTATIONARY COLORED NOISE

The general problem we address in this paper is the following detection problem:

$$\begin{cases} H_0 : y(t) = n(t) + b(t), \\ H_1 : y(t) = s(t) + n(t) + b(t), \end{cases} \quad (1)$$

where  $y(t)$  is the observed or received signal,  $n(t)$  is the reverberation noise generated by the emitted signal, and  $b(t)$  represents any other kind of noise.  $s(t)$  is the signal to be detected. We assume here that it is linked to the emitted signal in a simple way:  $s(t)$  is equal to the emitted signal up to a time delay, a Doppler shift and an amplitude attenuation, or  $s(t) = A e^{i(t - \tau)} \exp(2i\pi f_d t)$  where the amplitude  $A$ , the delay  $\tau$ , and the Doppler shift  $f_d$  are unknown parameters. In other words, the extent of the target is assumed to be very

<sup>a)</sup> Author to whom correspondence should be addressed. Electronic mail: Bidou.Amblard@lis.inpg.fr

small compared to the extent of the surface that generates the reverberation noise. Furthermore, the emitted signal is a phase- or frequency-modulated signal. But, it should be noted that the ratio  $B/\nu_0$  of the band  $B$  of the emitted signal to the carrier frequency  $\nu_0$  used during the emission should be small, so that the received waveform is not distorted too much. Finally, in the detection problem (1),  $y(t)$  represents the output of the sonar after complex demodulation (complex envelope). Hence, signals considered throughout the paper are complex valued.

When dealing with reverberation noise, the most difficult problems to solve are the dependence of the noise on the emitted signal, and the nonstationary character of the noise.

When the reverberation noise is modeled as a time-varying stochastic filtering process,<sup>1,2</sup> a general theory of detection can be written. This theory shows that the optimal strategy is to first whiten the signal, and to then apply the matched filter to the desired signal modified by the whitening filter. However, the whitening filter is the solution of an integral equation with a time-varying kernel. This equation may be solved in rare cases, and in all cases, this solution requires the knowledge of the covariance function of the reverberation noise. This function is very difficult to estimate since the noise is nonstationary. Therefore, suboptimal strategies must be developed.

Two main possibilities appear:

- (1) The whitening step is ignored, and the matched filter is applied to the desired signal. This approach will be denoted as the classical approach in the sequel.
- (2) The general structure of the optimal detector is retained (i.e., whitening and matched filtering), but the structure is implemented in a suboptimal way. This second approach requires some strong hypotheses on the reverberation noise.

We now detail a strategy that fits the second approach.

### A. A suboptimal strategy

One possibility<sup>3</sup> relies on two assumptions:

- (1) The signal to be detected  $[s(t)]$  has a short duration compared to the reverberation noise. This assumption is realistic and most often verified in practice.
- (2) The reverberation noise is locally stationary. This means that the reverberation noise may be considered as stationary over intervals of some duration. This assumption will be discussed in Sec. IB.

These assumptions lead to a natural solution (though suboptimal) of the detection problem. The received signal is cut into blocks of equal length, and the noise on each block is assumed to be stationary. The detection problem introduced above is thus transformed into the following one. Suppose that we have performed the detection on  $K$  blocks and that no signal has been detected. The detection problem is then to determine if the signal is present or not in the  $(K+1)$ th block, or

$$\begin{cases} H_0 : y(t) = n(t) + b(t) \\ H_1 : y(t) = s(t) + n(t) + b(t) \end{cases} \quad \text{for } t \in (K+1)\text{th block.} \quad (2)$$

Note that we implicitly assume here that the support of  $s(t)$  is shorter than (or equal to) the length of a block, and that this support is entirely included in one block. This is of course unrealistic, but we will discuss this point in the last section, and show how to overcome the weakness of this assumption.

The detection strategy is now simpler, since it consists of deciding on the presence of a signal in a stationary colored noise. Furthermore, this allows the elimination of the estimation of parameter  $\tau$ , which represents the time delay of the signal to be detected. The estimated  $\tau$  will correspond to the number of the block in which the signal is detected.

A second difficult problem is the dependence of the noise on the emitted signal. This leads to great complications. However, it also gives some useful information on the noise. For example, since the reverberation process is assumed to be linear and time varying, the power spectrum of the noise has a support which is larger than the support of the power spectrum of the emitted signal. In what follows, we ignore the dependence of the noise on the emitted signal, but we keep this dependence in mind for *a priori* information (e.g., support of the power spectrum). To detect a signal in colored noise, it is well-known that the optimal strategy begins by a whitening of the noise. But, whitening requires the knowledge of the covariance function of the noise. Hence, we need the covariance function on block  $K+1$  to derive an optimal detector on that block. Estimating the covariance on the block may lead to great difficulties if the signal is present on this block. The idea proposed by Kay and Salisbury<sup>3</sup> is to use the estimation of the covariance on block  $K$  (remember that the signal has not been detected in the first  $K$  blocks) to whiten the noise on block  $K+1$ . This is very simple, but induces a complication in the local stationarity assumption: The noise must now be stationary on the scale of two blocks. This problem is addressed in Sec. IB.

If the noise has been whitened on block  $K+1$  by a filter  $w(\cdot)$ , the problem of detection is thus to decide if signal  $w(s(t))$  is present in white noise. This is a classic problem, and we know that the optimal detector in the Neymann–Pearson sense is the matched filter (if the noise is Gaussian).

We finally have the following suboptimal strategy:

- (1) Cut the signal into blocks.
- (2) If no signal has been detected up to block  $K$ , use the estimation of the covariance of the noise on block  $K$  to whiten with filter  $w(\cdot)$  the noise on block  $K+1$ .
- (3) Apply the optimal detector of signal  $w(s(t))$  in white Gaussian noise on block  $K+1$ .
- (4) If no signal is detected, increment  $K$  and go to 2; otherwise, stop.

The “stop” in the last item is not compulsory and the process of detection may continue. However, in that case, block  $K+2$  cannot be tested since an estimation of the covariance of the noise on block  $K+1$  is not available (a signal is present). The detection scheme is synthesized in Fig. 1.

We now discuss the local stationarity assumption.

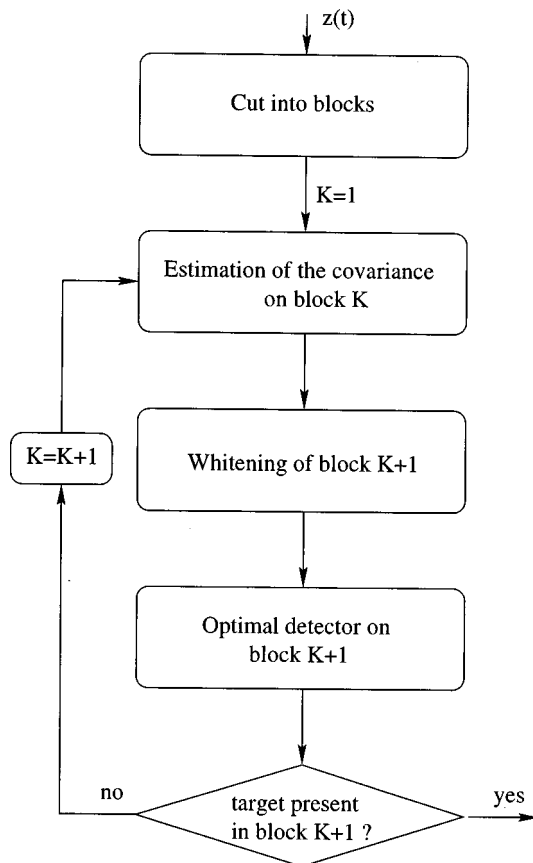


FIG. 1. Diagram of the detection strategy used in this paper.

### B. Local stationarity hypothesis

We present in Fig. 2 an example of reverberation noise generated by a linear chirp (sine wave with frequency varying linearly in time). The nonstationary property of the noise is clearly seen. Since we split the signal into blocks, Fig. 2 also represents the splitting by vertical lines. The length of the blocks corresponds to the length of the emitted signal. It is now necessary to study the stationarity property of the signal on each block and between two successive blocks. An examination of Fig. 2 shows that the stationarity of the signal on one block seems valid. For example, the variation of the instantaneous power on one block is small.

The second and more important point is the stationarity on the scale of two blocks, since we need an estimate of the

covariance of the noise on one block to whiten the noise on the following block. Hence, we concentrate here on second-order stationarity, i.e., we address the problem of comparing the covariances of two signals. Since knowing the covariance is equivalent to knowing the power spectrum, we will consider that the signal is stationary on two consecutive blocks if the spectrum of the noise does not change significantly from one block to the next. Comparing power spectral densities is a vast subject. There are many ways to deal with the question, but standard approaches consist of evaluating a distance between two spectral densities.

The distance we choose here is the “symmetrized” Itakura distance.<sup>5,6</sup> This distance was also used in underwater acoustics for classification purposes.<sup>7</sup> In this work, some distances were compared, and the Itakura distance gave the best results. It allowed the separation of signals whose spectral densities were very close.

Consider two discrete time signals  $x_1(t)$  and  $x_2(t)$  with respective power spectral densities  $s_1(\lambda)$  and  $s_2(\lambda)$ . The Itakura distance is defined as follows:

$$d_I(s_1, s_2) = \ln \left[ \int_{-\pi}^{+\pi} \frac{1}{2\pi} \frac{s_1(\lambda)/\sigma_1^2}{s_2(\lambda)/\sigma_2^2} d\lambda \right], \quad (3)$$

where  $\sigma_i^2$  is the asymptotic variance of the linear prediction error of signal  $x_i(t)$

$$\left( \sigma_i^2 = \exp \left( \int_{-\pi}^{\pi} \log(s_i(\lambda)) / (2\pi) d\lambda \right),^{6,8} \right).$$

However, this definition is not symmetrical, and is not *stricto sensu* a distance. We thus work with the symmetrized version of it, which is defined as

$$d'_I(s_1, s_2) = \frac{1}{2}(d_I(s_1, s_2) + d_I(s_2, s_1)). \quad (4)$$

If the signals are proportional, it is easy to see that  $d'_I(s_1, s_2) = 0$ . Hence, the distance is not only zero in the case of equal signals, but also in the case of proportional signals. This is of great interest, since the power of one block may increase on the following block. Therefore, the comparison we perform between two blocks is “power blind.”

To evaluate  $d'_I(s_1, s_2)$  between two blocks, we need to estimate the power spectra of the signals on these blocks. In our application, this is a difficult task since the length of the

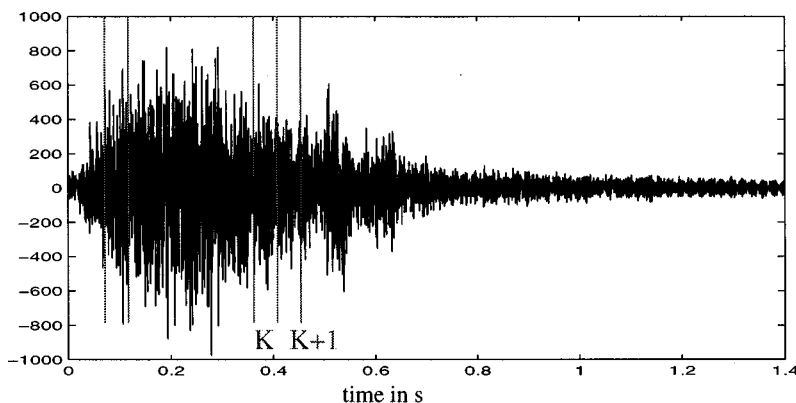


FIG. 2. Example of reverberation noise generated by the emission of a linear frequency modulation signal (LFM). The splitting of the signal into blocks is represented by the vertical lines (three blocks are represented: The first one around 0.1 s, and the other two around 0.4 s). Each block is then considered as stationary.



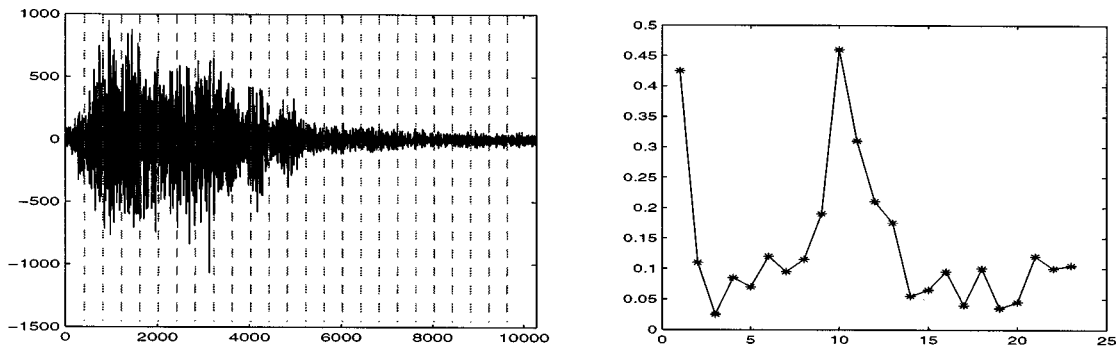


FIG. 3. Left: Reverberation noise resulting from the emission of a linear chirp, and its splitting into blocks. Right: Evolution of the Itakura distance between two successive blocks. The distance stays below 0.2, except for blocks 1 and 2 (high nonstationarity) and for blocks 10–11 and 11–12 (presence of a target in the sea).

block is in general quite short; we will not have a lot of samples to make a good estimate of the power spectral densities.

We therefore adopt another approach (that will also be used for the whitening, see Sec. II): We model the reverberation noise on one block as an autoregressive model. Consider an autoregressive model of order  $p$ , whose parameters are denoted by  $a_k$ . Let  $\varepsilon(t)$  be the white noise that drives the model. Then, the signal generated by the model is written as

$$x(t) = - \sum_{j=1}^p a_j x(t-j) + \varepsilon(t). \quad (5)$$

Consider

$$A(z) = \sum_{k=0}^p a_k z^{-k}, \quad \text{with } a_0 = 1. \quad (6)$$

Then, the spectral density of signal  $x(t)$  can be written as

$$s(\lambda) = \frac{\sigma^2}{|A(e^{i\lambda})|^2}, \quad (7)$$

where  $\sigma^2$  is the variance of  $\varepsilon(t)$ . We now assume that the reverberation noise on block  $K$  is an autoregressive model with spectral density  $s_K(\lambda)$ . Now, using Eq. (7), the Itakura distance between blocks  $K$  and  $L$  can be put in the following form:

$$d_I(s_K, s_L) = \ln \left[ \int_{-\pi}^{+\pi} \frac{1}{2\pi} \frac{|A_L(e^{i\lambda})|^2}{|A_K(e^{i\lambda})|^2} d\lambda \right]. \quad (8)$$

We now apply this distance on a reverberation signal obtained after the emission of a linear chirp. We split this signal into 25 blocks, and evaluate the Itakura distance between two successive blocks. The result is shown in Fig. 3. The signal and its splitting into 25 blocks are shown in the left panel. The evolution of the symmetrized Itakura distance between two successive blocks appears in the right panel. In the experiment, a target was present 400 m from the sonar, and appears around block #11 of the signal. This explains the peak observed in the Itakura distance around that block. Except for block #11, the greatest distance is observed for blocks #1 and #2, where the nonstationarity is the most violent, since it corresponds to the beginning of the reverberation process. For the other blocks, the distance is lower than 0.2 and roughly constant. This means that the AR models

from one block to the next do not change very much. For a better understanding of the meaning of a distance lower than 0.2, we plot in Fig. 4 the distance between block  $K$  and block #1 for the reverberation signal of Fig. 3. We observe that the distance is almost always greater than 0.2, and this shows that the stationarity cannot be considered on the scale of several blocks.

In light of these results, we conclude that the local stationarity assumption on the scale of two blocks is a reasonable assumption. Therefore, we can use the estimation of the covariance of the reverberation noise on block  $K$  to whiten the data of block  $K+1$ . This is done using the AR model obtained in block  $K$ .

### C. Whitening using AR modeling

We now want to implement the whitening filter on block  $K+1$  using the estimation of the covariance of block  $K$ . In order to estimate this covariance, we proceed as in the previous section, and adopt an AR modeling of the signal on block  $K$ . The estimated AR model may be written as

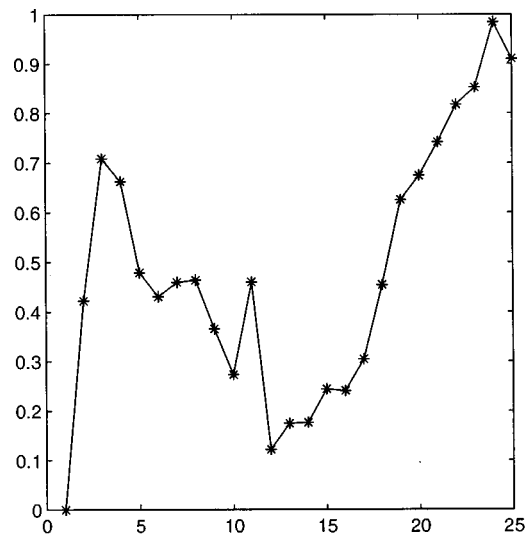


FIG. 4. Itakura distance between block  $K$  and block 1 for the signal of Fig. 3. The distance stays above 0.2, showing the nonstationarity on the scale of several blocks.

$$y_K(t) = -\sum_{j=1}^{\hat{p}_K} \hat{a}_K(j)y_K(t-j) + \varepsilon_K(t),$$

where  $\varepsilon_K(t)$  is a white noise. The caret marks ( $\hat{\cdot}$ ) mean that the parameters have been estimated. The estimation of parameter  $a_K$  is performed using the modified covariance method.<sup>9</sup> The order  $p_K$  of the model is chosen according to the minimum description length (MDL) criterion. Let us mention that in all our experiments, the shortest length of the block was 250 samples. For this length, the estimation of the model can be considered as correct since the orders chosen were small (maximum 15). Therefore, the whitening filter  $w$  associated with this AR model has the following transfer function:

$$W_K(z) = 1 + \sum_{j=1}^{\hat{p}_K} \hat{a}_K(j)z^{-j}.$$

To use this filter to whiten the signal on block  $K+1$ , we invoke the local stationarity assumption. In the AR modeling context, this assumption means that the signal on block  $K+1$  follows an AR model whose parameters coincide with the parameters estimated on block  $K$ . Therefore, signal  $y_{K+1}(t)$  is passed through filter  $w$  to give

$$z_{K+1}(t) = w_K(y_{K+1}(t)).$$

The detection problem is now simpler since it consists of detecting a signal in white noise. This is the goal of the following section.

## II. OPTIMAL DETECTION OF A SIGNAL IN AUTOREGRESSIVE NOISE

The initial detection problem on block  $K+1$  is now replaced by

$$\begin{cases} H_0: y_{K+1}(t) = n_{K+1}(t) + b_{K+1}(t) \\ \text{for } t=0, \dots, N-1, \\ H_1: y_{K+1}(t) = s_{K+1}(t) + n_{K+1}(t) + b_{K+1}(t) \end{cases} \quad (9)$$

where  $b_{K+1}(t)$  is white Gaussian noise and  $n_{K+1}(t)$  is the Gaussian reverberation noise. The ‘‘reverberation-to-noise’’ ratio is assumed high, and the situation is called ‘‘reverberation dominant.’’ Signal  $s_{K+1}(t)$  is a modified version of the emitted signal and may be written as  $s_{K+1}(t) = Ae(t)\exp(2i\pi f_d t)$ . Note that parameter  $\tau$  has been dropped since we assume that it corresponds to the number of the block in which we are testing.

For convenience, we will now drop the subscript  $K+1$  in our notation. However, the reader must keep in mind that we are working on block  $K+1$ .

### A. Optimal detector

The optimal detector whitens the data  $y(t)$  and performs a matched filtering. We therefore implicitly assume that the ‘‘total’’ noise  $n(t) + b(t)$  is well modeled as an autoregressive signal, so that the whitening approach described above can be applied. The detection problem after whitening may then be written as

$$\begin{cases} H_0: z(t) = \varepsilon(t) \\ H_1: z(t) = s^m(t) + \varepsilon(t) \end{cases} \text{ for } t=0, \dots, N-1, \quad (10)$$

where  $\varepsilon(t)$  is complex white Gaussian noise of variance  $\sigma^2$ . Signal  $s^m(t)$  corresponds to the filtering of signal  $s(t)$  (by the whitening filter  $w(\cdot)$ ) and thus reads

$$\begin{aligned} s^m(t) &= \sum_{j=0}^{\hat{p}_K} \hat{a}_K(j)s(t-j), \\ &= A \sum_{j=0}^{\hat{p}_K} \hat{a}_K(j)e^{(t-j)} \exp(2i\pi f_d(t-j)), \\ &= A\mu_\Theta(t), \end{aligned}$$

where by convention  $\hat{a}_K(0) = 1$ . The subscript  $\Theta$  stands for all the unknown parameters that rule signal  $\mu(t)$ : these are the Doppler shift  $f_d$ , but can also be parameters of the emitted signal (frequency growth rate—or slope—for a linear chirp, for example). If we adopt a vector notation, the detection problem is written as

$$\begin{cases} H_0: \mathbf{z} = \boldsymbol{\varepsilon} \\ H_1: \mathbf{z} = A\boldsymbol{\mu}_\Theta + \boldsymbol{\varepsilon}, \end{cases} \quad (11)$$

where vectors  $\mathbf{z}$ ,  $\boldsymbol{\mu}_\Theta$ ,  $\boldsymbol{\varepsilon}$  contain the  $N$  samples of the corresponding signals. Since parameters  $A$  and  $\Theta$  are unknown, we are faced with a composite hypothesis testing problem. We further assume that the variance  $\sigma^2$  of the noise is unknown. The detector is therefore given by the generalized likelihood ratio test (GLRT), which consists of using the classical likelihood ratio test in which unknowns are replaced by their estimators in the maximum likelihood sense.<sup>1</sup> In our case, the test statistic reads

$$L(\mathbf{z}) = \frac{\max_{A, \Theta, \sigma^2} p_{\mathbf{z}/H_1}(\mathbf{z})}{\max_{\sigma^2} p_{\mathbf{z}/H_0}(\mathbf{z})}.$$

The noise is Gaussian and complex circular (the real and imaginary parts of  $\mathbf{z}$  are statistically independent) so that

$$\begin{aligned} p_{\mathbf{z}/H_0} &= \frac{1}{\pi^N \sigma^{2N}} \exp\left(-\frac{\mathbf{z}^\dagger \mathbf{z}}{\sigma^2}\right), \\ p_{\mathbf{z}/H_1} &= \frac{1}{\pi^N \sigma^{2N}} \exp\left(-\frac{(\mathbf{z} - A\boldsymbol{\mu}_\Theta)^\dagger (\mathbf{z} - A\boldsymbol{\mu}_\Theta)}{\sigma^2}\right), \end{aligned}$$

where  $\dagger$  stands for complex conjugation and vector transposition. The explicit derivation of the GLRT test is provided in the Appendix. The test statistic is  $\mathcal{L} = \max_\Theta \mathcal{L}(\Theta)$ , where

$$\mathcal{L}(\Theta) = 2N \log \left[ \left( 1 - \frac{|\boldsymbol{\mu}_\Theta^\dagger \mathbf{z}|^2}{(\boldsymbol{\mu}_\Theta^\dagger \boldsymbol{\mu}_\Theta)(\mathbf{z}^\dagger \mathbf{z})} \right)^{-1} \right]. \quad (12)$$

This test can be shown to be equivalent to  $\mathcal{L}' = \max_\Theta \mathcal{L}'(\Theta)$ , where

$$\mathcal{L}'(\Theta) = 2N \frac{|\boldsymbol{\mu}_\Theta^\dagger \mathbf{z}|^2}{(\boldsymbol{\mu}_\Theta^\dagger \boldsymbol{\mu}_\Theta)(\mathbf{z}^\dagger \mathbf{z})}. \quad (13)$$

We will use this form in the sequel. We recognize the matched filter to signal  $\boldsymbol{\mu}_\Theta$  followed by a squarer and a power normalization.

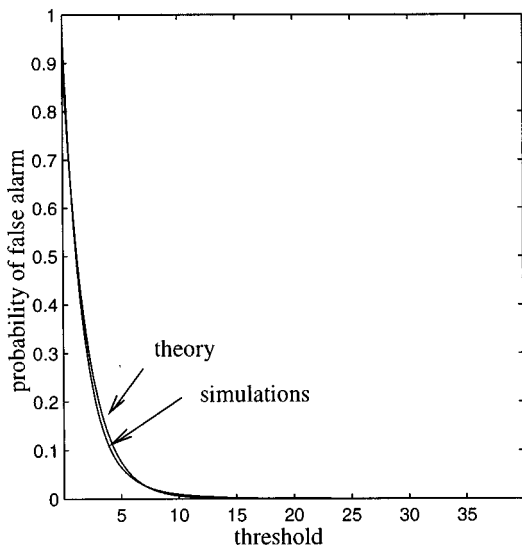


FIG. 5. False-alarm probability when parameters  $\Theta$  are known. 150 snapshots of an AR model were used. The first  $N$  samples were used to identify the whitening filter of the second part of the signal. The theoretical curve is added in this figure.

The decision is therefore taken after we compare the test statistic to a threshold  $\eta$ . If the test is greater than the threshold, hypothesis  $H_1$  is accepted, i.e., signal  $\mu_\Theta$  is declared present in block  $K + 1$ . If the test is lower than  $\eta$ , hypothesis  $H_0$  is accepted, i.e., signal  $\mu_\Theta$  is declared absent from block  $K + 1$ .

### B. Theoretical performances

The theoretical performance of the test  $\max_{\Theta} \mathcal{L}(\Theta) \geq_{H_0}^{H_1} \eta$  is difficult to study. It will be studied experimentally in Sec. III.

Here, we will present the theoretical performance of the detector for the case in which parameters  $\Theta$  are known, i.e., for the test  $\mathcal{L}(\Theta) \geq_{H_0}^{H_1} \eta$ .

When  $\Theta$  is known, it can be shown that  $\mathcal{L}(\Theta)$  is asymptotically a chi-square random variable with two degrees of freedom under hypothesis  $H_0$ .<sup>10</sup> The same result holds for  $\mathcal{L}'(\Theta)$ . This result allows an easy computation of the false alarm probability

$$P_{fa} = \text{Proba}(\mathcal{L}'(\Theta) > \eta / H_0) = \exp(-\eta/2).$$

Therefore, the false alarm probability depends only on the threshold. Hence, the detector we propose has a constant false-alarm rate over different blocks.

For the detection probability, it can be shown that the test  $\mathcal{L}(\Theta)$  follows asymptotically a noncentral chi-square law, with two degrees of freedom and a noncentrality parameter equal to  $\lambda = 2|A|^2 \mu_\Theta^\dagger \mu_\Theta / \sigma^2$ .<sup>10</sup> Hence, this parameter is directly linked to the signal-to-noise ratio. We have performed Monte Carlo simulations to verify these results. The signal chosen was a linear chirp. An AR model identified from a real reverberation noise (resulting from the emission of the chirp) was used to create our snapshots. Each snapshot

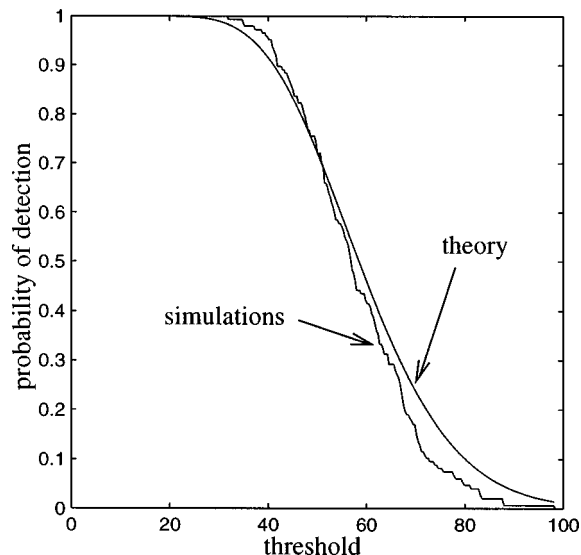


FIG. 6. Detection probability when parameters  $\Theta$  are known. 150 snapshots of an AR model were used. The first  $N$  samples were used to identify the whitening filter of the second part of the signal. The signal was placed in the second part of the snapshot with a signal-to-noise ratio of  $-12$  dB. The theoretical curve is added in this figure.

was made of  $2N$  samples to simulate two blocks. The whitening filter was estimated on the first  $N$  samples, and the test performed on the  $N$  remaining samples. To evaluate the detection probability, an echo was placed in the second block, whereas no signal was placed to evaluate the false-alarm probability. We used 150 snapshots to obtain the plot of Figs. 5 and 6. For the detection probability, the signal-to-noise ratio, evaluated classically as the ratio of the power of the signal to the power of the noise (in the support where the signal is located), is chosen at  $-12$  dB (note that the fre-

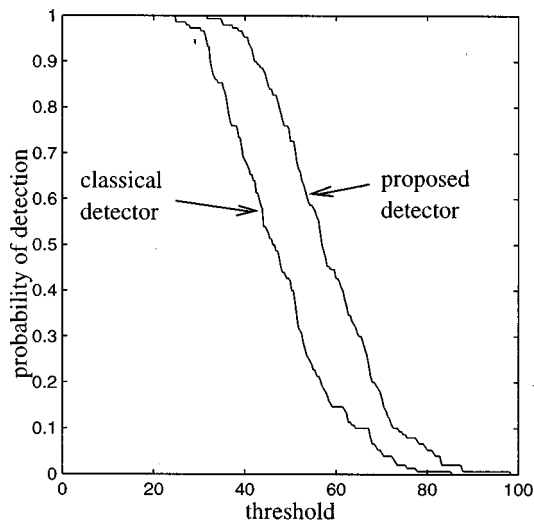


FIG. 7. Detection probability when parameters  $\Theta$  are known. Comparison between the classical detector and the proposed detector. 150 snapshots of an AR model were used. For the proposed detector, the first  $N$  samples were used to identify the whitening filter of the second part of the signal. The signal was placed in the second part of the snapshot with a signal-to-noise ratio of  $-12$  dB. In the case of the classical detector, no whitening is performed.

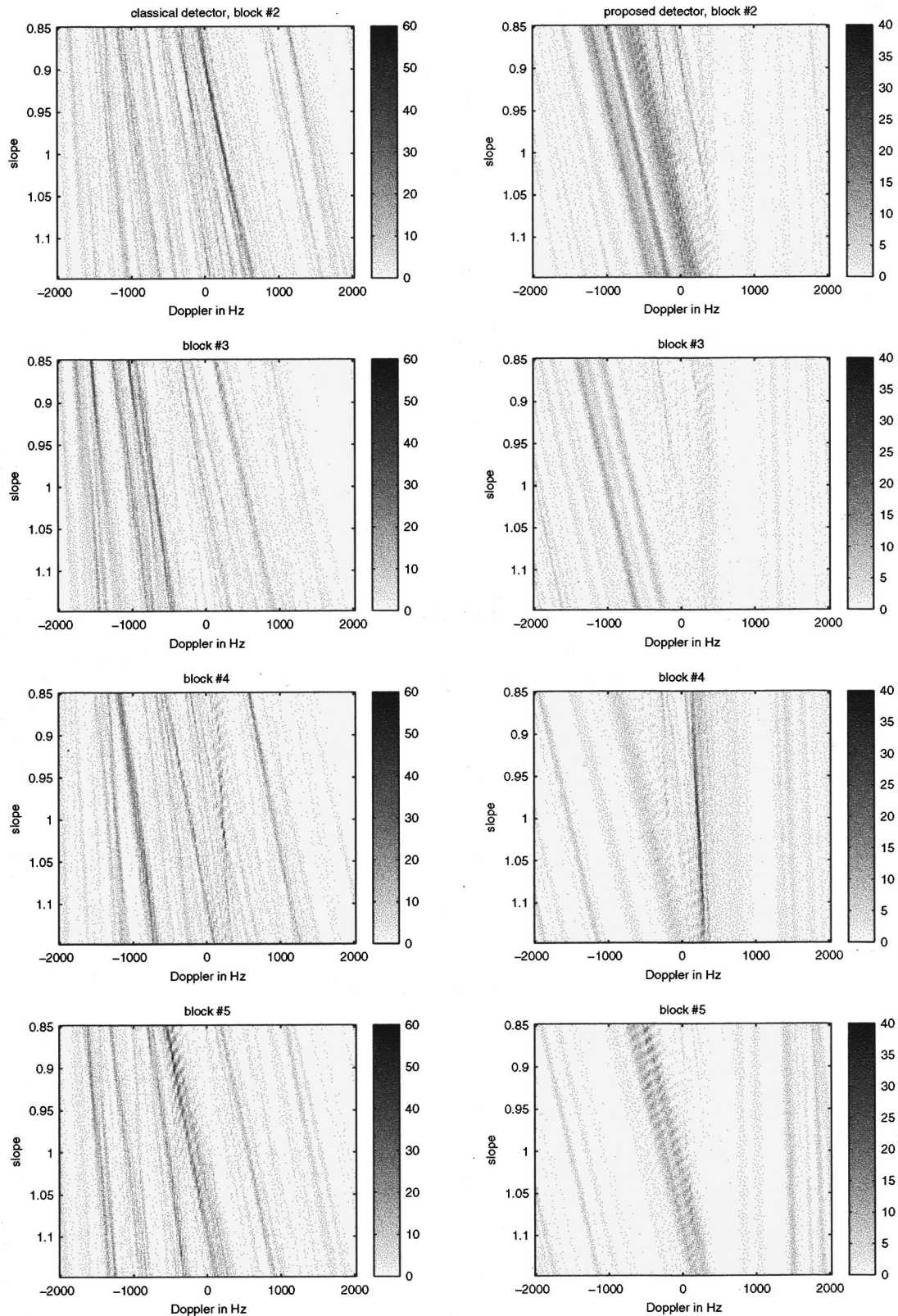


FIG. 8. Proposed (right) and classical (left) detectors of LFM signals in reverberation noise. Blocks 2, 3, 4, and 5. The target is in block 4. The classical detector has many false alarms.

quency supports of the signal and of the noise are roughly the same). Theoretical and simulation results show good agreement.

More interesting is a comparison between the proposed

detector and the classical detector. We must remember that this detector ignores the whitening operation. It is therefore restricted to a matched filter to signal  $s(t)$ , followed by a squarer and power normalization. It reads

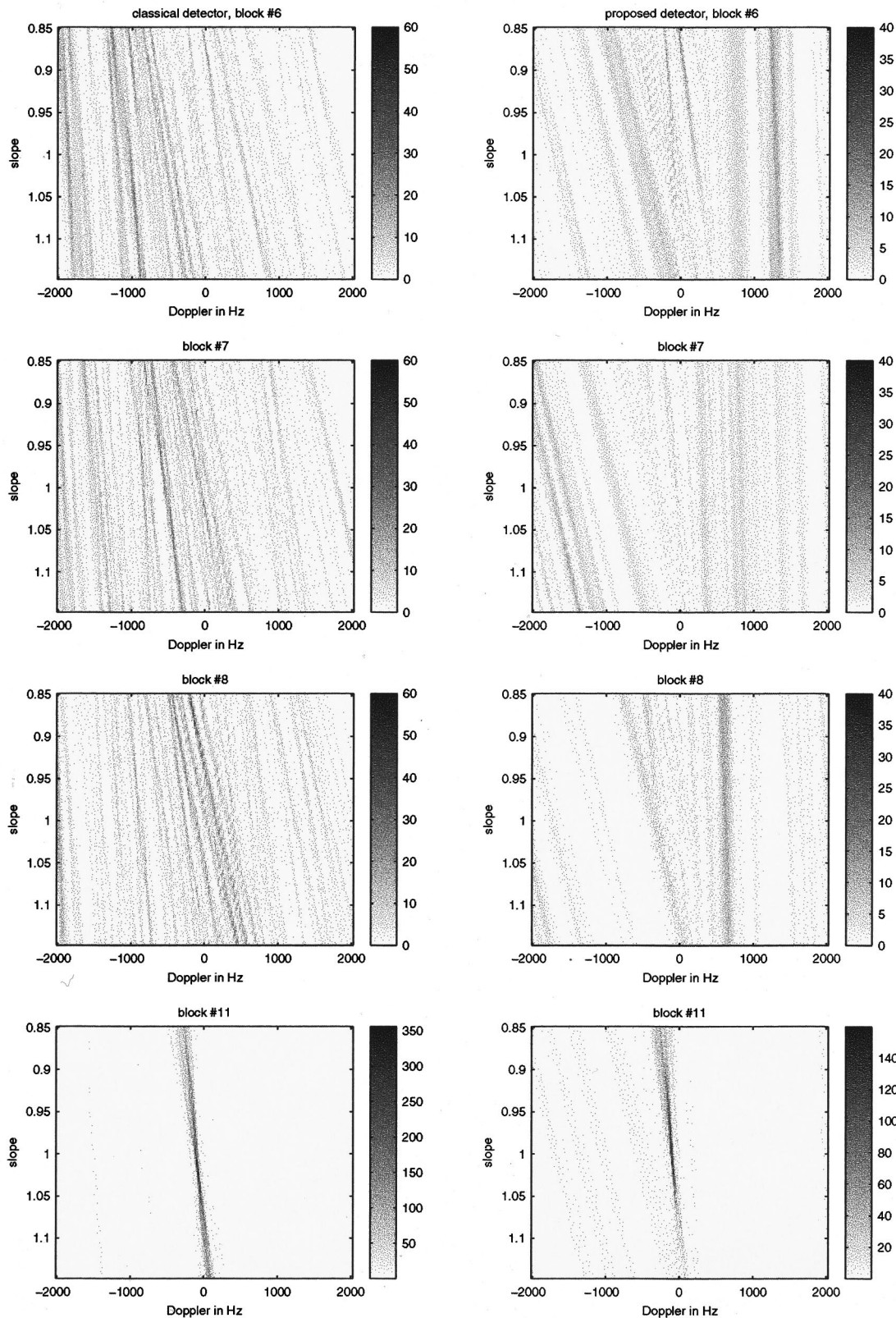


FIG. 9. Proposed (right) and classical (left) detectors of LFM signals in reverberation noise. Blocks 6, 7, 8, and 11. In this last block, the detectors find an immobile target.

$$\mathcal{L}_c(\Theta) = 2N \frac{|\mathbf{s}^\dagger \mathbf{y}|^2}{\mathbf{y}^\dagger \mathbf{y}}, \quad (14)$$

where  $\mathbf{s}$  is the vector that contains the samples of signal  $s(t)$ . The false-alarm probability of the classical detector is

equivalent to the false-alarm probability of the proposed detector. However, the detection probability of the proposed detector is much better than that of the classical detector, as illustrated in Fig. 7. This figure was obtained in the same way as before for a signal-to-noise ratio of  $-12$  dB. The

improvement is clear in this figure. Further comparisons will be made in Sec. III for the more realistic case of unknown parameters  $\Theta$ .

### III. REAL-CASE STUDY

In this section, we present real-case experiments. In the first paragraph, we will present the detector for three different types of signals: linear frequency modulation (LFM), hyperbolic frequency modulation (HFM), and binary-phase shift keying signals (BPSK). The case of the sinusoid has already been treated.<sup>3</sup> The behavior of the detector on real experiments will then be shown. Finally, the experimental performance of these detectors will be evaluated, and this will allow us to compare the three signals in a detection perspective.

#### A. Structure of the detector for LFM, HFM, and BPSK

In the following paragraph, we will show the results of the proposed approach for three different signals, and hence, three different types of reverberation noise. The first signal is a linear frequency modulated signal that reads

$$e(t) = B \exp(2\pi i(f_0 + \alpha_0 t)t) \quad \text{for } t \in [0, T], \quad (15)$$

where  $f_0$  is the carrier frequency used in the experiment. After complex demodulation, the received echo reads

$$s(t) = A \exp(2\pi i(f_d + \alpha t)t), \quad (16)$$

where  $f_d = f_0(\pm 2v/c)$  and  $\alpha = \alpha_0(1 \pm 2v/c)$  if  $v$  is the velocity of the target and  $c$  the speed of sound in the ocean. Note here that  $f_0$  and  $\alpha_0$  are known, and therefore,  $f_d$  and  $\alpha$  are linked. Therefore, signal  $s$  depends on only one parameter ( $v/c$ ). However, in order to detect any kind of linear chirps, the detector is evaluated as a two-dimensional function of parameters  $f_d$  and  $\alpha$  (see Barbarossa's<sup>11</sup> general discussion on the detection of LFM signals).

For the hyperbolic frequency modulation, the emitted signal reads

$$e_\gamma(t) = C \exp(2\pi i \alpha \log(1 + \gamma_0 t)) \quad \text{for } t \in [0, T]. \quad (17)$$

If this signal has duration  $T$ , maximum and minimum frequency  $f_M$  and  $f_m$ , and bandwidth  $B = f_M - f_m$ , then the parameters are  $\alpha = T f_M f_m / B$  and  $\gamma_0 = B / [T(f_0 + B/2)]$ , where  $f_0$  is the carrier frequency used in the experiment. After complex demodulation, the echo reads

$$s(t) = A \exp(2\pi i [f_0 t + \alpha \log(1 + \gamma t)]), \quad (18)$$

where  $\gamma = (1 \pm 2v/c)\gamma_0$ . Note again that  $f_0$ ,  $\alpha$ , and  $\gamma_0$  are known parameters.

The case of the BPSK signal does not lead to an explicit formulation. A BPSK is a pseudorandom sequence of  $\pm 1$  which modulates the phase of a sinusoid. It is characterized by its number of symbols and its bandwidth.<sup>12</sup> Since we do not have a closed-form expression for a BPSK, the receiver needs a copy of the emitted signal in order to construct the detector. Finally, we assume that the received signal is a modified version of the emitted signal by an attenuation factor  $A$  and a Doppler shift  $f_d$ .

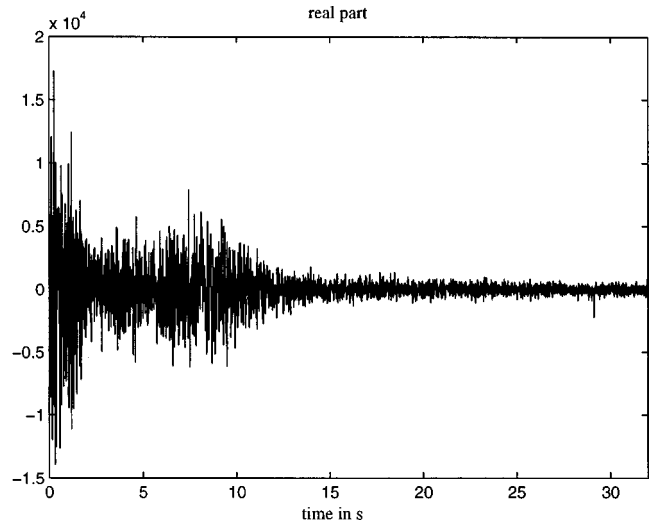


FIG. 10. Example of reverberation noise generated by the emission of an HFM. The reverberation is caused by the bottom of the ocean.

In all cases, the signal to be detected is in one block of the received signal. According to the detection procedure described above, this block will be whitened by a finite impulse response filter. This filter is obtained as the inverse of the AR model identified on the preceding block. The detector is then a filter matched to the target signal *modified by the whitening filter*. Note that this signal has been denoted by  $s^m(t)$  and reads

$$s^m(t) = \sum_{j=0}^{\hat{p}_K} \hat{a}_K(j) s(t-j) = A \mu_\Theta(t).$$

We now give the expression of  $s^m(t)$  for the three different cases studied here.

- (1) Linear frequency modulation: It can be shown that the modified signal remains a linear frequency modulation signal,<sup>13,14</sup> but with a time-varying amplitude. It reads

$$\begin{aligned} s^m(t) &= A \mu_{f_d, \alpha}(t), \\ &= A \exp(2\pi i (f_d + \alpha t)t) \sum_{j=0}^{\hat{p}_K} \hat{a}_K(j) \\ &\quad \times \exp(-2\pi i (f_d j + \alpha j(2t-j))). \end{aligned}$$

- (2) For the hyperbolic frequency modulation,<sup>13</sup> we obtain

$$\begin{aligned} s^m(t) &= A \mu_\gamma(t), \\ &= A \sum_{j=0}^{\hat{p}_K} \hat{a}_K(j) \exp(-2i\pi (f_0(t-j) + \alpha \log(1 + \gamma(t-j)))). \end{aligned}$$

- (3) Once again, for the BPSK, we do not have an explicit expression. Therefore, the receiver must have a copy  $e(t)$  of the emitted BPSK after demodulation. The ‘‘Dopplerized’’ signal is  $s(t) = A e(t) \exp(2i\pi f_d t)$ , so that

$$\begin{aligned} s^m(t) &= A \sum_{j=0}^{\hat{p}_K} \hat{a}_K(j) s(t-j), \\ &= A \mu_{f_d}(t). \end{aligned}$$

For each signal, we have stated the parameters to be estimated:  $\Theta = (f_d, \alpha)$  for an LFM,  $\Theta = \gamma$  for an HFM, and  $\Theta = f_d$  for the BPSK. The detector on block  $K+1$  is therefore

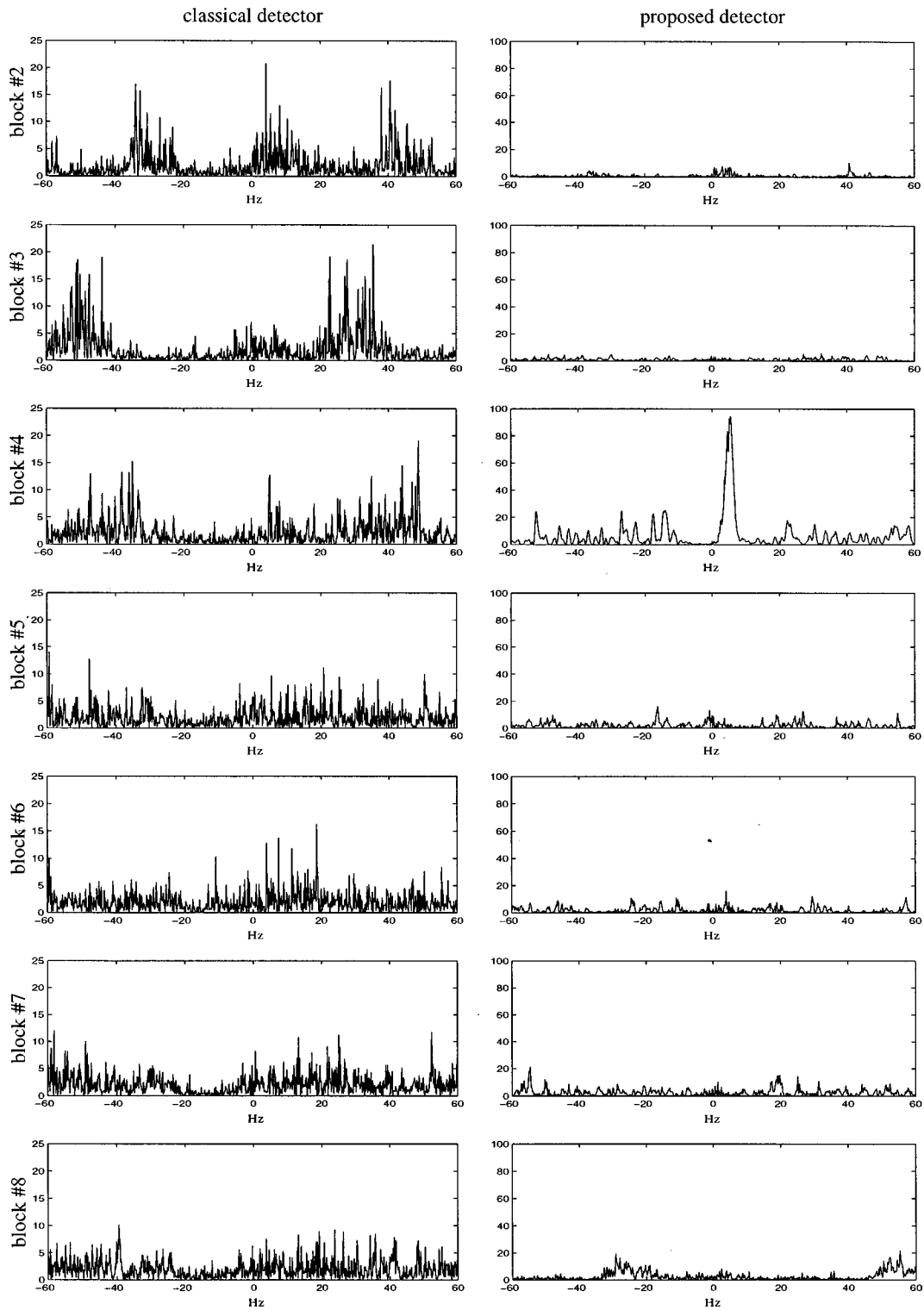


FIG. 11. Proposed (right) and classical (left) detectors of HFM signals in reverberation noise.

$$\mathcal{L}' = \max_{\Theta} \mathcal{L}'(\Theta),$$

$$\mathcal{L}' = \max_{\Theta} 2N \frac{|\mu_{\Theta}^{\dagger} \mathbf{z}|^2}{(\mu_{\Theta}^{\dagger} \mu_{\Theta})(\mathbf{z}^{\dagger} \mathbf{z})} \underset{H_0}{\geq} \eta,$$

where  $\mathbf{z}$  is the vector of the observations,  $\mu_{\Theta}$  contains the samples of signal  $\mu_{\Theta}(t)$ , and where  $\eta$  is the threshold. Note that for the LFM signals, the detector is closely linked to the so-called Wigner–Hough transform.<sup>11</sup>

Note that the detectors are quite heavy in terms of calculations. However, in the LFM case, one fast Fourier transform (FFT) can be used. But for HFM and BPSKs, we do not have rapid implementation of the detector.

## B. Real experiments

We now present the result of the suboptimal strategy described above on real experiments. In each experiment, we have created a synthetic low-speed moving target of velocity

$v$ . The echo created is then placed in a real reverberation noise in block  $\#K_t$  such that the distance between the target and the sonar is approximately  $cNK_tT_e/2$  m, where  $N$  is the number of samples per block,  $T_e$  is the sampling period, and  $c$  the sound velocity in the sea.

For each experiment, we present the output of the detector prior to the maximization over  $\Theta$ , i.e., the plot of  $\mathcal{L}'(\Theta)$  vs  $\Theta$ . The output of the detector is presented for several successive blocks. Therefore, the vertical axis represents the range between the sonar and the target, and hence, the time delay between the emission of the signal and the reception of the echo.

### 1. LFM signals

The LFM signal used has a decreasing frequency; its central frequency is 22.5 kHz, its band-width  $B$  is 3 kHz, and its duration  $T$  is 51 ms ( $BT$  product of 153). The sonar listens for 2 s. The sonar is used in emission and reception (monostatic situation), and is aimed at the sea surface with an aperture of about 20 deg. After sampling at a rate of 72.4 kHz, and after beamforming, the signal is complex demodulated, undersampled by a factor of 9 and cut into blocks of 411 samples.

The echo of a synthetic target is placed in block  $\#4$  with a signal-to-noise ratio of  $-12$  dB; it has a speed of  $v = 6.25$  m/s ( $f_d = 200$  Hz and  $\alpha = 1.0167$ ) and is located approximately at  $d = 150$  m from the sonar. In this experiment, a real immobile target is located at a distance of 400 m, and its echo is therefore in block  $\#11$ .

The output of the detector (before maximization) is shown in Fig. 8 for blocks  $\#2$  to  $\#5$ , and in Fig. 9 for blocks  $\#6$ ,  $\#7$ ,  $\#8$ , and  $\#11$ . There are two parameters to estimate: Doppler  $f_d$  is plotted in the  $x$ -axis and slope  $\alpha$  is plotted in the  $y$ -axis. Note that the detector has its responses concentrated on lines; this is due to the existing link between the slope and the Doppler frequency, as mentioned in the preceding section.

The synthetic target is easily detected in block  $\#4$ , with both the proposed detector and the classical detector. The parameters at the maximum of the detector are  $\alpha = 1.017$  and  $f_d = 202$  Hz and are thus good estimates of the true parameters. To see the superiority of the proposed detector, we look at other blocks, for which the classical detector presents many false alarms. This is evident for example in block  $\#7$ , where a false-alarm peak appears for the classical detector. Note that the Doppler corresponding to that false alarm is very small and corresponds to a realistic Doppler. We finally note that both detectors find the immobile target in block  $\#11$ . In this block, the signal-to-noise ratio is unknown, but we can argue that the level of the reverberation noise is very low since this target is far from the sonar. It is therefore easy for the detectors to find the target.

### 2. HFM signals

The HFM signal used has an increasing frequency, a central frequency of 1 kHz, and a bandwidth of 100 Hz. Its duration is 4 s ( $BT$  product of 400) and the array (composed of 27 sensors) listens for 34 s. After beamforming and com-

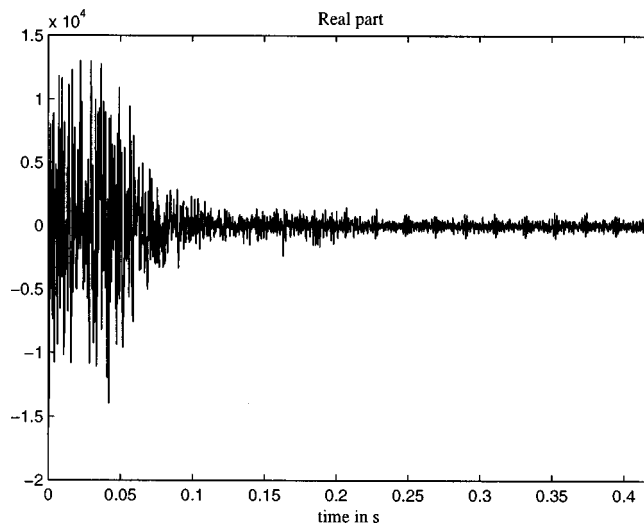


FIG. 12. Example of reverberation noise generated by the emission of a BPSK signal. The reverberation is caused by the sea surface.

plex demodulation, the signal is cut into blocks of 480 samples. In this experiment, the emission is done by a sonar that is aimed at the bottom of the sea. We are thus in the presence of bottom reverberation, and the bottom is of course immobile. A reverberation signal produced by the emission of an HFM is shown in Fig. 10.

The synthetic target is placed in block  $\#4$  and has a speed of  $v = 4$  m/s. The signal-to-noise ratio is chosen equal to  $-15$  dB, which is a very difficult situation. In fact, as we will see in the next section, this value corresponds to the limit in performance of the method for HFM. But, we have chosen this value since the result is spectacular, as illustrated by the output of the detector shown in Fig. 11.

Even with this very low signal-to-noise ratio, the proposed detector easily finds the target. The maximum of the detector gives the estimated Doppler frequency, and hence the estimated speed, which here is close to 4 m/s. Note that no false alarms occur in this experiment, but on other snapshots, false alarms may appear (this will be illustrated by the experimental receiver operating characteristic (ROC) curves).

### 3. BPSK signals

The duration of the emitted signal is 20.16 ms and its bandwidth 3.125 kHz ( $BT$  product of 60). The signal is emitted using a carrier frequency of 25 kHz. The sonar is aimed at the sea surface and has an aperture of about 30 deg. It is used for the emission and the reception. Before sampling at 25 kHz, the spectrum of the received signal is translated around frequency 0. After beamforming, complex demodulation, and undersampling by a factor of 2, the signal is cut into blocks of 252 samples.

The synthetic target is placed in block  $\#4$  with a signal-to-noise ratio of  $-20$  dB, and has a speed of  $v = 3$  m/s, corresponding to a Doppler frequency of 100 Hz. A reverberation signal produced by the emission of the BPSK is shown in Fig. 12. The output of the proposed detector and the classical detector are shown in Fig. 13.



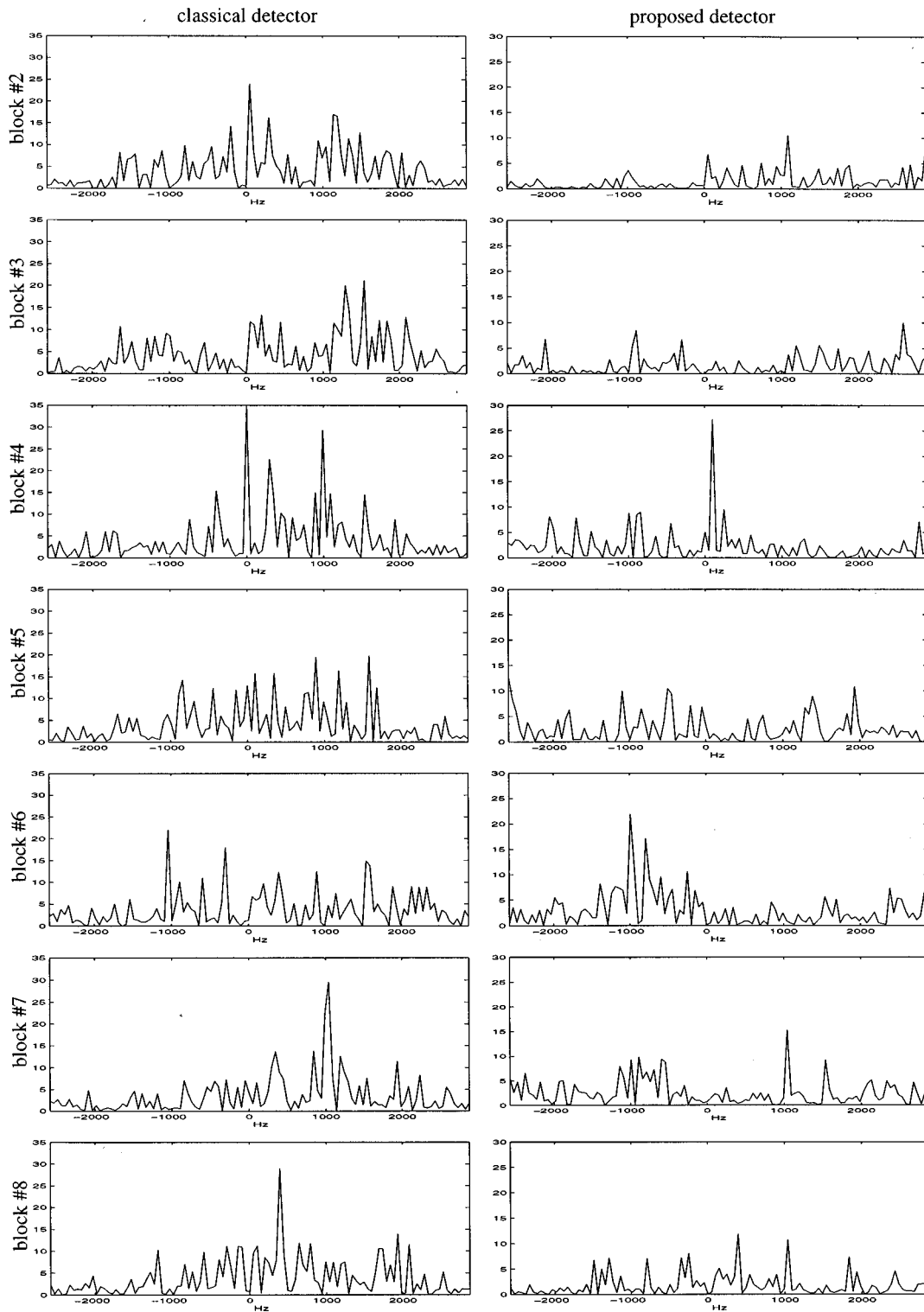


FIG. 13. Proposed (right) and classical (left) detectors of BPSK signals in reverberation noise.

Once again, the proposed detector is far better than the classical detector, even in the very low signal-to-noise ratio we study here. The proposed detector easily finds the target. However, high false-alarm peaks occur in other blocks, and since the test is  $\max_{\Theta} \mathcal{L}'(\Theta) \geq_{H_0}^{\eta} \eta$ , we conclude that the false-alarm rate will be high. But when we compare with the classical detector, we clearly see the superiority of the pro-

posed detector: The classical detector is unable to detect anything at this signal-to-noise ratio. Furthermore, in block #4 the classical detector finds a target at the zero Doppler frequency, whereas the proposed detector finds the target at  $f_d \approx 100$  Hz, close to the simulated Doppler.

For the three cases studied, the superiority of the proposed detection scheme is demonstrated. However, to make

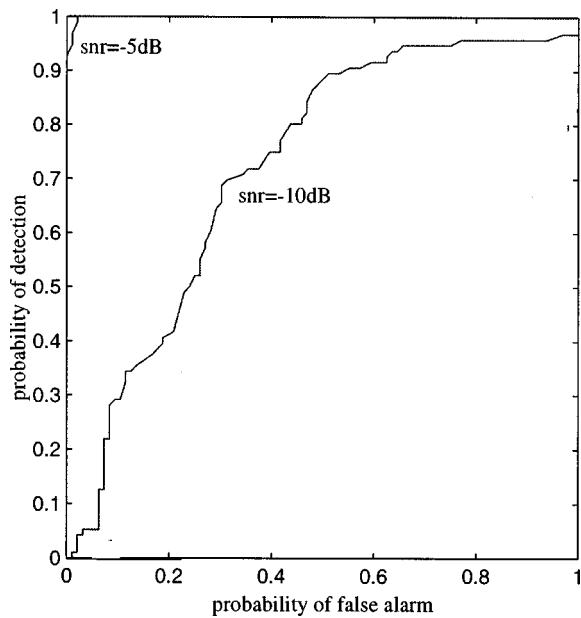


FIG. 14. Experimental ROC curves for the proposed detector of LFM.

this assessment more precise, we now turn to the evaluation of the experimental performance of these detectors.

### C. Experimental performance, comparisons

For the three previously described experiments, we have emitted the signal repeatedly, with a rate of repetition compatible with the listening duration. We have

- $n = 96$  occurrences for LFM signals.
- $n = 128$  occurrences for HFM signals.
- $n = 229$  occurrences for BPSK signals.

For LFM and BPSK signals, the total duration of the experiment to acquire  $n$  occurrences is about 3 min and 2 min,

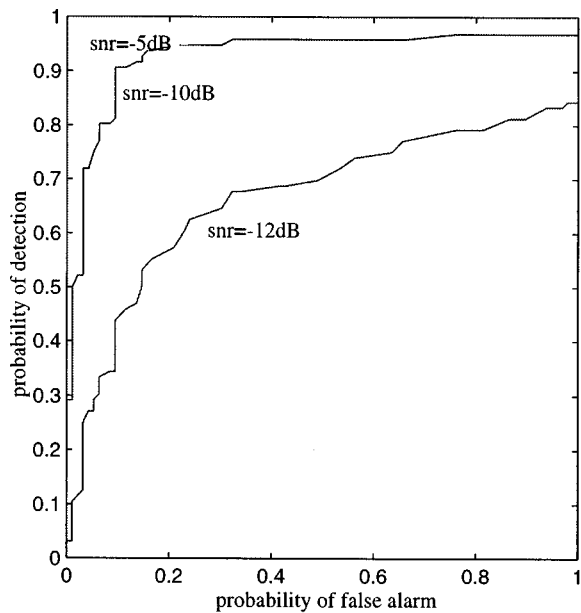


FIG. 15. Experimental ROC curves for the proposed detector of LFM without block #2. The effect of nonstationarity, which is high between blocks #1 and #2, is therefore forgotten. Note that the curve for  $-5$  dB is indeed plotted, but is nearly perfect.

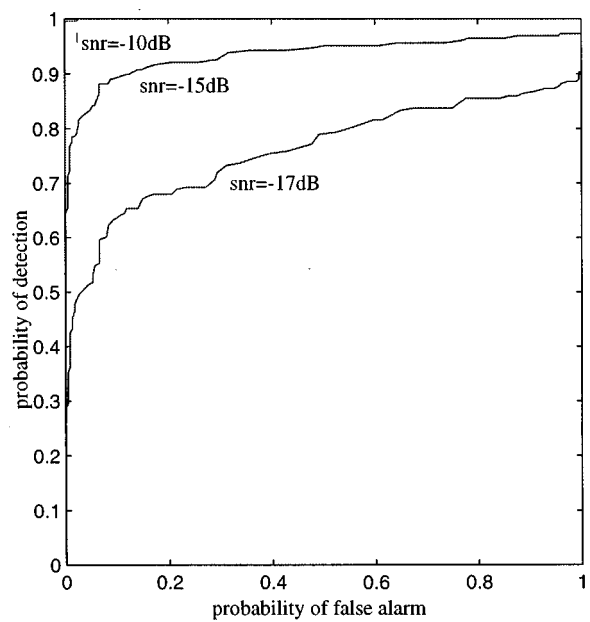


FIG. 16. Experimental ROC curves for the proposed detector of BPSK. Note that the curve for  $-10$  dB is indeed plotted, but is nearly perfect.

respectively. We hence assume that during these periods, the sea surface does not change very much. The  $n$  occurrences can then be viewed as  $n$  different realizations of a random signal. For HFM signals, the duration of the experiment is about 1 h. But, since we are in the presence of reverberation due to an immobile surface, the  $n$  occurrences can also be considered as snapshots of the same random phenomenon.

This allows us to use these occurrences to perform Monte Carlo simulations in order to obtain some experimental receiver operating characteristic curves (ROC curves). These curves represent the detection probability versus the false-alarm probability, at a fixed signal-to-noise ratio. We evaluate these curves experimentally on the real detector, i.e.,

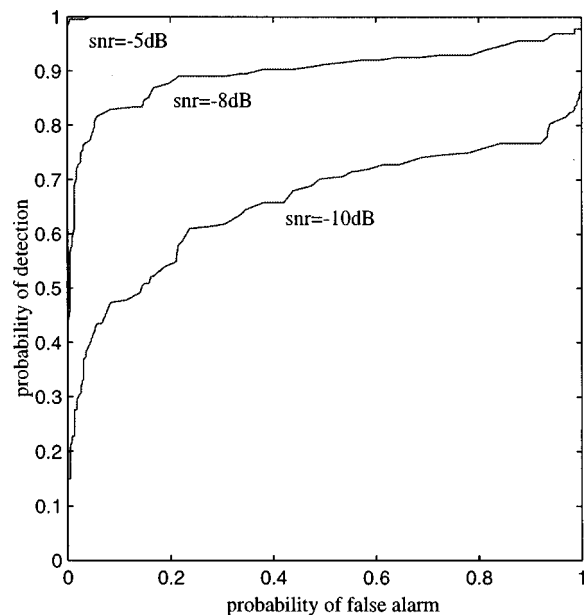


FIG. 17. Experimental ROC curves for the classical detector of BPSK.

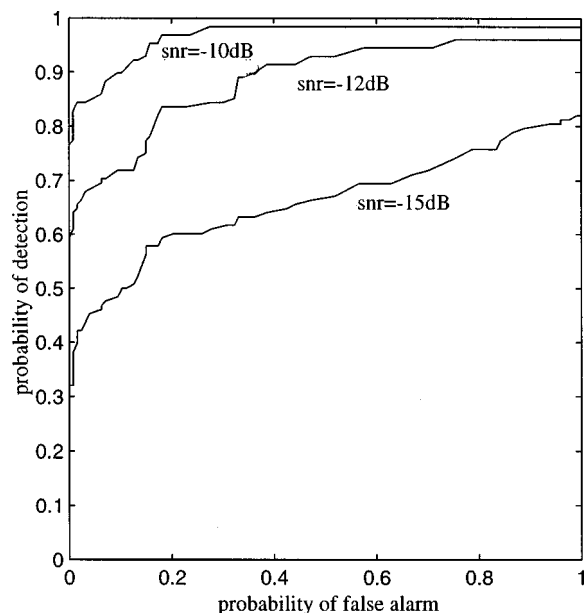


FIG. 18. Experimental ROC curves for the proposed detector of HFM.

$$\mathcal{L}' = \max_{\Theta} \mathcal{L}'(\Theta) \underset{H_0}{\geq} \underset{H_1}{\eta}.$$

The procedure is as follows. We apply the detection algorithm on each occurrence (calculation of  $\mathcal{L}'(\Theta)$  for each block of the occurrence), with and without a synthetic target. Then, for each block of each occurrence, we search for the maximum of  $\mathcal{L}'(\Theta)$ . When no target is present, this allows us to evaluate the probability of false alarms as a function of the threshold. When we have placed a target in block  $\#K_t$ , we take the maximum of  $\mathcal{L}'(\Theta)$  on block  $\#K_t$ . If the value of  $\Theta$  at the maximum is the true value (within a small interval of error), we count it as a good detection (if the maximum is greater than the threshold); otherwise, we count it as 0. Varying the threshold and averaging over the  $n$  occurrences allows us to obtain the estimation of the false-alarm and detection probabilities as functions of the threshold. We can therefore plot the ROC curves. This procedure is then repeated for some signal-to-noise ratios.

Four problems have been examined using the experimental performances of the detectors:

- (1) Influence of the nonstationarity. In Fig. 14 we plot the ROC curves for the LFM signal for a detection scheme that begins at block  $\#2$ . In Fig. 15, the ROC curves have been evaluated without block  $\#2$ , i.e., when the detection starts at block  $\#3$ . We see by comparing these figures that ignoring block  $\#2$  dramatically increases the performance of the detector. This fact is due to the high nonstationarity that exists between blocks  $\#1$  and  $\#2$ , as shown by a high Itakura distance between these two blocks. This nonstationarity induces a bad whitening of block  $\#2$ , and hence a poor quality of detection. More precisely, many false alarms will remain in this block (e.g., see block  $\#2$  of Fig. 8). This illustrates the importance of the local stationarity hypothesis we have made.
- (2) Comparison with the classical detector. This comparison has been made for the three signals, but we present the

results for BPSK signals only. The ROC curves are shown in Fig. 16 for the proposed detector and in Fig. 17 for the classical detector. Comparing these results, we see that for a fixed couple (probability of detection, probability of false alarm) we have a gain of 5 to 8 dB by passing from the classical to the proposed detector! But note that the classical detector has many false alarms corresponding to unrealistic high Dopplers  $f_d$ . Hence, reducing the  $f_d$  range to a more realistic one will attenuate the differences between the two detectors.

- (3) Best signal for detection. In Fig. 18, the ROC curves for HFM signals are plotted. Comparing Figs. 15, 16, and 18, we can conclude that LFM and HFM have roughly the same performances (HFM seem slightly superior, but the phenomenon of reverberation by an immobile surface can be considered as more stable), and that BPSK signals are far better, since we have a gain of about 5 dB for a fixed couple (probability of detection, probability of false alarm).
- (4) Very low signal-to-noise ratios. In the case of very low signal-to-noise ratios, experimental ROC curves are below the diagonal and therefore no longer valid. In this context, choosing this kind of detectors is not the best choice; it is better to toss between  $H_0$  and  $H_1$ !

#### IV. DISCUSSION

To conclude this paper, we would like to discuss some possible improvements to the detector. First, we mentioned the fact that the signal to be detected must be included entirely in one block. This is, of course, unrealistic since we do not know *a priori* its location, and therefore we cannot cut the received signal in such a way that the signal to be detected is in one block. To overcome this problem, we imagine using a sliding window instead of cutting the signals into blocks (of course, this cutting corresponds to a sliding window with a shift of length equal to the length of the window). The algorithm would then be the same: estimation of the covariance through a window, whitening of the signal in the next window, and matched filtering. This would provide a better estimation of the range between the target and the sonar. However, this would greatly increase the time of calculations. A reasonable choice can therefore be to double the size of the window (twice the length  $L$  of the emitted signal) and to make a shift of length  $L$ . In this way, we are sure that the signal to be detected will be entirely in one window. Note that the duration of local stationarity should then be longer than assumed in this paper.

Let us finally mention that the detector has been applied essentially for the detection of a synthetic target. Therefore, the detector was perfectly matched to the target echo. We need to apply it in real situations where a moving target effectively emits an echo. This opens the question of the robustness of the detector. We believe that using phase- or frequency-modulated signals such as BPSKs will lead to a robust detector.

#### ACKNOWLEDGMENT

The authors would like to thank the French DCN Saint-Tropez and DCN Toulon for the use of their facilities during

the experiments. Furthermore, we thank Martha Grand, a great professor of English.

## APPENDIX

In this appendix, we derive the structure of the proposed detector. Note that the problem is

$$\begin{cases} H_0 : \mathbf{z} = \boldsymbol{\varepsilon}, \\ H_1 : \mathbf{z} = A \boldsymbol{\mu}_\Theta + \boldsymbol{\varepsilon}, \end{cases} \quad (\text{A1})$$

where vectors  $\mathbf{z}$ ,  $\boldsymbol{\mu}_\Theta$ ,  $\boldsymbol{\varepsilon}$  contain the  $N$  samples of the corresponding signals. The generalized likelihood ratio test (GLRT) consists of using the classical likelihood ratio test, in which unknowns are replaced by their estimators in the maximum likelihood sense. In this case, the test statistic reads

$$L(\mathbf{z}) = \frac{\max_{A, \Theta, \sigma^2} p_{\mathbf{z}/H_1}(\mathbf{z})}{\max_{\sigma^2} p_{\mathbf{z}/H_0}(\mathbf{z})}.$$

The noise is Gaussian, complex valued, and circular. Therefore, the probability densities read

$$p_{\mathbf{z}/H_0} = \frac{1}{\pi^N \sigma^{2N}} \exp\left(-\frac{\mathbf{z}^\dagger \mathbf{z}}{\sigma^2}\right),$$

$$p_{\mathbf{z}/H_1} = \frac{1}{\pi^N \sigma^{2N}} \exp\left(-\frac{(\mathbf{z} - A \boldsymbol{\mu}_\Theta)^\dagger (\mathbf{z} - A \boldsymbol{\mu}_\Theta)}{\sigma^2}\right).$$

Under  $H_0$ , the maximum likelihood estimator of  $\sigma^2$  is obtained by maximizing  $p_{\mathbf{z}/H_0}(\mathbf{z})$  or equivalently, its logarithm. We classically find  $\sigma_0^2 = \mathbf{z}^\dagger \mathbf{z} / N$ . Under  $H_1$ , we obtain in the same way  $\sigma_1^2 = (\mathbf{z} - A \boldsymbol{\mu}_\Theta)^\dagger (\mathbf{z} - A \boldsymbol{\mu}_\Theta) / N$ . To obtain the estimation of  $A$ , we differentiate again the log-likelihood under  $H_1$  and obtain  $\hat{A} = \boldsymbol{\mu}_\Theta^\dagger \mathbf{z} / \boldsymbol{\mu}_\Theta^\dagger \boldsymbol{\mu}_\Theta$ . The maximization over  $\Theta$  cannot be performed explicitly. Therefore, the GLRT reads

$$L(\mathbf{z}) = \max_{\Theta} \frac{\sigma_0^{2N} \exp(-N)}{\sigma_1^{2N} \exp(-N)},$$

or equivalently

$$\begin{aligned} L(\mathbf{z}) &= \max_{\Theta} 2N \log \left[ \frac{\sigma_0^2}{\sigma_1^2} \right], \\ &= \max_{\Theta} 2N \log \left[ \frac{\mathbf{z}^\dagger \mathbf{z}}{(\mathbf{z} - A \boldsymbol{\mu}_\Theta)^\dagger (\mathbf{z} - A \boldsymbol{\mu}_\Theta)} \right], \\ &= \max_{\Theta} 2N \log \left[ \left( 1 - \frac{|\boldsymbol{\mu}_\Theta^\dagger \mathbf{z}|^2}{(\boldsymbol{\mu}_\Theta^\dagger \boldsymbol{\mu}_\Theta)(\mathbf{z}^\dagger \mathbf{z})} \right)^{-1} \right], \end{aligned}$$

where factor 2 is added for convenience.

- <sup>1</sup>H. L. Van Trees, *Detection, Estimation and Modulation Theory*, Vols. I and III (Wiley, New York, 1968).
- <sup>2</sup>G. Jourdain and J. Munier, "Détection en présence de réverbération," In Proc. Colloque Grets, Nice, 1975), pp. 604–610.
- <sup>3</sup>S. Kay and S. Salisbury, "Improved active sonar detection using autoregressive prewhiteners," J. Acoust. Soc. Am. **87**, 1603–1611 (1990).
- <sup>4</sup>D. E. Bowyer, P. K. Rajasekaran, and W. W. Gebhart, "Adaptive clutter filtering using autoregressive spectral estimation," IEEE Trans. Aerosp. Electron. Syst. **15**(4), 538–546 (1979).
- <sup>5</sup>F. Itakura, "Minimum prediction residual principle applied to speech recognition," IEEE Trans. Acoust., Speech, Signal Process. **ASSP-23**(1), 67–72 (1975).
- <sup>6</sup>M. Basseville, "Distance measures for signal processing and pattern recognition," Signal Process. **18**, 349–369 (1989).
- <sup>7</sup>K. Lashkari, B. Friedlander, J. Abel, and B. McQuiston, "Classification of transient signals," In Proceedings of ICASSP, New York, 1988, pp. 2689–2692.
- <sup>8</sup>B. Picinbono, *Random Signals and Systems* (Prentice-Hall, Englewood Cliffs, NJ, 1993).
- <sup>9</sup>S. Kay, *Modern Spectral Estimation* (Prentice-Hall, Englewood Cliffs, NJ, 1987).
- <sup>10</sup>M. Kendall and A. Stuart, *The Advanced Theory of Statistics, Inference and Relationship*, Vol. 2 (Griffin, London, 1978).
- <sup>11</sup>S. Barbarossa, "Analysis of multicomponent LFM signals by a combined Wigner–Hough transform," IEEE Trans. Signal Process. **43**(6), 1511–1515 (1995).
- <sup>12</sup>G. Jourdain, "Considérations sur la fonction d'ambiguïté dans le cas de signaux aléatoires. Etude et génération de certains de ces signaux." Ph.D. thesis, INP Grenoble, 1970.
- <sup>13</sup>V. Carmillet, "Contribution à la détection en présence de réverbération. Applications en acoustique sous-marine." Ph.D. thesis, INP Grenoble, 1998.
- <sup>14</sup>V. Carmillet and G. Jourdain, "Detection in reverberation," In Proceedings of Oceans'96 MTS/IEEE, Fort Lauderdale, 1996, Vol. 3, pp. 1335–1340.

# Imposing pattern nulls on broadband array responses

Peter J. Kootsookos<sup>a)</sup>

*CRASys, Department of Systems Engineering, RSISE, Australian National University, Canberra, ACT 0200, Australia*

Darren B. Ward<sup>b)</sup>

*Department of Engineering, FEIT, Australian National University, Canberra, ACT 0200, Australia*

Robert C. Williamson

*Department of Engineering, FEIT, Australian National University, Canberra, ACT 0200, Australia*

(Received 20 September 1996; revised 25 November 1998; accepted 25 February 1999)

This paper considers the problem of altering a quiescent design for an array of omni-directional sensors so that the altered design rejects a far-field broadband signal from a given direction. This problem occurs where microphone arrays are to be used to acquire speech signals for telecommunication and interfering signals must be rejected. Three main results are presented in this paper. First, conditions for imposing an exact broadband null upon any given quiescent array response are derived. Second, an analysis of the sensitivity of exact nulling array responses to sensor position error is presented. Third, frequency domain formulations for broadband nulls are obtained and it is shown that these are less restrictive than the exact null conditions which are imposed in the time domain. A Lagrange multiplier approach is used to impose the null conditions upon a quiescent array response to minimize the square error between the quiescent and nulled responses. A design example is given. © 1999 Acoustical Society of America. [S0001-4966(99)02706-X]

PACS numbers: 43.60.Gk [JCB]

## INTRODUCTION

In this paper we consider the problem of altering a far-field quiescent design for an array of omni-directional sensors so that the altered design rejects a far-field broadband signal from a given (known) direction. Without loss of generality, we will use quiescent designs based on the work presented by Ward *et al.*<sup>1</sup> While far-field designs are presented in this paper, similar results hold for the near-field problem.<sup>2</sup>

Microphone arrays for use in telecommunications applications<sup>3</sup> are the motivation for this work. The arrays must acquire a broadband signal, such as speech, while dealing with interference, such as other speakers or office noise.

No adaptive algorithms are examined here, although this is certainly an interesting extension of the current work. Rather, this paper examines what is achievable in the time-invariant setting. This can be considered as a bound on the performance of any adaptive extensions.

In the past, several methods of controlling the positions of the nulls have been employed, including modification of the element amplitude and phase,<sup>4,5</sup> modification of the element phase only,<sup>4,6</sup> and element position perturbation.<sup>7-9</sup>

In phased arrays, where the beamformer weights are fixed as a function of frequency, broadband pattern nulling is effected by imposing a null over a wide spatial region centered on the required broadband null direction. Methods of producing such a null trough include

- imposing multiple pattern nulls,<sup>10-12</sup>
- imposing derivative constraints at the desired direction,<sup>10,13,14</sup> and
- constraining the average power over an angular region,<sup>15,16</sup>

in the vicinity of the required broadband null direction.

The broadband nulling problem this paper addresses is more precisely defined in Sec. I. The problem decomposes into two sub-problems: choice of the cost functional (a squared error cost is chosen) and determination of appropriate constraints. A standard Lagrange multiplier approach is used to solve this optimization problem.

In Sec. II, the conditions for an exact broadband null are derived. These conditions are also presented in a form compatible with the solution approach used. An analysis of the sensitivity of the exact null to sensor position uncertainty is presented. This result allows greater insight into the structural enforcement of pattern nulls and improves on the performance of previous techniques.

Problems with the exact null approach, namely the necessity for oversampling and overly restrictive sensor position constraints, lead to an examination of an alternative approach, multiple frequency nulling, shown in Sec. III.

Both nulling approaches are applied to a frequency invariant quiescent design in Sec. IV. Conclusions are drawn in Sec. V.

## I. PROBLEM STATEMENT

We will now define some notation and define the problem addressed in this paper.

Consider a linear array of  $N$  omni-directional sensors, as illustrated in Fig. 1. Each sensor signal is filtered using a

<sup>a)</sup>Present address: e-Muse Corporation Ltd, 27 Upper Fitzwilliam St., Dublin 2, Ireland.

<sup>b)</sup>Present address: School of Electrical Engineering, University College, The University of New South Wales, Australian Defence Force Academy, Canberra ACT 2600, Australia.

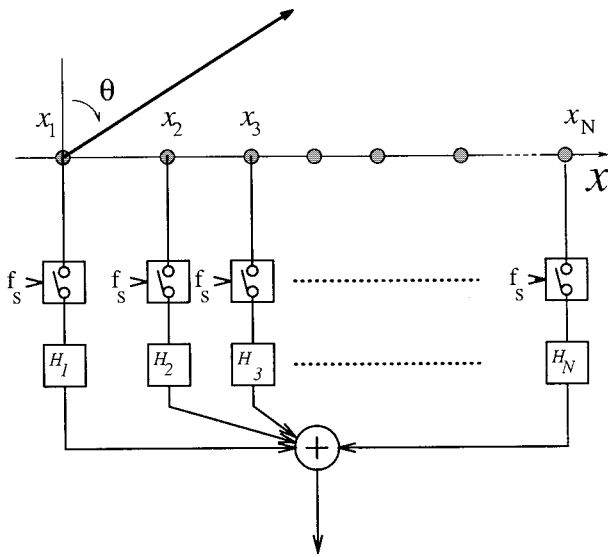


FIG. 1. The array geometry assumed.

causal finite impulse response (FIR) filter,  $h_n[k]$  for  $k = 0, \dots, L-1$ , and the filtered signals are combined to give the following spatial response for a plane wave impinging on the array from a direction  $\theta$  measured relative to broadside:

$$r(u, \omega) = \sum_{n=1}^N H_n(\omega) \exp(j\omega \tau_n(u)), \quad (1)$$

where

$$H_n(\omega) = \sum_{k=0}^{L-1} h_n[k] \exp(-j\omega k)$$

is the frequency response of the  $n$ th sensor filter,  $\omega = 2\pi f/f_s$  is the discrete-time frequency variable with a sampling rate of  $f_s$ ,  $\tau_n(u) = f_s c^{-1} x_n u$  is the relative propagation delay to the  $n$ th sensor with a propagation speed of  $c$ ,  $x_n$  is the location of the  $n$ th sensor, and  $u = \sin \theta$ .

Let the  $NL$ -dimensional delay vector be

$$\mathbf{d}(u, \omega) = \mathbf{e}(\omega) \otimes \mathbf{a}(u, \omega),$$

where  $\otimes$  indicates a Kronecker product, the  $L$ -dimensional discrete Fourier transform vector is

$$\mathbf{e}(\omega) = [1, e^{-j\omega}, \dots, e^{-j\omega(L-1)}]^T,$$

and the  $N$ -dimensional array response vector is

$$\mathbf{a}(u, \omega) = [e^{j\omega \tau_1(u)}, e^{j\omega \tau_2(u)}, \dots, e^{j\omega \tau_N(u)}]^T.$$

The array response may then be written in vector form as

$$r(u, \omega) = \mathbf{h}^H \mathbf{d}(u, \omega),$$

where the  $NL$  vector of complex FIR filter coefficients is

$$\mathbf{h} = [h_1[0] \cdots h_N[0] \cdots h_1[L-1] \cdots h_N[L-1]]^H,$$

and  $x^H$  denotes the Hermitian transpose of  $x$ .

We wish to find the coefficients  $\hat{\mathbf{h}}$  that produce a similar array response to  $\mathbf{h}$ , but with a broadband null in a given direction. We will assume that coefficients  $\hat{\mathbf{h}}$  are formed by adding a set of perturbing coefficients to  $\mathbf{h}$ ; i.e.,  $\hat{\mathbf{h}} = \mathbf{h} + \mathbf{b}$ .

The resulting response  $\hat{r}(u, \omega)$  is equivalent to the quiescent response  $r(u, \omega)$  plus the response of a nulling beamformer with coefficients  $\mathbf{b}$ :

$$\begin{aligned} \hat{r}(u, \omega) &= \hat{\mathbf{h}}^H \mathbf{d}(u, \omega) \\ &= \mathbf{h}^H \mathbf{d}(u, \omega) + \mathbf{b}^H \mathbf{d}(u, \omega) \\ &= r(u, \omega) + \mathbf{b}^H \mathbf{d}(u, \omega). \end{aligned}$$

**Problem 1 (Imposing a Broadband Null).** The general broadband nulling problem can be formulated as: determine the coefficient vector  $\mathbf{b}$  that minimizes

$$\min_{\mathbf{b}} J(\mathbf{b}) \quad (2)$$

$$\text{subject to } |\hat{r}(u_0, \omega)| \leq \epsilon, \quad \forall \omega \in \Omega, \quad (3)$$

where  $J$  is some suitably defined cost functional that measures the distance between  $r(u, \omega)$  and  $\hat{r}(u, \omega)$ ,  $\epsilon$  is the desired null depth in the nulling direction  $u_0$ , and  $\Omega$  is the set of frequencies of interest (typically the bandwidth of the source and interfering signals).

Several candidate cost functionals and null-imposing constraints will now be considered.

## A. Cost functionals

The choice of  $J$  in Eq. (2) is very application dependent. In many situations use of a min-max (or  $L_\infty$ ) error criterion is of interest, however, computationally simple procedures to solve such problems occur only rarely<sup>17</sup> and closed-form solutions generally exist only for trivial cases.

As we wish to find closed-form, computationally simple functionals we will only consider square-error ( $L_2$ ) techniques. Any one of several candidate functionals could be used, including:

- minimizing the weighted square distance between the quiescent response and the perturbed response,
- minimizing the output power of the nulling beamformer (which would be appropriate in an adaptive environment), and
- minimizing the gain of the nulling beamformer.

Although there is no one ‘‘right’’ choice of  $J$ , the last one listed above is by far the simplest to compute, and it will be considered exclusively in the remainder of this paper (for a more complete treatment of cost functionals see Ref. 18). Specifically, we wish to minimize the  $L_2$  norm between the quiescent coefficients and the perturbed coefficients:

$$J_{NG} = \|\hat{\mathbf{h}} - \mathbf{h}\|^2 = \|\mathbf{b}\|^2,$$

where  $\|\mathbf{b}\| = (\mathbf{b}^H \mathbf{b})^{1/2}$  is the  $L_2$  norm of  $\mathbf{b}$ .

## B. Constraints

Having determined a cost functional, the practicality of the constraint in Eq. (3) must be assessed. The main results of this paper allow the following constraints to be imposed.

## 1. Exact nulling

Consider the following problem:

$$\min_{\mathbf{b}} \mathbf{b}^H \mathbf{b} \quad (4)$$

$$\text{subject to } \hat{r}(u_0, \omega) = 0, \quad \forall \omega. \quad (5)$$

This corresponds to a constraint on the time domain impulse responses of the sensor filters, as shown in Sec. II.

## 2. Multiple frequency nulling

An alternative to an exact broadband null is to impose multiple zeros in the frequency domain response of the beamformer at the null angle. In this case, the problem is formulated as:

$$\min_{\mathbf{b}} \mathbf{b}^H \mathbf{b} \quad (6)$$

$$\text{subject to } \hat{r}(u_0, \omega) = 0, \quad \omega = \omega_1, \dots, \omega_M. \quad (7)$$

The challenge now is in choosing the set of  $M$  frequencies  $\{\omega_m\}_{m=1}^M$  at which to impose the frequency domain zeros in order to obtain a broadband null of a given depth. This problem is considered in Sec. III.

## C. Producing a broadband null in a quiescent response

The problem of producing a broadband null in a quiescent broadband response while perturbing that pattern the least with regard to  $J_{NG}$  may be tackled as follows. If  $r(u, \omega) = \mathbf{h}^H \mathbf{d}(u, \omega)$  is the quiescent beamformer response, find coefficients  $\mathbf{b}$  such that

$$\hat{r}(u, \omega) = (\mathbf{h}^H + \mathbf{b}^H) \mathbf{d}(u, \omega) \quad (8)$$

is close to  $r(u, \omega)$  as measured by  $J_{NG}$  and a broadband null is imposed.

The optimum  $\mathbf{b}$  is found as the solution to the constrained optimization problem:

$$\min_{\mathbf{b}} \mathbf{b}^H \mathbf{b} \quad (9)$$

$$\text{subject to } \mathbf{C}^H (\mathbf{h} + \mathbf{b}) = \mathbf{0}_K, \quad (10)$$

where  $\mathbf{C}$  is a  $NL \times K$  constraint matrix,  $\mathbf{0}_K$  is the  $K$  vector of zeros, and  $K$  represents the finite number of constraints to be imposed.

The solution to this constrained optimization problem may be found using Lagrange multipliers as stated in the following Proposition; see Frost<sup>19</sup> for a more complete treatment.

**Proposition 1 (Coefficients for a Broadband Null).** *The optimum perturbations  $\mathbf{b}_{\text{opt}}$  from the quiescent weights  $\mathbf{h}$  to ensure a broadband null in the required direction are given by*

$$\mathbf{b}_{\text{opt}} = -\mathbf{C} [\mathbf{C}^H \mathbf{C}]^{-1} \mathbf{C}^H \mathbf{h}. \quad (11)$$

Having defined a suitable cost functional and examined the solution technique to be used, the next two sections show how the two broadband null constraints may be placed in the

solution framework. That is, we identify suitable choices for  $\mathbf{C}$  to impose broadband nulls.

## II. EXACT NULLING

We now show how it is possible through proper array design to produce an exact pattern null (a pattern zero that is present over all frequencies). However, we also show that stringent requirements are imposed on the array geometry and sampling rate to achieve this. In Sec. IID we evaluate the degradation of an exact null that would occur in a practical setting with sensor location errors.

### A. Exact broadband nulling

The requirements for an exact null are presented in Proposition 2 and its Corollary.

**Proposition 2 (Condition for a Broadband Null).** *A broadband null at  $u_0$  will be available if and only if either*

$$r(u_0, \omega) = \sum_{n=1}^N H_n(\omega) e^{j\omega \tau_n(u_0)} = 0, \quad \forall \omega \quad (12)$$

or, equivalently,

$$\sum_{n=1}^N (h_n[k] * \text{sinc}(k + \tau_n(u_0))) = 0, \quad \forall k \quad (13)$$

where  $*$  denotes convolution in the  $k$  index and  $\text{sinc}(x) = \sin(\pi x)/(\pi x)$ .

**Proof:** From Eq. (1), the response in direction  $u_0$  is

$$\begin{aligned} r(u_0, \omega) &= \sum_{n=1}^N H_n(\omega) e^{j\omega \tau_n(u_0)} \\ &= \sum_{n=1}^N \sum_{k=0}^{L-1} h_n[k] e^{-j\omega k} e^{j\omega \tau_n(u_0)} \\ &= \sum_{k=0}^{L-1} \left( \sum_{n=1}^N h_n[k] * \text{sinc}[k + \tau_n(u_0)] \right) e^{-j\omega k}. \end{aligned} \quad (14)$$

Equation (14) yields Eq. (12) and the inverse discrete-time Fourier transform of Eq. (15) yields Eq. (13).

It is not immediately clear how Eqs. (12) and (13) may be easily enforced. The following result shows this.

**Corollary 1 (Integer Delay Property).** *If  $\tau_n(u_0)$  is an integer for all  $n$ , and*

$$\sum_{n=1}^N h_n[k + \tau_n(u_0)] = 0, \quad \forall k, \quad (16)$$

then Eq. (13) is satisfied.

**Proof:** If  $\tau_n(u_0)$  is an integer for all  $n$ , then Eq. (13) becomes

$$\sum_{n=1}^N h_n[k] * \delta[k + \tau_n(u_0)] = 0, \quad \forall k,$$

where

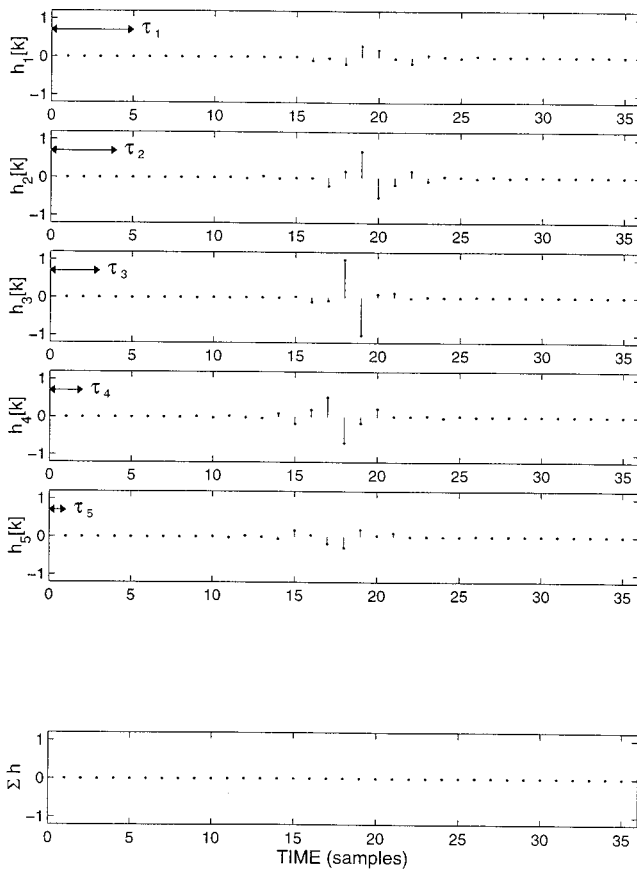


FIG. 2. Time domain nulling constraint.

$$\delta[k] = \begin{cases} 1, & k=0 \\ 0, & k \neq 0. \end{cases}$$

The null constraint Eq. (16) is illustrated in Fig. 2. Specifically, when  $\tau_n(u_0)$  is an integer for all  $n$ , the sum of the delayed filter impulse responses can be made exactly zero.

Frost<sup>19</sup> noted that, for a signal arriving from broadside, the broadband beamformer with  $N$  sensors and  $L$  taps per sensor was equivalent to an  $L$  tap filter whose coefficients were the sum over  $N$  sensors. The integer delay property of the Corollary is equivalent, but for delays other than  $\tau_n(u_0)=0, \forall n$ .

### B. Constraints for exact nulling

Note that it is always possible to place a null at  $u_0=0$  because in this case

$$\tau_n(u_0) = \frac{f_s}{c} x_n u_0 = 0, \quad \forall n$$

and Eq. (16) reduces to requiring

$$\sum_{n=1}^N h_n[k] = 0, \quad \forall k.$$

Using this idea, the condition for an exact null at broadside may be written as

$$\mathbf{C}_0^H \mathbf{h} = \mathbf{0}_L,$$

where

$$\mathbf{C}_0 = \mathbf{I}_L \otimes \mathbf{1}_N,$$

where  $\mathbf{I}_L$  is the  $L \times L$  identity matrix and  $\mathbf{1}_N$  is the  $N$ -vector of ones. The resulting  $\mathbf{C}_0$  is an  $NL \times L$  matrix.

To demonstrate how a null may be placed at directions other than broadside, we rewrite the integer delay property Eq. (16) as

$$\sum_{n=1}^N \tilde{h}_n[k] = 0,$$

where  $\tilde{h}_n[k] = h_n[k + \tau_n(u_0)]$ . The  $\tilde{h}_n[k]$  are effectively the filter coefficients that steer the null direction to broadside. Thus if the coefficients  $\tilde{h}_n[k]$  were used as the beamformer filter coefficients, the resulting response would be steered to  $-u_0$  and the null would appear at broadside. This null steering is only exact for  $\tau_n(u_0) \in \mathbb{Z}, \forall n$ .

The null constraint can now be written as

$$\mathbf{C}_0^H \tilde{\mathbf{h}} = \mathbf{0}.$$

Define a linear transformation matrix  $\mathbf{T}_u$  that satisfies  $\tilde{\mathbf{h}} = \mathbf{T}_u \mathbf{h}$ . The null constraint now becomes

$$\mathbf{C}^H \mathbf{h} = \mathbf{0}, \tag{17}$$

where  $\mathbf{C} = \mathbf{T}_u^H \mathbf{C}_0$  is the transformed constraint matrix.

### C. Implications for practical array design

It is instructive to consider what the integer delay property implies for a practical design. The requirement for a broadband null is

$$\tau_n(u_0) = f_s c^{-1} x_n u_0 \in \mathbb{Z}, \quad n = 1, 2, \dots, N.$$

Let  $x_1 = 0$  and  $x_n > 0, n > 1$ , i.e., we are considering a single-sided linear array. The first constraint on the array geometry is

$$\frac{x_n}{d_0} \in \mathbb{Z}, \quad n > 1, \tag{18}$$

regardless of the desired null angle, where we will refer to  $d_0$  as the *fundamental spacing*. Second, we require

$$d_0 = \frac{mc}{f_s u_0}, \quad m \in \mathbb{Z}. \tag{19}$$

Assume that the sensor spacings are logarithmically increasing,<sup>20</sup> and the part of the array closest to the origin is used only at the highest frequency, with more and more elements becoming active at lower frequencies. Assuming we want to use the minimum number of sensors to avoid spatial aliasing, then

$$d_0 = \frac{c}{2f_U},$$

where  $f_U$  is the upper frequency of interest. The directions in which we may form a broadband null are then

$$u_0 = \frac{m2f_U}{f_s}, \quad m \in \mathbb{Z}.$$

Clearly, for a minimum sampling rate of  $f_s = 2f_U$  it is only possible to produce a broadband null at  $u_0 \in \{-1, 0, 1\}$ , or



equivalently  $\theta_0 \in \{-\pi/2, 0, \pi/2\}$ . Hence, the minimum sampling rate required for a null at  $u_0$  is

$$f_s = \frac{2f_U}{u_0} \geq 2f_U. \quad (20)$$

This demonstrates the major disadvantages of time domain nulling: Oversampling is required to produce a null that is away from broadside or endfire, and the sensor locations must be quantized to multiples of the fundamental spacing. These problems can be overcome by using interpolation beamforming<sup>21</sup> or fractional sample delay FIR filters.<sup>22,23</sup> However, since FIR filters must be quite long to obtain accurate fractional sample delays, these techniques will result in significantly longer sensor filters.

#### D. Effect of sensor position errors on a broadband null

Given that the optimum nulling coefficients have been determined to produce an exact null in a given direction, we now consider the effect of sensor position errors on the expected null depth. We follow the standard method in the antenna literature.<sup>24</sup>

Assume the actual position of the  $n$ th sensor is  $\hat{x}_n = x_n + \chi_n$ , where  $x_n$  is the ideal sensor location and the  $\chi_n$  are independent zero-mean Gaussian random variables with variance  $\sigma_x^2$ . We assume that the location errors are uncorrelated from sensor to sensor.

The actual broadband response is

$$\hat{r}_a(u, \omega) = \sum_{n=1}^N \hat{H}_n(\omega) e^{j\omega \hat{\tau}_n(u)}, \quad (21)$$

where

$$\hat{H}_n(\omega) = \sum_{k=0}^{L-1} (h_n[k] + b_n[k]) e^{-j\omega k},$$

with the  $b_n[k]$  calculated from Eq. (11) assuming ideal sensor locations. The actual propagation delay to the  $n$ th sensor is

$$\hat{\tau}_n(u) = f_s c^{-1} (x_n + \chi_n) u = \tau_n(u) + \Delta_n(u). \quad (22)$$

We now have

$$\begin{aligned} |\hat{r}_a(u, \omega)|^2 &= \sum_{n=1}^N \sum_{m=1}^N \hat{H}_n(\omega) \hat{H}_m^*(\omega) e^{j\omega[\tau_n(u) - \tau_m(u)]} \\ &\quad \times e^{j\omega[\Delta_n(u) - \Delta_m(u)]} + \sum_{n=1}^N |\hat{H}_n(\omega)|^2. \end{aligned}$$

The expected value of the beampattern is

$$\begin{aligned} E[|\hat{r}_a(u, \omega)|^2] &= \sum_{n=1}^N \sum_{m=1}^N \hat{H}_n(\omega) \hat{H}_m^*(\omega) e^{j\omega[\tau_n(u) - \tau_m(u)]} \\ &\quad \times |E_\Delta(u, \omega)|^2 + \sum_{n=1}^N |\hat{H}_n(\omega)|^2, \quad (23) \end{aligned}$$

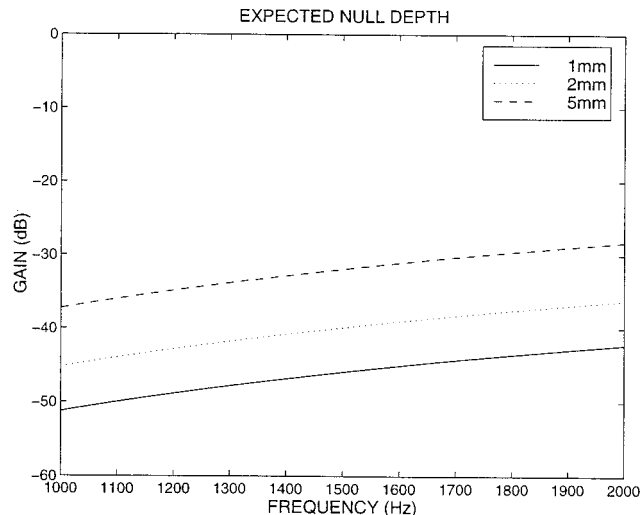


FIG. 3. Expected null depth for  $\sigma_x = 1$  mm (—),  $\sigma_x = 2$  mm (···), and  $\sigma_x = 5$  mm (--).

where  $E_\Delta(u, \omega) = E[e^{j\omega f_s c^{-1} \chi u}]$ . Noting that  $E_\Delta(u, \omega)$  is the characteristic function of the Gaussian random variable  $\chi$ , gives<sup>25</sup>

$$E_\Delta(u, \omega) = \exp(-\sigma_x^2 (\omega f_s c^{-1} u)^2 / 2). \quad (24)$$

Recall that the ideal beampattern is

$$\begin{aligned} |\hat{r}(u, \omega)|^2 &= \sum_{n=1}^N \sum_{m=1}^N \hat{H}_n(\omega) \hat{H}_m^*(\omega) e^{j\omega[\tau_n(u) - \tau_m(u)]} \\ &\quad + \sum_{n=1}^N |\hat{H}_n(\omega)|^2. \end{aligned}$$

Hence, the actual beampattern can be expressed as

$$\begin{aligned} E[|\hat{r}_a(u, \omega)|^2] &= |\hat{r}(u, \omega)|^2 |E_\Delta(u, \omega)|^2 \\ &\quad + \sum_{n=1}^N |\hat{H}_n(\omega)|^2 (1 - |E_\Delta(u, \omega)|^2). \end{aligned}$$

Assuming that the ideal pattern has an exact null at  $u_0$ , then  $\hat{r}(u_0, \omega) = 0, \forall \omega$ , and the expected actual null direction response is

$$\begin{aligned} E[|\hat{r}_a(u_0, \omega)|^2] &= \sum_{n=1}^N |\hat{H}_n(\omega)|^2 \\ &\quad \times \left( 1 - \exp\left(-\sigma_x^2 \frac{(\omega f_s c^{-1} u_0)^2}{2}\right) \right)^2 \quad (25) \\ &\approx \sigma_x^2 (\omega f_s c^{-1} u_0)^2 \sum_{n=1}^N |\hat{H}_n(\omega)|^2, \end{aligned}$$

for small  $\sigma_x$ .

(26)

The expected null depth for several different positioning errors, for the array design detailed in Sec. IV, is shown in Fig. 3. Unsurprisingly, the depth of the null decreases as the sensor position errors increase.

### III. MULTIPLE FREQUENCY NULLING

In this section we consider an alternative to exact broadband nulling that avoids the oversampling problem identified in exact nulling. Specifically, we consider the problem of placing multiple zeros in the null direction frequency response such that a given null depth is achieved while minimizing the disturbance to the quiescent broadband response.

#### A. Constraints for multiple zeros in null direction frequency response

Let  $r(u, \omega) = \mathbf{h}^H \mathbf{d}(u, \omega)$  be the quiescent broadband response. The constraint we wish to impose here is

$$\mathbf{d}(u_0, \omega_m)^H (\mathbf{h} + \mathbf{b}) = 0, \quad (27)$$

for  $m = 1, \dots, M$ . The cost functional is still the nulling gain  $J_{NG}$ .

This problem is identical to Eqs. (9) and (10) with  $\mathbf{C} = [\mathbf{d}(u_0, \omega_1), \dots, \mathbf{d}(u_0, \omega_M)]$ . The resulting  $\mathbf{C}$  is an  $NL \times M$  matrix. The solution is again given by Eq. (11).

Clearly, we are not limited to only placing multiple frequency nulls at the direction  $u_0$ . A wider null may be specified by placing further nulls at a direction  $u_0 + \Delta u_0$  (this is demonstrated by an example in Sec. IV).

It is now necessary to determine the zero locations  $\omega_m$ ,  $m = 1, \dots, M$  in order to achieve the required null depth. As a simplification we will assume the  $M$  zeros are equally spaced within the design frequency band, so the problem now is only to choose the number of zeros.

#### B. Determining the required number of frequency zeros

Consider the problem of determining the number  $M$  of equally spaced zeros to impose in the null direction frequency response in order to achieve a given null depth within the design frequency band  $\Omega = [\omega_L, \omega_U]$ . An analogous problem was considered by Steyskal<sup>11</sup> in which he determined the number of constraints required to achieve a required null depth over a spatial sector for a phased array. Steyskal's method can be modified as follows.

The constraint Eq. (27) may be rewritten as

$$\hat{r}(u_0, \omega_m) = 0, \quad (28)$$

for  $m = 1, \dots, M$ .

Steyskal<sup>10</sup> showed that the weights which solve the constrained optimization are given by

$$\hat{\mathbf{h}} = \mathbf{h} - \sum_{m=1}^M \alpha_m \mathbf{d}_m,$$

where  $\mathbf{d}_m = \mathbf{d}(u_0, \omega_m)$ . Equivalently, the optimum response is given by

$$\hat{r}(u, \omega) = r(u, \omega) - \sum_{m=1}^M \alpha_m q^{(m)}(u, \omega), \quad (29)$$

where

$$q^{(m)}(u, \omega) = \mathbf{d}_m^H \mathbf{d}(u, \omega). \quad (30)$$

The parameters  $\alpha_m$  are obtained by solving the set of  $M$  simultaneous equations

$$\begin{bmatrix} r(u_0, \omega_1) \\ \vdots \\ r(u_0, \omega_M) \end{bmatrix} = \begin{bmatrix} q^{(1)}(u_0, \omega_1) & \cdots & q^{(M)}(u_0, \omega_1) \\ \vdots & & \vdots \\ q^{(1)}(u_0, \omega_M) & \cdots & q^{(M)}(u_0, \omega_M) \end{bmatrix} \times \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_M \end{bmatrix}. \quad (31)$$

Define the null depth as

$$\epsilon = \max_{\omega \in \Omega} |\hat{r}(u_0, \omega)|^2, \quad (32)$$

where  $\hat{r}(u, \omega)$  is given by Eq. (29), and  $q^{(m)}(u, \omega)$  is given by Eq. (30).

#### 1. Simplification for a frequency invariant beamformer

It is extremely difficult to gain any insight into the variation of  $\epsilon$  with  $M$  using the general equations above. Hence, in this section our aim is to determine  $M$  for the specific case where the quiescent beamformer is a frequency invariant beamformer (FIB). We will see that in this case it is relatively simple to determine *a priori* the required value of  $M$  for a desired null depth.

Consider the following assumptions.

- (1) The  $M$  nulls are equally spaced over the region  $\Omega$ .
- (2) For an FIB, the quiescent response in the nulling direction is approximately constant over the nulling frequency band, that is,  $r(u_0, \omega) \approx A$ ,  $\omega \in \Omega$ , where  $A = |r(u_0, \omega_L)|$ .
- (3)  $\hat{r}(u_0, \omega)$  is oscillatory and crosses the  $\omega$  axis at  $\omega = \omega_1, \dots, \omega_M$ .
- (4)  $\hat{r}(u_0, \omega)$  attains its maximum value at  $\omega_{\max} = (\omega_1 + \omega_2)/2$ .

Of these assumptions, only the last one requires further comment. When  $M$  equally spaced frequency nulls are imposed within a region  $\Omega$ , the null-direction response will exhibit  $(M-1)$  lobes within  $\Omega$ . The peak of each lobe will occur approximately midway between the adjacent pair of null frequencies. From many simulations, we have observed that it is generally the first or last lobe that has maximum height. For simplicity, we will assume that the first lobe has maximum height.

With these assumptions, the null depth is given by

$$\epsilon = \left| \hat{r}\left(u_0, \frac{\omega_1 + \omega_2}{2}\right) \right|^2, \quad (33)$$

where

$$\hat{r}(u_0, \omega) = A - \sum_{m=1}^M \alpha_m q^{(m)}(u_0, \omega), \quad (34)$$

with  $q^{(m)}(u_0, \omega)$  given by Eq. (30), and  $\alpha_m$  given by the solution of Eq. (31) with  $r(u_0, \omega_m) = A$ ,  $m = 1, \dots, M$ .

#### 2. Incremental null depth

At this point it is useful to define the *incremental null depth*,  $\epsilon_{\text{inc}}$  as the null depth imposed in the direction  $u_0$  relative to the quiescent FIB response, i.e.,

$$\epsilon_{\text{inc}} \triangleq |\hat{r}(u_0, \omega_{\text{max}}) - r(u_0, \omega_{\text{max}})|^2, \quad (35)$$

which for a FIB reduces to

$$\epsilon_{\text{inc}} = |\hat{r}(u_0, \omega_{\text{max}}) - A|^2. \quad (36)$$

Hence, from Eq. (34) we have

$$\epsilon_{\text{inc}} = \left| \sum_{m=1}^M \alpha_m q^{(m)}(u_0, \omega_{\text{max}}) \right|^2, \quad (37)$$

with  $\alpha_m$  given by the solution of Eq. (31) with  $r(u_0, \omega_m) = 1$ ,  $m = 1, \dots, M$ .

Although Eq. (37) falls short of providing an explicit expression for  $M$  in terms of  $\epsilon_{\text{inc}}$ , it does, however, provide a simple means of calculating  $\epsilon_{\text{inc}}$  for a given  $M$ . This calculation depends on the array geometry, number of filter coefficients, sampling rate, design bandwidth, and null direction. For a given beamformer all of these variables will be fixed. Importantly, however, the calculation does not depend on the filter coefficients themselves (only on the number of filter coefficients used). Thus the incremental null depth achieved for a given  $M$  is independent of the quiescent beampattern; even if the FIB filter coefficients are determined from an adaptive algorithm, the incremental null depth in the null direction will be as determined by Eq. (37).

### C. Effect of sensor position errors

In Sec. II we derived an expression, Eq. (26), for the expected null depth in a practical situation where sensor position errors are present. In the case of multiple frequency nulls this equation becomes

$$E[|\hat{r}_a(u_0, \omega)|^2] = |\hat{r}(u_0, \omega)|^2 |E_{\Delta}(u_0, \omega)|^2 + \sum_{n=1}^N |\hat{H}_n(\omega)|^2 (1 - |E_{\Delta}(u_0, \omega)|^2),$$

where  $\hat{r}(u_0, \omega)$  is the ideal null direction response (assuming ideal sensor locations).

## IV. DESIGN EXAMPLES

In order to illustrate the methods developed above, consider the following example.

A broadband array was designed to operate in the band 1–2 kHz, with an aperture size of six half-wavelengths (for acoustic waves in air) and a sampling rate of  $f_s = 8$  kHz. The null was to be located at  $\theta = 30^\circ$ . For exact nulling, this required a fundamental spacing of  $d_0 = 0.085$  m. The resulting sensor positions are tabulated in Table I.

### A. Quiescent response

The quiescent response was found following the frequency invariant beamformer (FIB) ideas presented by Ward *et al.*<sup>1,26</sup> The FIR filters each had  $L = 32$  coefficients. The quiescent array response is shown in Fig. 4.

TABLE I. Sensor locations and sensor-to-sensor delays relative to center sensor for a source at  $30^\circ$ .

Sensor position (m)	Time delay (samples)
-0.5100	-6
-0.4250	-5
-0.3400	-4
-0.2550	-3
-0.1700	-2
-0.0850	-1
0	0
0.0850	1
0.1700	2
0.2550	3
0.3400	4
0.4250	5
0.5100	6

### B. Exact nulling

Placing a null at  $\theta = 30^\circ$  corresponds to one of the quiescent array sidelobes. Using the techniques of Sec. II, the array response of the exact nulling design is shown in Fig. 5.

Note, from Table I, that the sensor-to-sensor delays (for a signal arriving at  $\theta = 30^\circ$ ) in this array are integer numbers of samples.

### C. Multiple frequency nulling

In order to avoid the problem of integer delays (or the associated problem of large increases in filter order required by fractional delay FIR filters) the multiple frequency nulling technique can be used. Applied in this example, with  $M = 10$  and nulling direction again  $\theta = 30^\circ$ , the resulting array response is shown in Fig. 6.

Figure 7 shows the multiple frequency technique applied at bearings of  $27.5^\circ$  and at  $32.5^\circ$ . Ten nulls were placed at each bearing. The purpose of this is to produce a null over a wider range of bearings.

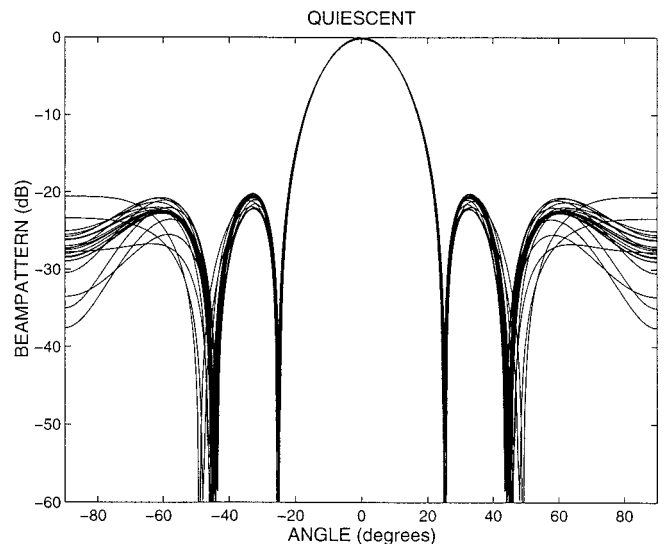


FIG. 4. Quiescent array response at 21 equispaced frequencies within the design band.

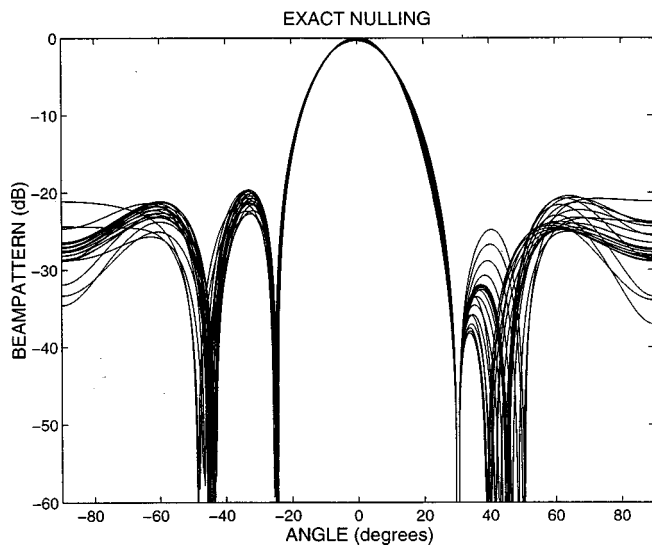


FIG. 5. Array response with exact null at  $30^\circ$  at 21 equispaced frequencies within the design band.

#### D. Null depths

The array responses of the quiescent and multiple frequency nulls at the required null direction are displayed in Fig. 8, normalized with respect to the quiescent null direction response at 1 kHz, i.e., these plots show the incremental null depth that will be obtained relative to the quiescent response. Observe that the predicted null depth ( $-46$  dB) is very close to the actual null depth ( $-49$  dB).

The exact null depth is not shown as it was zero for all intents and purposes (lower than 200 dB over the frequency range of interest).

The predicted incremental null depth for the multiple frequency nulling approach Eq. (37) is plotted versus number of nulls in Fig. 9. Also shown is the actual incremental null depth, which demonstrates that the predicted null depth is in good agreement with the actual depth over a large range of  $M$ . This demonstrates the validity of the prediction technique derived in Sec. III.

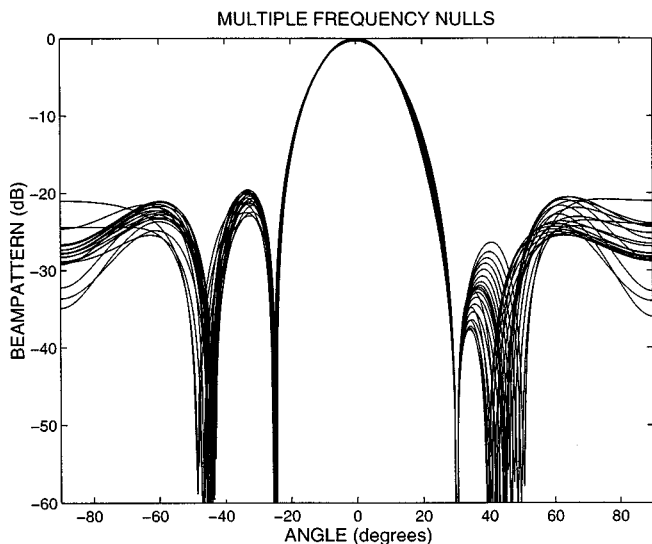


FIG. 6. Array response with multiple frequency nulls ( $M=10$ ) at  $30^\circ$  at 21 equispaced frequencies within the design band.

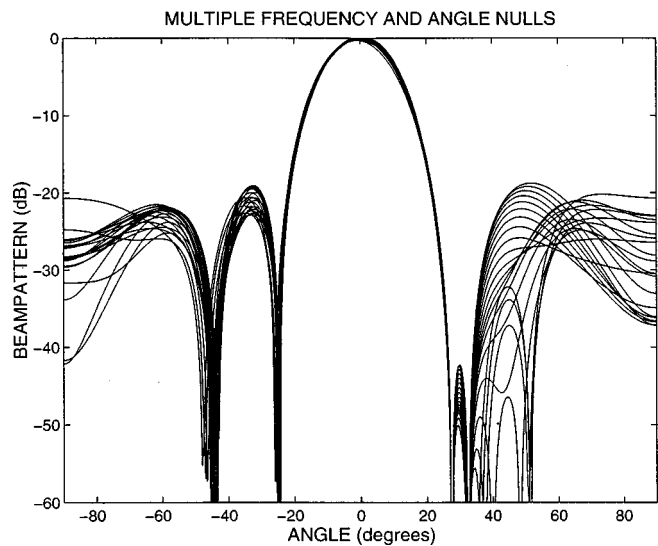


FIG. 7. Array response with multiple frequency nulls ( $M=10$ ) at  $27.5^\circ$  and multiple frequency nulls ( $M=10$ ) at  $32.5^\circ$  at 21 equispaced frequencies within the design band.

#### V. CONCLUSION

The broadband nulling problem considered in this paper was formulated as follows. Given an  $NL$  vector of filter coefficients that produces some desired broadband response  $r(u, \omega)$  for a beamformer with  $N$  sensors and  $L$  filter taps per sensor, find the coefficients that produce a broadband response  $\hat{r}(u, \omega)$  which has a broadband null in a specified direction and was close in some respect to the original response  $r(u, \omega)$ . The problem was formulated in terms of the following constrained minimization problem: minimize the distance between  $r(u, \omega)$  and  $\hat{r}(u, \omega)$  subject to the constraint that  $\hat{r}(u, \omega)$  exhibits a broadband null in the direction  $u = u_0$ .

Two null constraints were considered: one which placed an exact null in the null direction over all frequencies and one which placed multiple single-frequency nulls in the required null direction.

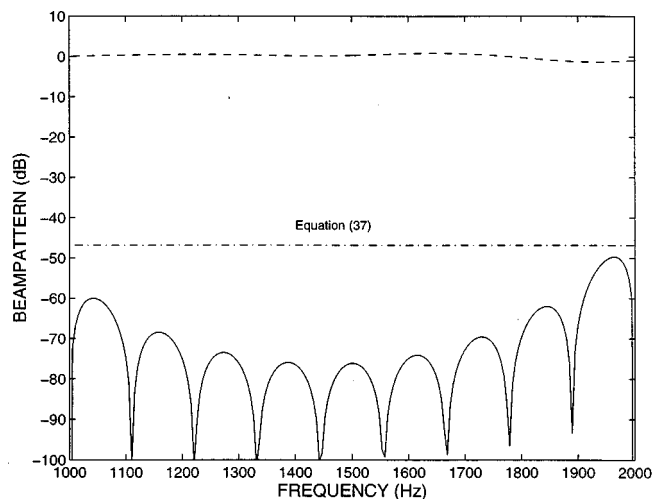


FIG. 8. Array responses of quiescent design (--) and multiple frequency null design (—) at the null direction ( $\theta=30^\circ$ ) over the array design frequency band. The predicted null depth from Eq. (37) for the multiple frequency nulling approach with  $M=10$  is also shown (-.-).

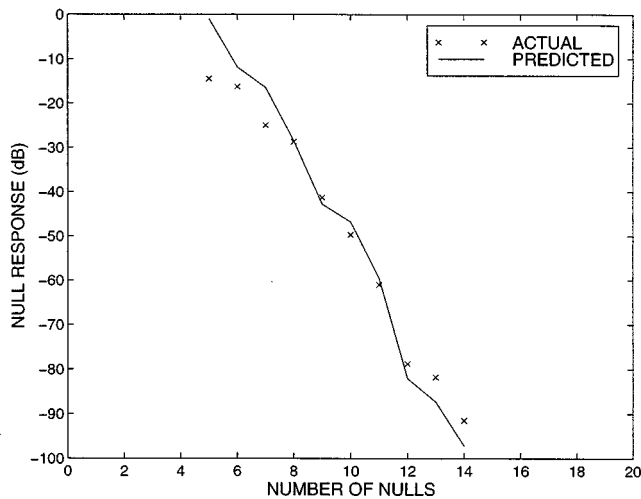


FIG. 9. The predicted null depth given by Eq. (37), and the actual obtained depth, versus number of nulls.

For the exact null constraint, it is possible to produce a pattern zero over all frequencies by formulating the constraint as a time domain constraint. This requires oversampling to produce a null at directions other than broadside or endfire, and places stringent constraints on the sensor locations. The degradation of the exact null that occurs in a practical setting with sensor positioning errors was considered. A reasonably deep null is still achieved for small sensor position perturbations. Sensor calibration errors [i.e., differences between the assumed array vector  $\mathbf{a}(u, \omega)$  and the actual array vector] would also affect the null depth. A similar technique to that presented for sensor position errors could also be used to analyze the effects of sensor calibration errors on the expected null depth.

For the multiple frequency nulling constraint, a relationship was derived between the number of zeros to impose in the quiescent beampattern, and the broadband null depth over the design bandwidth. In the specific case of a frequency invariant beamformer, a simple expression was provided to *a priori* determine the incremental null depth that will be achieved for a given number of frequency zeros.

Examples show the strengths and weaknesses of each technique.

Finally, we conclude by noting that there is no single "right" answer to the broadband nulling problem considered in this paper. We have attempted to give a brief review of some possible formulations of the problem, and have obtained some new results concerning both approximate and exact broadband pattern nulls. These new results are specifically aimed at the frequency invariant beamformer, although they have wider application to more general broadband beamformers.

## ACKNOWLEDGMENTS

The authors wish to acknowledge the funding of the activities of the Cooperative Research Center for Robust and Adaptive Systems by the Australian Commonwealth Government under the Cooperative Research Centers Program. The authors gratefully acknowledge support from the Australian Research Council.

- <sup>1</sup>D. B. Ward, R. A. Kennedy, and R. C. Williamson, "Theory and design of broadband sensor arrays with frequency invariant far-field beam patterns," *J. Acoust. Soc. Am.* **97**, 1023–1034 (1995).
- <sup>2</sup>R. A. Kennedy, T. Abhayapala, D. B. Ward, and R. C. Williamson, "Near field broadband frequency invariant beamforming," in *Proc. International Conference on Acoustics, Speech and Signal Processing*, Atlanta, GA, Vol. II, pp. 905–908 (1996).
- <sup>3</sup>J. L. Flanagan, D. A. Berkley, G. W. Elko, and M. M. Sondhi, "Autodirective microphone systems," *Acustica* **73**, 58–71 (1991).
- <sup>4</sup>H. Steyskal, R. A. Shore, and R. L. Haupt, "Methods for null control and their effects on the radiation pattern," *IEEE Trans. Antennas Propag.* **AP-34**, 404–409 (1986).
- <sup>5</sup>D. K. Cheng, "Optimization techniques for antenna arrays," *Proc. IEEE* **59**, 1664–1674 (1971).
- <sup>6</sup>B. E. Stuckman and J. C. Hill, "Method of null steering in phased array antenna systems," *Electron. Lett.* **26**, 1216–1218 (1990).
- <sup>7</sup>B. P. Ng, M. H. Er, and C. Kot, "Linear array geometry synthesis with minimum sidelobe level and null control," *IEE Proc., Part H: Microwaves, Antennas Propag.* **141**, 162–166 (1994).
- <sup>8</sup>T. H. Ismail and M. M. Dawoud, "Null steering in phased arrays by controlling the element positions," *IEEE Trans. Antennas Propag.* **39**, 1561–1566 (1991).
- <sup>9</sup>A. Tennant, M. M. Dawoud, and A. P. Anderson, "Array pattern nulling by element position perturbations using a genetic algorithm," *Electron. Lett.* **30**, 174–176 (1994).
- <sup>10</sup>H. Steyskal, "Synthesis of antenna patterns with prescribed nulls," *IEEE Trans. Antennas Propag.* **AP-30**, 273–279 (1982).
- <sup>11</sup>H. Steyskal, "Wide-band nulling performance versus number of pattern constraints for an array antenna," *IEEE Trans. Antennas Propag.* **AP-31**, 159–163 (1983).
- <sup>12</sup>R. J. Mailloux, "Covariance matrix augmentation to produce adaptive array pattern troughs," *Electron. Lett.* **31**, 771–772 (1995).
- <sup>13</sup>K. Takao and K. Komiyama, "An adaptive antenna for rejection of wide band interference," *IEEE Trans. Aerosp. Electron. Syst.* **AES-16**, 452–459 (1980).
- <sup>14</sup>C. C. Ko, "Broadband power inversion array with maximally flat response at null directions," *IEE Proc. F, Commun. Radar Signal Process.* **136**, 161–167 (1989).
- <sup>15</sup>M. H. Er, "Technique for antenna array pattern synthesis with controlled broad nulls," *IEE Proc., Part H: Microwaves, Opt. Antennas* **135**, 375–380 (1988).
- <sup>16</sup>T. P. Guella and R. M. Davis, "Synthesis of notched antenna patterns for wide band processing," *IEEE Trans. Antennas Propag.* **43**, 1465–1471 (1995).
- <sup>17</sup>T. W. Parks and J. H. McClellan, "Chebyshev approximation for nonrecursive digital filters with linear phase," *IEEE Trans. Circuit Theory CT-19*, 189–194 (1972).
- <sup>18</sup>D. B. Ward, "Theory and application of broadband frequency invariant beamforming," Ph.D. Thesis, Australian National University, Canberra, Australia (1996).
- <sup>19</sup>O. L. Frost, "An algorithm for linearly constrained adaptive array processing," *Proc. IEEE* **60**, 926–935 (1972).
- <sup>20</sup>J. H. Doles III and F. D. Benedict, "Broadband array design using the asymptotic theory of unequally spaced arrays," *IEEE Trans. Antennas Propag.* **36**, 27–33 (1988).
- <sup>21</sup>R. G. Pridham and R. A. Mucci, "A novel approach to digital beamforming," *J. Acoust. Soc. Am.* **63**, 425–434 (1978).
- <sup>22</sup>T. I. Laakso, V. Välimäki, M. Karjalainen, and U. K. Laine, "Splitting the unit delay," *IEEE Signal Process. Mag.* **13**, 30–60 (1996).
- <sup>23</sup>P. J. Kootsookos and R. C. Williamson, "FIR approximation of fractional sample delay systems," *IEEE Trans. Circuits Syst., II: Analog Digital Signal Process.* **43**, 269–271 (1996).
- <sup>24</sup>M. I. Skolnik, "Nonuniform arrays," in *Antenna Theory, Pt. 1*, edited by R. E. Collin and F. J. Zucker (McGraw-Hill, New York, 1969), Chap. 6, pp. 207–279.
- <sup>25</sup>K. L. Chung, *A Course in Probability Theory* (Harcourt, Brace, and World, New York, 1968).
- <sup>26</sup>D. B. Ward, R. A. Kennedy, and R. C. Williamson, "FIR filter design for frequency-invariant beamformers," *IEEE Signal Process. Lett.* **3**, 69–71 (1996).

# Adaptation of $2f_1-2f_2$ distortion product otoacoustic emission in young-adult and old CBA and C57 mice

X.-M. Sun and D. O. Kim<sup>a)</sup>

Division of Otolaryngology, Department of Surgery, Neuroscience Program, University of Connecticut Health Center, Farmington, Connecticut 06030 and Department of Communication Sciences, University of Connecticut, Storrs, Connecticut 06269

(Received 20 October 1998; accepted for publication 19 March 1999)

The phenomenon of efferent-mediated adaptation of  $2f_1-f_2$  distortion product otoacoustic emission (DPOAE) was investigated in two strains (CBA/JNia and C57BL/6JNia) of mice of various ages using stimuli presented monaurally or binaurally. The present study demonstrated the existence of the DPOAE adaptation phenomenon in mice analogous to that previously reported in cats. The present data were fitted with one- or two-exponential functions. With a one-exponential fit in 2-month old mice, the adaptation magnitude ranged from 0 to 4 dB with the average value of 0.5 to 1.6 dB and the average time constant was 0.5 to 2.3 s. With a two-exponential fit, the shorter time constant was 0.3 to 1.7 s. The adaptation magnitude and time constant were similar between the monaural and binaural stimulations. We observed that there was a statistically significant decrease of adaptation magnitude in older CBA mice with age-related hearing loss when compared with young adult mice. The results from the young adult mice should be useful in future studies, e.g., a study of developmental changes in post-natal mice, or changes accompanying an alteration in the central auditory system arising from any etiology. We suggest that this phenomenon can be used as a tool for advancing basic knowledge of the auditory system and for assessing an impairment of the olivocochlear system, e.g., in aging. © 1999 Acoustical Society of America. [S0001-4966(99)05006-7]

PACS numbers: 43.64.Ha, 43.64.Jb, 43.64.Kc, 43.64.Qh [BLM]

## INTRODUCTION

Medial olivocochlear (MOC) neurons in the brain stem make descending projections onto outer hair cells (OHCs) of the cochlea, forming a feedback or reflex system (e.g., Warr and Guinan, 1979). The MOC efferent system provides negative feedback, i.e., suppression of cochlear responses (e.g., Galambos, 1956; Desmedt, 1962). The MOC system is believed to influence OHCs not only electrophysiologically but also mechanically because activation of the efferent system was found to produce changes in distortion product otoacoustic emissions (DPOAEs), a mechanical signal (Mountain, 1980; Siegel and Kim, 1982). This notion is consistent with the subsequent finding that OHCs are electromotile (Brownell *et al.*, 1985; Zenner *et al.*, 1985) and thus potentially able to influence mechanics of the cochlea, as was postulated theoretically in “cochlear amplifier” models (e.g., Davis, 1983; Neely and Kim, 1983).

The effects of the MOC acoustic reflex have been demonstrated in a number of ways. Many studies assessed the effects of the MOC system by applying a contralateral sound while recording various cochlear responses, e.g., responses of single auditory-nerve fibers (e.g., Buño, 1978; Warren and Liberman, 1989), auditory evoked potentials (e.g., Folsom and Owsley, 1987), or click-evoked otoacoustic emissions (e.g., Collet *et al.*, 1990). Effects of the MOC system on DPOAEs have been investigated in two ways: (1) by electri-

cally activating the efferent system in animals (Mountain, 1980; Siegel and Kim, 1982); and (2) by applying contralateral acoustical stimulation in animals (Puel and Rebillard, 1990; Kujawa *et al.*, 1993; Puria *et al.*, 1996) and in humans (e.g., Moulin *et al.*, 1993; Williams and Brown, 1995). The contralateral influences were eliminated when the olivocochlear bundle was transected (Warren and Liberman, 1989; Puel and Rebillard, 1990; Liberman, 1990).

Contralateral suppression of DPOAEs is expected to be mediated by contralaterally responsive MOC neurons whereas an ipsilaterally evoked effect of the MOC system is expected to be mediated by ipsilaterally responsive MOC neurons (Warren and Liberman, 1989). Measuring ipsilaterally evoked effects of the MOC system would be inherently more difficult than measuring contralaterally evoked effects. For example, if one were to consider adding an ipsilateral modifier (“suppressor”) signal to a test signal, both signals would unavoidably interact peripherally in the same ear besides interacting centrally via olivocochlear neurons.

A method recently introduced by Liberman *et al.* (1996) provides a good way to observe an ipsilateral effect of the MOC system on DPOAEs. Liberman *et al.* observed a phenomenon in cats where the  $2f_1-2f_2$  DPOAE level changed gradually over a few seconds after the onset of the primary stimulus tones, which we refer to as DPOAE adaptation. This phenomenon was interpreted to be mainly mediated by the MOC system (further discussions about this are given in Sec. III). This method of examining the effects of the MOC system has the following advantages: (1) the phenomenon can be examined monaurally or binaurally, whereby ipsilat-

<sup>a)</sup>Corresponding author. Department of Surgery, University of Connecticut Health Center, Farmington, CT 06030-1110; Electronic-mail: kim@neuron.uchc.edu

eral, contralateral or binaural effects of the MOC system can be observed; and (2) the magnitude of the ipsilateral effect tends to be larger than that of the contralateral effect.

The main goal of the present study is to characterize the  $2f_1-2f_2$  DPOAE adaptation phenomenon in mice. Mice are an attractive species for auditory research for a number of reasons. For example, there are many naturally occurring (e.g., Bonaiuto *et al.*, 1996; Zheng *et al.*, 1999) and artificially created (e.g., Vetter *et al.*, 1998) mutant strains of mice which exhibit normal function or various functional abnormalities of the auditory system. This offers an opportunity to pursue multidisciplinary research of the auditory system, using mice, under the influence of various genetic changes. Mice are also desirable as an animal model for studies of age-related hearing loss (e.g., Willott, 1991). In the present study, we measured the  $2f_1-f_2$  DPOAE adaptation in two strains of mice, CBA/JNia and C57BL/6JNia, at various ages. CBA and C57 are two strains of genetically inbred mice commonly used as models of early- and late-onset age-related hearing loss, respectively (e.g., Henry, 1983; Parham, 1997; Sun *et al.*, 1998). The second goal of the present study is to examine the DPOAE adaptation in old mice as a means of discerning an alteration in the olivocochlear system as a function of age. Early results of the present study were reported by Kim and Sun (1998).

## I. METHODS

### A. Subjects

Two strains of mice were purchased from National Institute on Aging/Charles River Laboratories (Kingston, NY) at several age groups: (1) CBA/JNia mice, 2, 17 and 22 months of age; (2) C57BL/6JNia mice, 2, 10 and 12 months of age. Each group consisted of five mice, and both ears of each mouse were tested. The range of ages in each strain was selected to include 2 months as a young-adult age (Crispens, 1976) and to cover a large portion of age span where age-related hearing changes were observed (e.g., Henry and Chole, 1980; Willott, 1986) including changes of DPOAE level (Parham, 1997; Sun *et al.*, 1998). Only data having DPOAE levels of at least 5 dB above the noise floor were included for the adaptation analysis. Because of imposing the signal/noise ratio criterion and because of death or middle-ear disorder during the study, the final number of ears represented tended to decrease for older mice as indicated in the figures presented in Sec. II. Regarding 12-month C57 mice, we were not able to discern a clear pattern of response in those ears because we were able to obtain data with a sufficiently high signal/noise ratio from only a small number of ears. Consequently, results from 12-month C57 are not presented. The care and use of the animal subjects reported in this study was approved by the Institutional Animal Care and Use Committee of the University of Connecticut Health Center (No. 92-032-96).

### B. Recording procedure

Prior to the recordings, the ear canals and ear drums were microscopically inspected for signs of cerumen obstruction or infection, and only those ears with clear ear ca-

nal and tympanic membrane were used. For the recording sessions, the mice were anesthetized by intraperitoneal injection of a mixture of ketamine/xylazine (0.12 and 0.01 mg/g body weight, respectively) with supplementary doses (20% of the initial dose) every 40 to 60 min as necessary. All recordings took place in an acoustic chamber (Model 1204, Industrial Acoustics Co., Bronx, NY). The body temperature of the animals was maintained near 37 °C with a regulated heating pad and monitored with a rectal thermoprobe.

### C. Recording systems and protocol

The present DPOAE measurement system was described in detail in a previous report (Kim *et al.*, 1996). Briefly, the system consists of an Ariel DSP-16 board in a personal computer, Etymotic Research ER-2 earphones, and an ER-10B microphone system with a built-in amplifier providing a gain of +40 dB. CUB<sup>®</sup>DISC software (Mimosa Acoustics, Mountainside, NJ) was employed.

The paradigm for generation of the stimuli and measurement of the  $2f_1-f_2$  DPOAE was similar to that used by Liberman *et al.* (1996). In the present study, the sampling rate was 50 kHz, and the primary tones were presented as a train of brief bursts. Each stimulus burst had a duration of 143.5 ms corresponding to seven 20.5-ms periods including a rise time of 20.5 ms. The interburst silent period was 8.1 ms. The CUB<sup>®</sup>DISC DPOAE system required a minimum silent period of 8.1 ms between bursts. Each data point of the DPOAE level, measured as a function of time, represented the last 102.5-ms portion of each burst. A ‘‘block’’ of data, displaying the time course of the DPOAE adaptation, consisted of 40 consecutive data points encompassing a total elapsed time of about 6 s following the onset of a block of the stimulus bursts. To reduce the noise level, averages were taken over five blocks of data which were separated by an interblock silent gap of 30 s. The resulting data were used to display the time course of DPOAE adaptation.

The responses were measured monaurally and binaurally. In the binaural measurements, identical stimuli were simultaneously presented into the two ears while recording in one ear. Data were collected at  $f_2=8, 11.3$  and 16 kHz, with  $f_2/f_1=1.2$ . The primary tone levels,  $L_1/L_2$ , were 65/50 dB SPL (*re*: 20  $\mu$ Pa) for  $f_2=8$  and 11.3 kHz and  $L_1/L_2=50/35$  dB SPL for  $f_2=16$  kHz. The use of a lower stimulus level for 16 kHz was due to an equipment limitation. For  $f_2=11.3$  kHz, measurements were also made with  $L_1/L_2=65/65$  dB SPL to observe the effect of  $L_1-L_2$  on DPOAE adaptation. The test time for each ear was about 1 h.

### D. Data analysis and statistical tests

To quantify the magnitude and time constant of the DPOAE adaptation, the data of DPOAE level (in dB relative to the first data point) versus time were fit with a one-exponential function,  $y(t)=\alpha(1-e^{-(t/\tau)})$ , where  $y(t)$ ,  $\alpha$ , and  $\tau$  represent the DPOAE level at time  $t$ , adaptation magnitude (in dB) and time constant (in s), respectively. To examine the rapid and slow components of the DPOAE adaptation observed by Liberman *et al.* (1996), we also used a two-exponential fitting function,

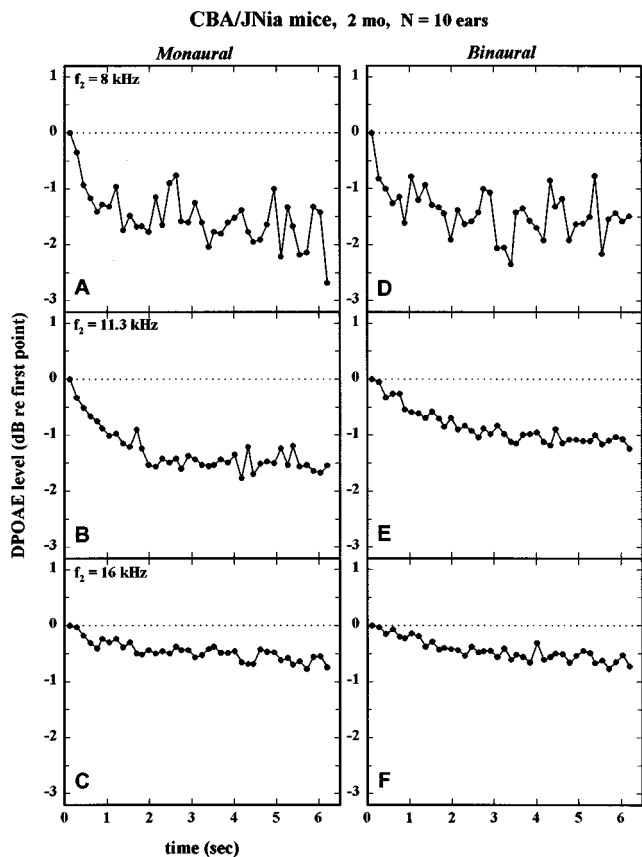


FIG. 1. Time course of  $2f_1-f_2$  DPOAE level relative to the first point following the onset of the primary tones in 2-month old CBA mice. Data represent averages of 10 ears (N). The left and right columns represent monaural and binaural stimulations, respectively. The first point of DPOAE level with monaural stimulation for 8, 11.3 and 16 kHz was 4.6, 13.2 and 18 dB SPL, respectively, and that with binaural stimulation for the same three frequencies was 4.5, 12.7 and 16.9 dB SPL, respectively. The noise level associated with these DPOAE measurements was about  $-10$  dB SPL (not shown in the figure). Figures 3, 4, 7 and 8 are presented using this same format.

$y(t) = \alpha_1(1 - e^{-(t/\tau_1)}) + \alpha_2(1 - e^{-(t/\tau_2)})$ , where  $\alpha$  and  $\tau$  with subscript 1 represent the parameters of the rapid component and those with subscript 2 the slow component. The variability of the DPOAE signal was measured in terms of the standard error of the mean (SEM) of DPOAE level at each time point relative to the first point among a group of ears. The SEM values at various time points were then averaged.

Paired  $t$ -test was performed by using Systat 7.0 software (SPSS, Inc.) to evaluate the statistical significance ( $p < 0.05$ ) of the difference between monaural and binaural stimulations regarding the DPOAE adaptation magnitude and time constant. A one-way analysis of variance (ANOVA) was performed to evaluate differences among the age groups regarding the DPOAE adaptation magnitude and time constant. Significant ( $p < 0.05$ ) main effects were evaluated with the Tukey *post hoc* pairwise comparisons.

## II. RESULTS

### A. DPOAE adaptation in young CBA mice

The mean time courses of the  $2f_1-f_2$  DPOAE level following the onset of primary tones are shown in Fig. 1 for 2-month old CBA mice. The ordinate represents DPOAE

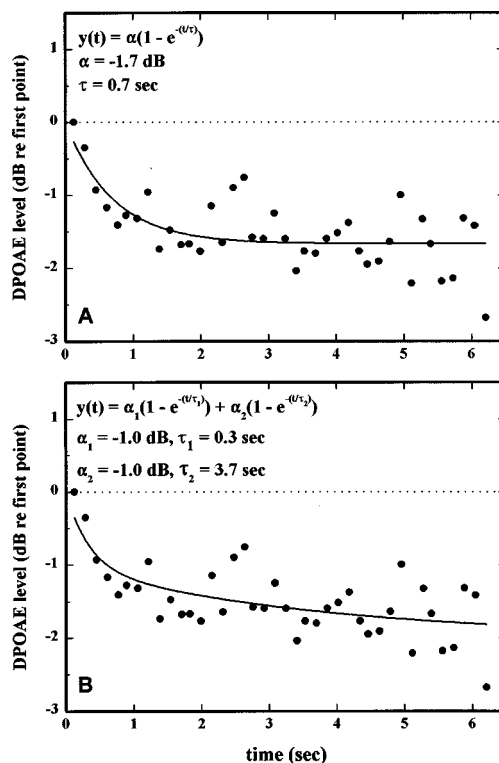


FIG. 2. Examples of fitting functions of DPOAE level versus time; 2-month old CBA mice, average of 10 ears; monaural stimulation,  $f_2 = 8$  kHz. (A) one-exponential function,  $y(t) = \alpha(1 - e^{-(t/\tau)})$ . (B) Two-exponential function,  $y(t) = \alpha_1(1 - e^{-(t/\tau_1)}) + \alpha_2(1 - e^{-(t/\tau_2)})$ . Parameter values of the fitting functions are shown in the figure.

level in dB relative to the first point in time. The absolute ordinate value of the first point in dB SPL for each panel is indicated in the figure caption. It is seen that the DPOAE level gradually declined over time while the stimulus level remained the same, i.e., there was adaptation of DPOAE level. In each of the six panels, a gradual decline of the emission level occurred over a few seconds following the onset of primary tones, eventually reaching a steady-state value.

An example of the DPOAE adaptation time course together with a one-exponential fitting function is shown in Fig. 2(A). The magnitude,  $\alpha$ , and time constant,  $\tau$ , of the DPOAE adaptation for this example were  $-1.7$  dB and  $0.7$  s, respectively. The negative polarity of  $\alpha$  here signifies a “decreasing adaptation.” In the study by Liberman *et al.* (1996), two components, rapid and slow, were described for DPOAE adaptation in cats. They observed that the rapid adaptation component was more dominant and abolished by transection of the olivocochlear bundle, whereas the slow component was of a smaller magnitude and persisted after transection of the olivocochlear bundle. Figure 2(B) illustrates an example of a DPOAE adaptation time course fitted with a two-exponential function. Values of the adaptation magnitude and time constant for this example were  $-1.0$  dB and  $0.3$  s for the rapid component and  $-1.0$  dB and  $3.7$  s for the slow component.

Mean values of the parameters of the one-exponential fit for 2-month old CBA mice are listed in the upper part of Table I. Regarding the results derived from the one-



TABLE I. Average parameter values for a one-exponential fit for the DPOAE adaptation in CBA mice of various ages;  $y(t) = \alpha(1 - e^{-t/\tau})$ .

Age (months)	$f_2$ (kHz)	Monaural		Binaural	
		$\alpha$ (dB)	$\tau$ (s)	$\alpha$ (dB)	$\tau$ (s)
2	8	-1.6	0.5	-1.5	0.6
	11.3	-1.5	1.0	-1.2	1.7
	16	-0.6	1.9	-0.7	2.3
17	8	-0.8	1.4	-1.2	0.5
	11.3	-1.4	1.9	-1.4	2.0
	16	-0.5	1.1	-0.4	1.6
22	8	-0.5	7.1	-1.1	2.1
	11.3	-0.5	2.9	-0.6	1.1
	16	-0.2	2.4	-0.2	0.3

exponential fit, the adaptation magnitude,  $\alpha$ , was negative,  $-0.6$  to  $-1.6$  dB, for all of the six cases (three frequencies, monaural and binaural). The time constant,  $\tau$ , was  $0.5$  to  $2.3$  s. Although a comparison across frequencies in this study was complicated by the fact that a lower stimulus level was used for  $16$  kHz than for the other two lower frequencies, a trend was seen such that the adaptation magnitudes were larger and time constants were shorter at lower frequencies. Differences between the monaural and binaural stimulations were small for both  $\alpha$  and  $\tau$ . We performed a paired  $t$ -test on values of  $\alpha$  and  $\tau$  of individual ears of 2-month CBA mice comparing the monaural and binaural stimulus conditions. The test result indicated that there was no significant difference (i.e.,  $p > 0.05$ ) between the monaural and binaural stimulations regarding either  $\alpha$  or  $\tau$ . Regarding the results derived from the two-exponential fit (Table II), magnitudes of the rapid and slow adaptation,  $\alpha_1$  and  $\alpha_2$ , respectively, were similar to each other. Also, both of them were smaller than the overall adaptation magnitude,  $\alpha$ , derived from the one-exponential fit. The rapid time constant was  $0.3$  to  $1.7$  s and the slow time constant was  $2.2$  to  $3.9$  s.

### B. DPOAE adaptation in old CBA mice

Figures 3 and 4 show the DPOAE adaptation functions, in the same format as that of Fig. 1, for 17- and 22-month old CBA mice, respectively. The adaptation functions of the 17-month CBA mice (Fig. 3) were generally similar to those of the 2-month CBA mice (Fig. 1). A trend was seen such that the 22-month CBA mice (Fig. 4) exhibited smaller adaptation magnitude than the 2-month CBA mice (Fig. 1) although the 22-month mice data were more variable than the younger mice data. The parameters of the one-exponential fit for the two older age groups of CBA mice are listed in the

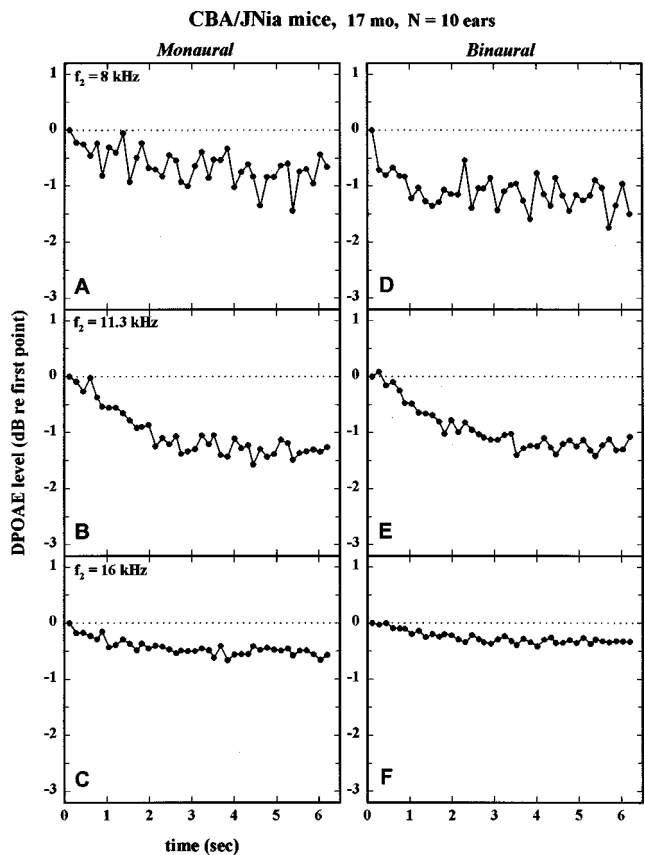


FIG. 3. Data for 17-month old CBA mice. The first point of the DPOAE level with monaural stimulation for 8, 11.3 and 16 kHz was 7.7, 12.2 and 19.6 dB SPL, respectively, and that with binaural stimulation for the same three frequencies was 7.5, 12.2 and 18.9 dB SPL, respectively.

lower part of Table I, where it is seen that the aged mice tended to have a smaller adaptation magnitude (as noted above) and a longer time constant than the younger mice.

The variability of DPOAE level relative to the first point was measured in terms of the SEM for various ages and frequencies. The SEM for CBA mice, listed in Table III, was  $0.1$  to  $0.5$  dB, being larger for lower frequencies and similar between the monaural and binaural stimulations. There was a tendency for the SEM to be larger for the 22-month data, at least partly as a result of a smaller number of ears represented.

A one-way analysis of variance (ANOVA) was performed to determine whether age had statistically significant effects on the DPOAE adaptation. The results are summarized in Table IV. With monaural stimulation, the reduction in the adaptation magnitude with age was statistically significant for the following conditions: at  $8$  kHz for 2 vs 17

TABLE II. Average parameter values for a two-exponential fit for the DPOAE adaptation in 2-month old CBA mice;  $y(t) = \alpha_1(1 - e^{-t/\tau_1}) + \alpha_2(1 - e^{-t/\tau_2})$ .

$f_2$ (kHz)	Monaural				Binaural			
	$\alpha_1$ (dB)	$\tau_1$ (s)	$\alpha_2$ (dB)	$\tau_2$ (s)	$\alpha_1$ (dB)	$\tau_1$ (s)	$\alpha_2$ (dB)	$\tau_2$ (s)
8	-1.0	0.3	-1.0	3.9	-0.9	0.3	-0.8	2.2
11.3	-0.9	0.7	-0.8	2.5	-0.6	1.2	-0.6	2.9
16	-0.3	1.1	-0.3	3.9	-0.3	1.7	-0.3	3.2

CBA/JNia mice, 22 mo

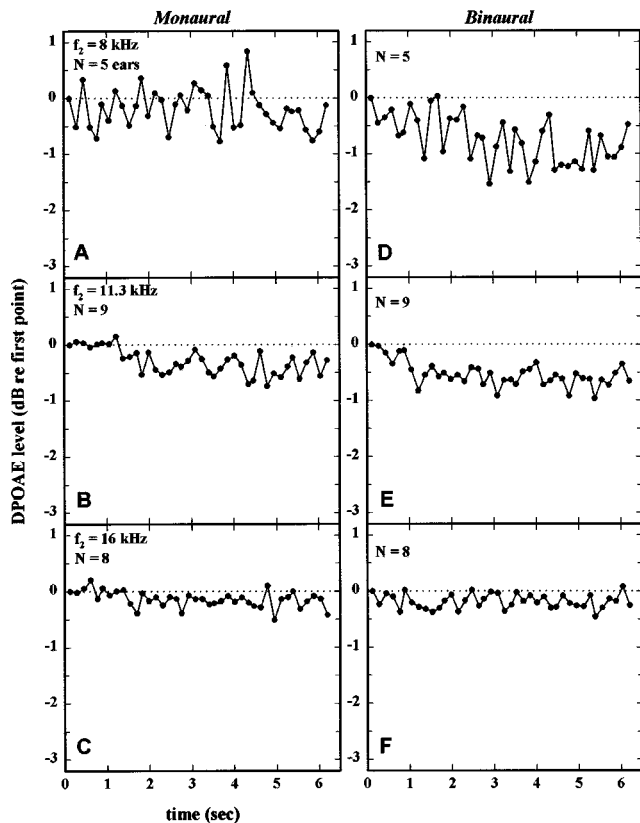


FIG. 4. Data for 22-month old CBA mice. The first point of the DPOAE level with monaural stimulation for 8, 11.3 and 16 kHz was 5.3, 12.3 and 14.3 dB SPL, respectively, and that with binaural stimulation for the same three frequencies was 6.4, 13.5 and 14.0 dB SPL, respectively. The number of ears averaged (N) is shown in each panel in this figure and in Figs. 5, 8 and 9.

months and 2 vs 22 months; at 11.3 kHz for 2 vs 22 months and 17 vs 22 months; at 16 kHz for 2 vs 22 months and 17 vs 22 months. With binaural stimulation, it was significant at 16 kHz for 2 vs 17 months and 2 vs 22 months. The effect of age on the adaptation *time constant* was not statistically significant for any of the age comparisons. A larger number of ears and greater averaging of the DPOAE signal to achieve higher signal/noise ratios may reveal more statistically significant age-dependent changes in the DPOAE adaptation parameters.

**C. Effects of  $L_1-L_2$  and increasing adaptation**

The effect of level difference of primary tones,  $L_1-L_2$ , on the DPOAE adaptation process was examined by comparing results with  $L_1-L_2=15$  dB and 0 dB with  $L_1$

TABLE III. Standard error of the mean (in dB) of DPOAE level relative to the first time point for CBA mice. Various ages, frequencies, monaural and binaural stimulations are represented. The number in parentheses, in each case, corresponds to the number of ears represented.

Age (months)	$f_2$ (kHz)	Monaural	Binaural
2	8	0.43(10)	0.44(10)
	11.3	0.22(10)	0.12(10)
	16	0.12(10)	0.09(10)
17	8	0.28(10)	0.30(10)
	11.3	0.34(10)	0.27(10)
	16	0.09(10)	0.06(10)
22	8	0.35 (5)	0.49 (5)
	11.3	0.20 (9)	0.18 (9)
	16	0.17 (8)	0.14 (8)

=65 dB SPL at  $f_2=11.3$  kHz (Fig. 5). The adaptation magnitude was larger with  $L_1-L_2=15$  dB (i.e.,  $L_1/L_2=65/50$  dB SPL) than with  $L_1-L_2=0$  dB (i.e.,  $L_1/L_2=65/65$  dB SPL) for most cases. For example, the emission level changed little over time with  $L_1/L_2=65/65$  binaural stimulation [Fig. 5(D)]. An interesting phenomenon observed in this study is “increasing adaptation” where the DPOAE level *increased* over time, e.g., Fig. 5(E) with  $L_1/L_2=65/65$ , binaural, age =17 months, opposite to the usual “decreasing adaptation.” The increasing type of adaptation was not restricted to 17 months of age because, even among 2-month CBA mice, some ears exhibited the increasing type of adaptation with  $L_1/L_2=65/65$  while exhibiting the decreasing type of adaptation with  $L_1/L_2=65/50$ , both under binaural stimulation (Fig. 6).

**D. DPOAE adaptation in young C57 mice**

The DPOAE adaptation functions for 2-month old C57 mice are shown in Fig. 7. The general appearance of the adaptation functions of the C57 mice is similar to that of the CBA mice (Fig. 1). Mean values of the parameters of the one-exponential fit for 2-month old C57 mice are listed in the upper part of Table V. Values of the adaptation magnitude,  $\alpha$ , were all negative,  $-0.5$  to  $-0.9$  dB, for the six cases (three frequencies, monaural and binaural). The time constant,  $\tau$ , was 0.6 to 1.6 s. There was a trend such that the adaptation time constant was shorter at lower frequencies. Paired *t*-test on  $\alpha$  and  $\tau$  of individual ears of 2-month C57 mice indicated that there was no significant difference ( $p > 0.05$ ) between the monaural and binaural stimulations regarding either  $\alpha$  or  $\tau$ . Results from 2-month C57 mice summarized in Table V (upper part) are generally similar to those

TABLE IV. Results of a statistical test, one-way ANOVA, of the effect of age on the DPOAE adaptation magnitude with monaural and binaural stimulations in CBA mice. The “-” symbol represents that  $p > 0.05$ .

$f_2$ (kHz)	Monaural			Binaural		
	2 vs 17 months	2 vs 22 months	17 vs 22 months	2 vs 17 months	2 vs 22 months	17 vs 22 months
8	$p < 0.05$	$p < 0.05$	-	-	-	-
11.3	-	$p < 0.05$	$p < 0.05$	-	-	-
16	-	$p < 0.01$	$p < 0.05$	$p < 0.01$	$p < 0.01$	-

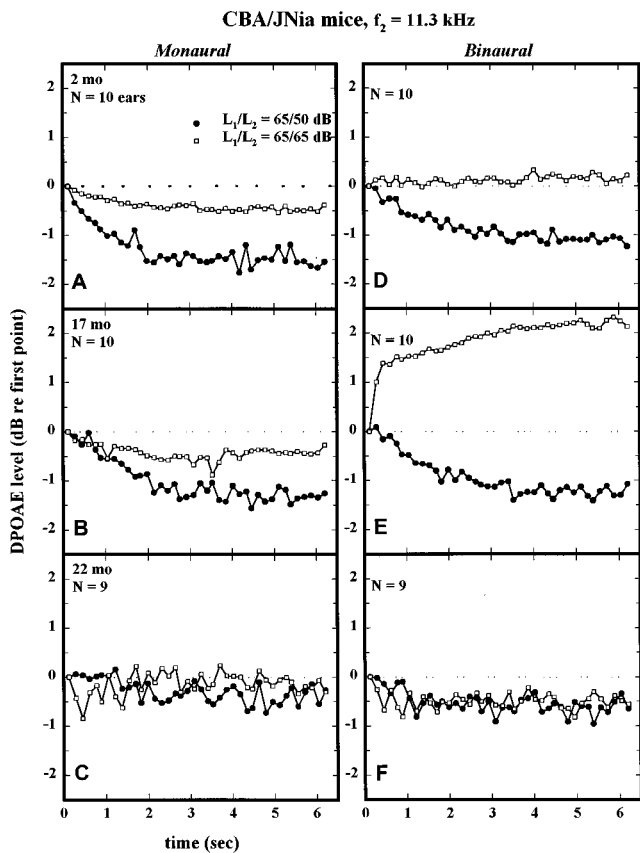


FIG. 5. Comparison of time courses of  $2f_1 - f_2$  DPOAE level of CBA mice with  $L_1/L_2 = 65/50$  dB (closed circles) and  $65/65$  dB SPL (open squares). Ages: 2, 17 and 22 months for the top, middle and bottom rows, respectively. The first point of the DPOAE level in dB SPL for  $65/50$  and  $65/65$  dB SPL stimuli was 13.2 and 21.1 (panel A), 12.2 and 24.4 (panel B), 12.3 and 15.0 (panel C), 12.7 and 17.2 (panel D), 12.2 and 20.0 (panel E), 13.5 and 17.2 (panel F), respectively.

from 2-month CBA mice summarized in Table I except that the adaptation magnitude,  $\alpha$ , tended to be smaller for C57 than for CBA by 0.5 dB on the average; average of  $\alpha$  was  $-1.2$  dB and  $-0.7$  dB for CBA and C57, respectively.

Regarding the results derived from the two-exponential fit (Table VI), magnitudes of the rapid and slow adaptation,  $\alpha_1$  and  $\alpha_2$  respectively, were similar to each other and both

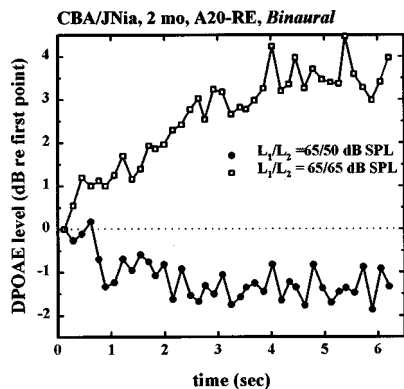


FIG. 6. Examples of time courses of the DPOAE level in an individual ear of a 2-month old CBA mouse which exhibited increasing adaptation with  $L_1/L_2 = 65/50$  dB SPL and declining adaptation with  $65/65$  dB SPL;  $f_2 = 11.3$  kHz. The first point of the DPOAE level was 11.7 and 12.5 dB SPL for the  $65/50$  and  $65/65$  conditions, respectively.

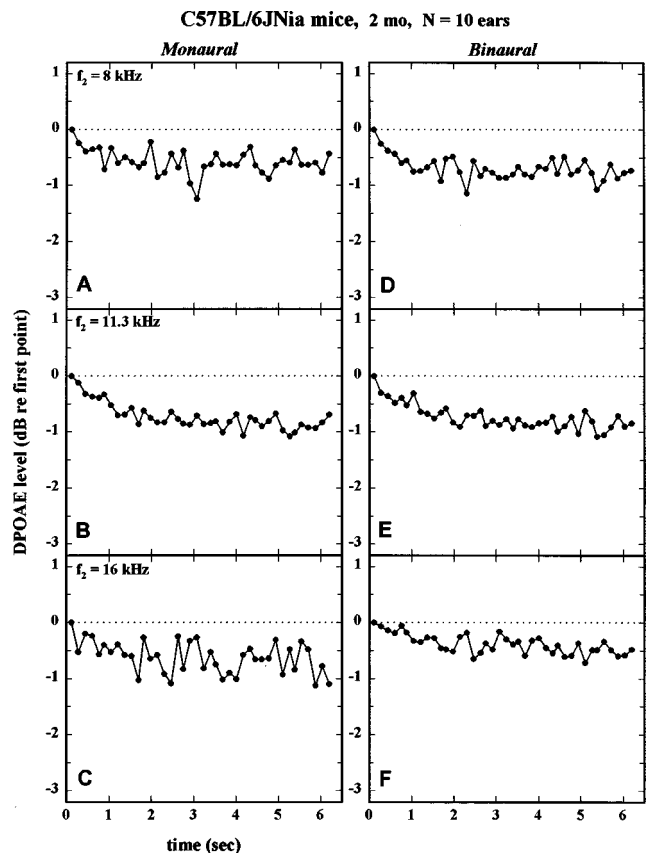


FIG. 7. Data for 2-month old C57 mice. The first point of the DPOAE level with monaural stimulation for 8, 11.3 and 16 kHz was 7.6, 12.8 and 11.8 dB SPL, respectively, and that with binaural stimulation for the same three frequencies was 8, 13.1 and 14.3 dB SPL, respectively.

of them were smaller than the overall adaptation magnitude,  $\alpha$ , derived from the one-exponential fit. The rapid time constant was 0.3 to 1.0 s and the slow time constant was 1.8 to 3.5 s.

### E. DPOAE adaptation in old C57 mice

Figure 8 shows the DPOAE adaptation functions, in the same format as that of Fig. 1, for 10-month old C57 mice. The parameters of the one-exponential fit for 10-month C57 mice are listed in the lower part of Table V. In Figs. 7 and 8 and Table V, one can note that 10-month C57 mice tended to have a slightly smaller adaptation magnitude,  $\alpha$ , than

TABLE V. Average parameter values for a one-exponential fit for the DPOAE adaptation in C57 mice of various ages;  $y(t) = \alpha(1 - e^{-(t/\tau)})$ .

Age (months)	$f_2$ (kHz)	Monaural		Binaural	
		$\alpha$ (dB)	$\tau$ (s)	$\alpha$ (dB)	$\tau$ (s)
2	8	-0.6	0.6	-0.8	0.6
	11.3	-0.9	1.2	-0.9	1.0
	16	-0.5	1.5	-0.5	1.6
10	8	-0.5	0.5	-0.6	0.4
	11.3	-0.6	4.0	-0.3	0.3
	16	-0.3	6.3	-0.4	0.2

TABLE VI. Average parameter values for a two-exponential fit for the DPOAE adaptation in 2-month old C57 mice;  $y(t) = \alpha_1(1 - e^{-t/\tau_1}) + \alpha_2(1 - e^{-t/\tau_2})$ .

$f_2$ (kHz)	Monaural				Binaural			
	$\alpha_1$ (dB)	$\tau_1$ (s)	$\alpha_2$ (dB)	$\tau_2$ (s)	$\alpha_1$ (dB)	$\tau_1$ (s)	$\alpha_2$ (dB)	$\tau_2$ (s)
8	-0.4	0.4	-0.3	1.8	-0.5	0.3	-0.5	2.3
11.3	-0.5	0.8	-0.5	2.4	-0.5	0.5	-0.5	2.6
16	-0.3	1.0	-0.3	3.1	-0.3	0.9	-0.3	3.5

2-month C57 mice. This is analogous to the age-related change in  $\alpha$  noted above in CBA mice (Table I). A pattern of changes in  $\tau$  with age was not clear.

The DPOAE adaptation functions obtained with  $L_1/L_2 = 65/50$  and 65/65 dB SPL at  $f_2 = 11.3$  kHz are shown in Fig. 9. The stimuli with 65/50 dB SPL yielded larger adaptation magnitudes than those with 65/65 dB SPL in both 2-month and 10-month C57 mice. In contrast to the results from CBA mice (Fig. 5), *increasing* adaptation functions were not observed in C57 mice with either of the two types of stimuli at any age.

The variability of DPOAE level relative to the first point in C57 mice is listed in Table VII in terms of the SEM. The SEM was 0.1 to 0.5 dB except for one condition, being larger for lower frequencies and generally similar between the monaural and binaural stimulations. There was a tendency for the SEM to be larger for the 10-month data, at

least partly as a result of a smaller number of ears represented. The reason for one exceptionally large SEM, 0.9 dB for the 10-month, 8 kHz monaural condition, is not clear.

We performed a one-way ANOVA test to determine whether there were significant effects of age on the DPOAE adaptation in C57 mice. The test results indicated that the age-related changes were not statistically significant regarding either the magnitude or time constant of the DPOAE adaptation in C57 mice. As noted above, greater sample size and signal averaging may reveal statistically significant effects of age on the DPOAE adaptation in C57 mice.

### III. DISCUSSION

#### A. Main findings

The main findings of the present study are that a DPOAE adaptation phenomenon exists in mice and that the DPOAE adaptation changes with age. The pattern of the post-onset time course of DPOAE level in mice is qualitatively similar to that in cats (Liberman *et al.*, 1996). Even though the magnitude of the DPOAE adaptation in mice was found to be relatively small, ranging from near 0 to about 4 dB in young adults, the temporal pattern of change in the

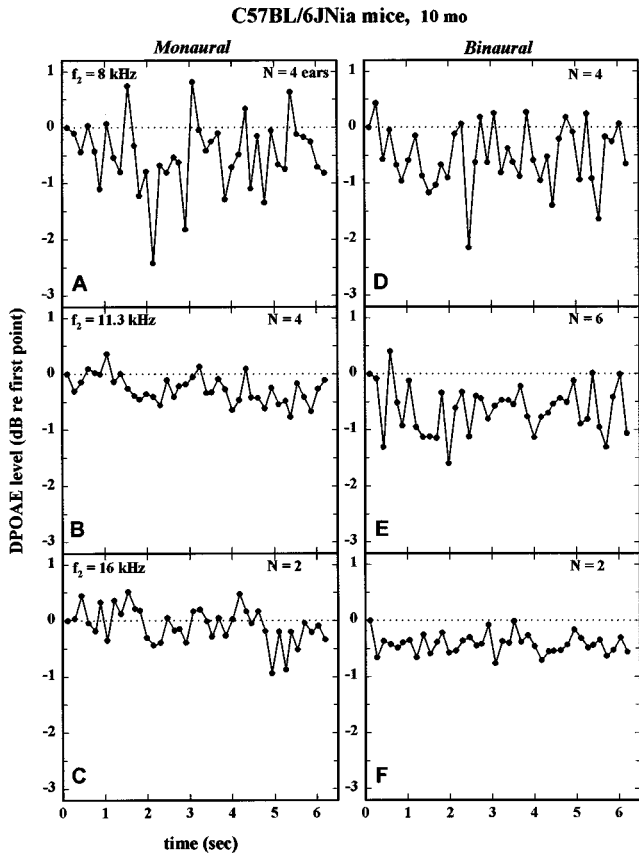


FIG. 8. Data for 10-month old C57 mice. The first point of the DPOAE level with monaural stimulation for 8, 11.3 and 16 kHz was 3.9, 10.6 and 11.1 dB SPL, respectively, and that with binaural stimulation for the same three frequencies was 4.8, 8.4 and 12.8 dB SPL, respectively.

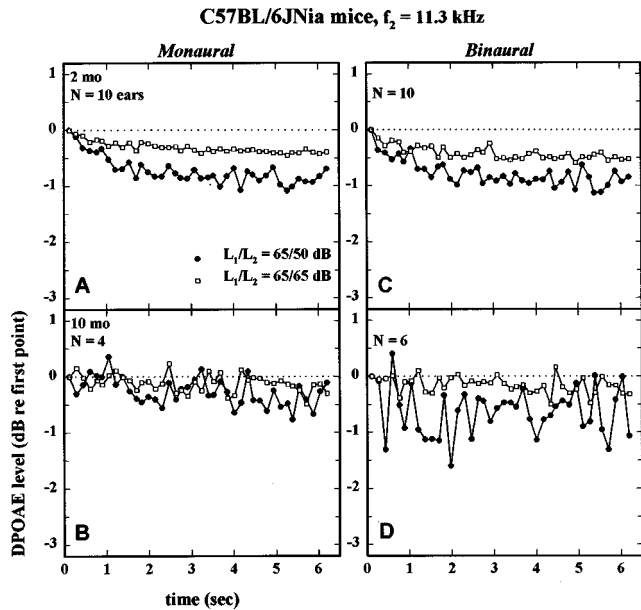


FIG. 9. Comparison of time courses of  $2f_1 - f_2$  DPOAE level of C57 mice with  $L_1/L_2 = 65/50$  dB (closed circles) and 65/65 dB SPL (open squares). Ages: 2 and 10 months for the top and bottom rows, respectively. The first points of DPOAE level in dB SPL for 65/50 and 65/65 dB SPL stimuli were 12.3 and 22.5 (panel A), 10.6 and 17.3 (panel B), 13.1 and 22.2 (panel C), 8.4 and 15.4 (panel D), respectively.

TABLE VII. Standard error of the mean (in dB) of DPOAE level relative to the first time point for C57 mice. Various ages, frequencies, monaural and binaural stimulations are represented. The number in parentheses, in each case, corresponds to the number of ears represented.

Age (months)	$f_2$ (kHz)	Monaural	Binaural
2	8	0.22(10)	0.22(10)
	11.3	0.16(10)	0.16(10)
	16	0.14(10)	0.15(10)
10	8	0.94 (4)	0.55 (4)
	11.3	0.21 (4)	0.48 (6)
	16	0.23 (2)	0.18 (2)

emission level as a function of time was clearly recognizable (e.g., Fig. 1) because the random variability was much smaller (the SEM being often a small fraction of 1 dB, Tables III and VII) than the adaptation magnitude.

### B. Middle-ear muscle reflex

Adaptation of DPOAE may be possibly mediated by a middle-ear muscle reflex. This is because middle-ear muscles can be activated by ipsilateral and/or contralateral sound stimulation, which would attenuate both a forward transmission of the primary tones and a reverse transmission of the emission. Although the present study did not include conditions where the middle-ear muscles were severed, the study by Liberman *et al.* (1996) provides strong evidence against middle-ear muscle reflex as a major source for DPOAE adaptation by demonstrating the following in the cat: (1) DPOAE adaptation in animals with the middle-ear muscles completely severed was not substantially different from that in animals with intact middle-ear muscles (e.g., Fig. 9 of Liberman *et al.*, 1996); and (2) transection of the olivocochlear bundle mostly abolished the DPOAE adaptation. Furthermore, moderate level stimuli ( $L_1/L_2=65/50$  and  $65/65$  dB SPL) used in this study are not expected to produce a middle-ear muscle reflex sufficient to produce the observed DPOAE adaptation.

### C. Olivocochlear system

The observation that the DPOAE adaptation was mostly abolished by olivocochlear transection (cited above) supports a view that the DPOAE adaptation is mainly mediated by the olivocochlear system. Liberman *et al.* also observed that a minor component of the DPOAE adaptation, having a slower time constant, remained after olivocochlear transection. This observation suggests that the minor component of the DPOAE adaptation is mediated by a local cochlear process, possibly a process of distortion-product generation associated with OHC motility. Because the present study did not include transection of the olivocochlear bundle in mice, we are unable to ascertain directly how much of the DPOAE adaptation in mice is mediated by the olivocochlear system and how much by a local cochlear process. We interpret, however, that the DPOAE adaptation of the mouse described in this study is mediated, to a large extent, by the olivocochlear system because the DPOAE adaptation behaviors of the mouse and cat are qualitatively similar. In addition, the

cat study showed that the second component of DPOAE adaptation (remaining after olivocochlear transection) was small at 8–16 kHz, the frequency range which was used in the present mouse study. Although accurately converting frequencies between the cat and mouse is difficult, the finding in the cat is consistent with the present interpretation.

### D. MOC system

The olivocochlear system consists of the medial and lateral olivocochlear (MOC and LOC) neurons projecting to the outer hair cells (OHCs) and to afferent fibers beneath the inner hair cells (IHCs), respectively (e.g., Warr, 1992; Guinan, 1996). Between the two olivocochlear subsystems, the MOC, rather than the LOC, subsystem is believed to underlie the DPOAE adaptation for the following reasons. First, Liberman *et al.* (1996) found that sectioning only the crossed olivocochlear bundle (COCB) at the midline was as effective in disrupting the ipsilaterally evoked DPOAE adaptation as sectioning the complete olivocochlear bundle. This finding suggests that ipsilaterally responsive MOC neurons, most of which are expected to travel in the COCB, mediate the adaptation and that LOC neurons that are spared by the midline section do not contribute to the adaptation. In addition, the present view is consistent with the observations of OHC electromotility and influence on DPOAEs by electrical activation of the COCB cited in the Introduction.

### E. Prediction about basilar-membrane motion

According to a combined hypothesis about OHCs and MOC neurons (e.g., Kim, 1984; Kim *et al.*, 1995, 1998), one postulates that, when the ear is subjected to the onset of a steady-state stimulus lasting for several seconds, the cochlear amplifier gain is controlled by reflex activities of MOC neurons whereby a cochlear mechanical response undergoes a gradual change reflecting MOC neurons' activities. The existence of the DPOAE adaptation phenomenon provides strong support for this hypothesis. There is an important prediction of the hypothesis, i.e., the amplitude of basilar-membrane motion (e.g., in response to single or two tones) should undergo an adaptation over the time course of a few seconds analogous to the DPOAE adaptation. To our knowledge, there are presently no reports of such adaptation in basilar-membrane motion. A reliable measurement should reveal the existence of such an adaptation in basilar-membrane motion. This is because, according to current understanding, the presence of the DPOAE adaptation phenomenon in the ear canal is incompatible with the absence of a parallel adaptation phenomenon in basilar-membrane motion.

### F. Effects of frequency and monaural/binaural stimulation

The magnitudes of the DPOAE adaptation in the mice tended to be larger at lower frequencies, a phenomenon also observed in cats (Liberman *et al.*, 1996). The adaptation magnitude with binaural stimulation is generally expected to be larger when compared with ipsilateral stimulation since more MOC neurons are expected to be activated when

sounds are presented to both ears. In cats, the DPOAE adaptation magnitude was larger with binaural stimulation than with monaural stimulation. In the present study of mice, however, no significant difference of adaptation magnitude was found between monaural and binaural stimulations. The reason for this difference is not clear. The present finding of similar magnitudes of the DPOAE adaptation in the mouse between monaural and binaural stimulations implies that contralaterally responsive MOC neurons (which may be mostly uncrossed) are relatively fewer in the mouse than in the cat. Neuroanatomical studies (e.g., Warren and Liberman, 1989; Campbell and Henson, 1988; review by Warr, 1992) indicate that uncrossed (i.e., ipsilateral) MOC neurons constitute 25% and 30% of all MOC neurons in the mouse and cat, respectively. A fewer percentage of uncrossed MOC neurons in the mouse than in the cat is consistent with the above implication but it is not clear whether the slight difference in these percentages is sufficient to account for the difference in the DPOAE adaptation between the mouse and cat regarding the contralateral influence.

### G. Comparison of mouse, cat and other species

The average adaptation magnitude (obtained with a one-exponential fit) for the mouse studied here ranged from near 0 to about 4 dB in young adults. These values tend to be smaller than those in the adult cat, ranging from near 0 to about 9 dB (Liberman *et al.*, 1996, Fig. 11). This may partly represent a species difference. The difference, however, may be explained, at least in part, by differences in animal preparation methods. For example, the bulla was opened and the pinna and ear canal were removed in the cat study whereas the outer and middle ears were kept intact in the present mouse study. When viewed from a source located inside the ear, placement of an acoustic probe close to the tympanic membrane after removing the ear canal is expected to make the load impedance higher. Since the signal source amplitude is divided between the source impedance and the load impedance, a higher load impedance would lead to a larger amplitude of the recorded emission signal.

In the cat, the rapid component of the DPOAE adaptation was much larger than the slow component (cited above). In the mouse, the magnitudes of the rapid and slow components were similar (Tables II and VI). The reason for this difference is not clear. Regarding time constants of the adaptation, we note the following from Tables II and VI of this paper and Figs. 11 and 12 of Liberman *et al.* (1996): (1) the rapid time constant was 0.3 to 1.7 s in the mouse and 0.05 to 0.5 s in the cat; and (2) the slow time constant was 1.8 to 3.9 s in the mouse and 0.4 to 2.5 s in the cat. Thus the time constants in the two species were in overlapping ranges with those of the mouse being longer. The reason for the difference in the time constants is not clear. The time constants were similar between the monaural and binaural stimulations in the mouse (Tables II and V), which is in agreement with the results from the cat (Liberman *et al.*, 1996, Fig. 11).

Besides the cat and mouse, ipsilaterally produced DPOAE adaptation was also observed in the guinea pig (Kujawa and Liberman, 1998), where the adaptation behavior was qualitatively similar to those of the other two species.

The magnitude of the DPOAE adaptation observed in the present study is comparable to that of the contralateral influence on DPOAEs observed in guinea pigs (e.g., 0 to 3 dB, Kujawa *et al.*, 1993; Puel and Rebillard, 1990) and in humans (e.g. 0 to 2 dB, Moulin *et al.*, 1993; Williams and Brown, 1995), and those of contralateral influence on transiently evoked emissions in chinchillas (0 to 1 dB, Kakigi *et al.*, 1997) and in humans (0 to 2 dB, Hood *et al.*, 1997).

### H. Effects of $L_1-L_2$

Properties of the DPOAE adaptation were observed to be much affected by stimulus parameters. For example, not only the magnitude but also the polarity of DPOAE adaptation was affected by a change in  $L_2$  while  $L_1$  was fixed (Figs. 5, 6, and 9). Complicated patterns of change in DPOAEs observed in the present study are analogous to the early observations by Siegel and Kim (1982) of changes in DPOAEs produced by electrical activation of the OCB at the floor of the fourth ventricle. The latter study demonstrated that DPOAEs were decreased, increased or remained unaffected by activation of the OCB under various acoustic stimulus conditions. As suggested by Siegel and Kim (1982), we consider that, even if the effect of OCB activation on cochlear-partition motion at a local place is always suppressive, complicated patterns, such as an increase of the emission level in the ear canal by OCB activation, can arise as a result of vector-signal summation along the length of the cochlea which may sometimes include out-of-phase cancellations. Further studies would be needed to delineate the overall picture of the magnitude and polarity of the DPOAE adaptation as a function of stimulus parameters and underlying mechanisms.

### I. Relationship between the adaptation and DPOAE level

In the present results, there was a tendency for the absolute DPOAE level (indicated in the figure captions) to decrease with increasing age. This, combined with a decrease of the adaptation magnitude with increasing age described above, suggests a correlation between the two measures, i.e., the adaptation magnitude should be larger for higher DPOAE level. A part of the present data, e.g., the CBA data obtained with 16 kHz (described in the captions of Figs. 1, 3, and 4, and Table I), and the C57 data obtained with monaural stimulation (Figs. 7 and 8, and Table IV), bear out such a correlation. The overall examination of all of the present results, however, did not yield a clear correlation between the two measures. The reason for this is not clear.

### J. Effects of aging

The present study examined properties of the DPOAE adaptation as a function of age. To our knowledge, effects of OC activity on DPOAEs in ears with hearing impairment have not been previously reported. The general appearance of the DPOAE adaptation process in aged ears was similar to that in young ears. In comparison to young mice, both the aged CBA and C57 mice tended to have smaller adaptation magnitudes and longer time constants. For human subjects,

contralateral suppression of transiently evoked OAEs was also reported to be less in older than in younger subjects (Castor *et al.*, 1994; Hood *et al.*, 1997). Results of the present and these previous studies tentatively suggest that, in old subjects, both cochlear OHC function and olivocochlear efferent system function are impaired.

A definitive conclusion concerning the olivocochlear function in aged ears, however, cannot be made until the following two issues are resolved. The first issue regards the stimulus level. Because old ears with hearing loss have higher thresholds than young ears, the sensation level is reduced in old ears when the stimulus is presented at a constant dB SPL, as was done in the present study. Ideally, results from various age groups obtained with matched sensation levels (in addition to constant SPL) would help address questions regarding effects of both sensation level and SPL. The second issue regards the source(s) of age-related change in the DPOAE adaptation. The source(s) of the age-related reduction of the adaptation magnitude observed in this study may possibly be alteration(s) in the olivocochlear system, in the cochlea (i.e., distortion-generation process) or both. Surgical or pharmacological disruptions of the olivocochlear system would help resolve this issue.

### K. C57 versus CBA mice

Comparing the results from C57 (Table V) with those from CBA (Table I), both recorded at 2 months of age, we note that C57 mice exhibited a smaller average magnitude (0.7 dB) of the DPOAE adaptation than CBA mice (1.2 dB). This is an unexpected finding because, at 2 months of age, both strains of mouse are commonly believed to have normal function of the auditory system. Regarding the absolute DPOAE levels in dB SPL, the two strains generally exhibited similar values (11.7 and 11.3 dB SPL on the average for CBA and C57, respectively). The auditory brainstem responses have been also observed to be generally similar between the two strains at 2 months of age (e.g., Henry and Chole, 1980; Sun, 1998). A study by Li and Borg (1991), however, observed that the auditory brainstem response threshold of C57 mice was increased by several dB for frequencies of 8 to 32 kHz from the age of 1 month to 2 months. The smaller magnitude of the DPOAE adaptation in C57 mice at 2 months of age may be due to a slight impairment of the auditory function observed by Li and Borg. Further studies would be needed to elucidate this question.

### L. Continuous stimulation and analysis

Regarding the method employed in the present study, only quasi-continuous stimulation was used due to a minimum 8.1-ms interburst time, a limitation imposed by the digital signal processing board (Ariel DSP-16) used. The interburst time was found to have a significant effect on the DPOAE adaptation, i.e., the DPOAE adaptation magnitude was larger when 61.4-ms two-tone bursts were presented with a shorter (e.g., 8.1 ms) interburst time than with a longer (e.g., 500 ms) one (Lieberman *et al.*, 1996). In studies of the contralateral influence of click-evoked OAEs (Veuillet *et al.*, 1991), an analogous observation was made, i.e., con-

tralateral stimulation with clicks having a shorter inter-click interval was found to increase the suppression magnitude. It is expected that the magnitude of the DPOAE adaptation would be larger than that seen in the present and Liberman *et al.* studies if the interburst time is reduced below 8.1 ms or, more ideally, if a two-tone stimulus is presented continuously for several seconds. In the latter case, one would need to do short-term Fourier analysis with a window width on the order of 50 to 100 ms. Improved instrumentation that can perform such a continuous stimulation and analysis (Neely *et al.*, 1999) would help make measurement of the DPOAE adaptation a more useful tool.

### M. Future studies

The present results were obtained with the upper primary frequency,  $f_2$ , ranging from 8 to 16 kHz. Mice can hear sounds at frequencies reaching 100 kHz (Fay, 1988). This implies that the maximum characteristic frequency of a mouse cochlea may be on the order of 70 kHz. Therefore, it will be desirable that future studies extend the frequency range for the DPOAE adaptation behavior of mice to higher frequencies reaching such high frequencies.

The present results from young adult mice constitute a normal baseline of the DPOAE adaptation phenomenon in mice. These results should be useful in future studies that may use this phenomenon as a noninvasive assay of the function of the olivocochlear system, e.g., a study of developmental changes in postnatal mice, or changes accompanying an alteration in the central auditory system arising from any etiology including aging. We suggest that this phenomenon can be used as a tool for advancing basic knowledge of the auditory system and for assessing an impairment of the olivocochlear system. Application of this method to humans will require an improvement in instrumentation because a lower signal/noise ratio in humans, compared with other mammals, makes it difficult to measure the DPOAE adaptation in humans (Agrama *et al.*, 1998; Neely *et al.*, 1999).

### ACKNOWLEDGMENTS

This study was supported partly by NIH Grants Nos. DC 00360 and P60-AG13631 and HCRAC Faculty Grant of University of Connecticut Health Center. We thank S. Puria for providing us with CUB<sup>o</sup>DISC program, and Y. Ye, K. Parham, D.G. Machado, and J.K. Joseph for assistance in the preparation of the manuscript and comments on the manuscript.

- Agrama, M. T., Waxman, G. M., Stagner, B. B., Martin, G. K., and Lonsbury-Martin, B. L. (1998). "Effects of efferent activation on distortion-product otoacoustic emissions in normal humans using ipsilateral acoustic stimulation," *Assoc. Res. Otolaryngol. Meeting Abstr.* **21**, 152.
- Bonaiuto, G., Parham, K., and Kim, D. O. (1996). "Response properties of dorsal cochlear nucleus neurons of the Purkinje cell degeneration (pcd) mutant and control mice," *Assoc. Res. Otolaryngol. Meeting Abstr.* **19**, 169.
- Brownell, W. E., Bader, C. R., Bertrand, D., and de Ribaupierre, Y. (1985). "Evoked mechanical responses of isolated cochlear outer hair cells," *Science* **227**, 194-196.
- Buño, W. J. (1978). "Auditory nerve fiber activity influenced by contralateral ear sound stimulation," *Exp. Neurol.* **59**, 62-74.

- Campbell, J. P. and Henson, M. M. (1988). "Olivocochlear neurons in the brainstem of the mouse," *Hearing Res.* **35**, 271–274.
- Castor, X., Veuillet, E., Morgan, A., and Collet, L. (1994). "Influence of aging on active cochlear micromechanical properties and on the medial olivocochlear system in humans," *Hearing Res.* **77**, 1–8.
- Collet, L., Kemp, D. T., Veuillet, E., Duclaux, R., Moulin, A., and Morgon, A. (1990). "Effect of contralateral auditory stimuli on active cochlear micro-mechanical properties in human subjects" *Hearing Res.* **43**, 251–262.
- Crispens, C. G. (1976). *Handbook on the Laboratory Mouse* (C. C. Thomas, Springfield, IL), Table 8, General data on reproduction and development, p. 63.
- Davis, H. (1983). "An active process in cochlear mechanics," *Hearing Res.* **9**, 79–90.
- Desmedt, J. (1962). "Auditory-evoked potentials from cochlea to cortex as influenced by activation of the efferent olivocochlear bundle," *J. Acoust. Soc. Am.* **34**, 1478–1496.
- Fay, R. R. (1988). *Hearing in Vertebrates: A Psychophysical Databook* (Hill-Fay, Winnetka, IL), pp. 367–368.
- Folsom, R. C., and Owsley, R. M. (1987). "N<sub>1</sub> action potentials in humans. Influence of simultaneous contralateral stimulation," *Acta Oto-Laryngol.* **103**, 262–265.
- Galambos, R. (1956). "Suppression of auditory nerve activity by stimulation of efferent fibers to cochlea," *J. Neurophysiol.* **19**, 424–437.
- Guinan, J. J. (1996). "Physiology of olivocochlear efferents" in *The Cochlea*, edited by P. Dallos, A. N. Popper, and R. R. Fay (Springer-Verlag, New York), pp. 435–502.
- Henry, K. R. (1983). "Ageing and audition," in *The Auditory Psychobiology of the Mouse*, edited by J. F. Willott (C. C. Thomas, Springfield, IL), pp. 470–493.
- Henry, K. R., and Chole, R. A. (1980). "Genotypic differences in behavioral, physiological and anatomical expression of age-related hearing loss in the laboratory mouse," *Audiology* **19**, 369–383.
- Hood, L. J., Hurley, A. E., Goforth, L., Bordelon, J., and Berlin, C. I. (1997). "Aging and efferent suppression of otoacoustic emissions," *Assoc. Res. Otolaryngol. Meeting Abstr.* **20**, 167.
- Kakigi, A., Hirakawa, H., Mount, R. J., and Harrison, R. V. (1997). "The effect of crossed olivocochlear bundle section on transient evoked otoacoustic emissions," *Hearing Res.* **110**, 34–38.
- Kim, D. O. (1984). "Functional roles of the inner- and outer-hair-cell subsystems in the cochlea and brainstem" in *Hearing Science: Recent Advances*, edited by C. I. Berlin (College-Hill Press, San Diego), pp. 241–262.
- Kim, D. O., and Sun, X.-M. (1998). "Efferent-mediated adaptation of 2f<sub>1</sub>-f<sub>2</sub> distortion product otoacoustic emission in mice," *Assoc. Res. Otolaryngol. Meeting Abstr.* **21**, 151.
- Kim, D. O., Parham, K., Zhao, H., and Ghoshal, S. (1995). "The olivocochlear feedback gain control subsystem: Ascending input from the small cell cap of the cochlear nucleus?" in *Active Hearing*, edited by A. Flock, D. Otteson, and M. Ulfendahl (Elsevier Science, New York), pp. 31–51.
- Kim, D. O., Paparello, J., Jung, M. D., Smurzynski, J., and Sun, X. (1996). "Distortion product otoacoustic emission test of sensorineural hearing loss: Performance regarding sensitivity, specificity and receiver operating characteristics," *Acta Oto-Laryngol.* **116**, 3–11.
- Kim, D. O., Ghoshal, S., and Ye, Y. (1998). "Integration of ascending and descending signals representing stimulus intensity in the marginal shell of the anteroventral cochlear nucleus" in *Psychophysical and Physiological Advances in Hearing*, edited by A. R. Palmer, A. Reese, A. Q. Summerfield, and R. Meddis (Whurr, London), pp. 195–203.
- Kujawa, S. G., and Liberman, M. C. (1998). "Olivocochlear reflex effects on DPOAE amplitude in anesthetized guinea pig," *Assoc. Res. Otolaryngol. Meeting Abstr.* **21**, 139.
- Kujawa, S. G., Glatke, T. J., Fallon, M., and Bobbin, R. P. (1993). "Contralateral sound suppression distortion product otoacoustic emissions through cholinergic mechanisms," *Hearing Res.* **68**, 97–106.
- Li, H.-S., and Borg, E. (1991). "Age-related loss of auditory sensitivity in two mouse genotypes," *Acta Oto-Laryngol.* **111**, 827–834.
- Liberman, M. C. (1990). "Effects of chronic de-efferentation of auditory nerve response," *Hearing Res.* **49**, 209–224.
- Liberman, M. C., Puria, S., and Guinan, Jr., J. J. (1996). "The ipsilaterally evoked olivocochlear reflex causes rapid adaptation of the 2f<sub>1</sub>-f<sub>2</sub> distortion product otoacoustic emission," *J. Acoust. Soc. Am.* **99**, 3572–3584.
- Moulin, A., Collet, L., and Duclaux, R. (1993). "Contralateral auditory stimulation alters acoustic distortion products in humans," *Hearing Res.* **65**, 193–210.
- Mountain, D. C. (1980). "Changes in endolymphatic potential and crossed olivocochlear bundle stimulation alters cochlear mechanics," *Science* **210**, 71–72.
- Neely, S. T., and Kim, D. O. (1983). "An active cochlear model showing sharp tuning and high sensitivity," *Hearing Res.* **9**, 123–130.
- Neely, S. T., Joseph, J., and Kim, D. O. (1999). "Method for continuous measurement of adaptation of distortion product otoacoustic emission," *Assoc. Res. Otolaryngol. Meeting Abstr.* **22**, 97.
- Parham, K. (1997). "Distortion product otoacoustic emissions in the C57BL/6J mouse model of age-related hearing loss," *Hearing Res.* **112**, 216–234.
- Puel, J.-L., and Rebillard, G. (1990). "Effects of contralateral sound stimulation on the distortion product 2f<sub>1</sub>-f<sub>2</sub>: Evidence that the medial efferent system is involved," *J. Acoust. Soc. Am.* **87**, 1630–1635.
- Puria, S., Guinan, Jr., J. J., and Liberman, M. C. (1996). "Efferent-mediated effects of contralateral sound: Suppression of CAP versus ear-canal distortion products," *J. Acoust. Soc. Am.* **99**, 500–507.
- Siegel, J. H., and Kim, D. O. (1982). "Efferent neural control of cochlear mechanics? Olivocochlear bundle stimulation affects cochlear biomechanical nonlinearity," *Hearing Res.* **6**, 171–182.
- Sun, X.-M. (1998). "Distortion product otoacoustic emissions in mouse models of early-onset and late-onset presbycusis," Ph.D. dissertation, University of Connecticut, Storrs, CT.
- Sun, X.-M., Kim, D. O., and Parham, K. (1998). "Effects of stimulus parameters on distortion product otoacoustic emissions in the CBA mouse model of age-related hearing loss," *Assoc. Res. Otolaryngol. Meeting Abstr.* **21**, 79.
- Vetter, D. E., Elgoyhen, A. B., Mann, J., Boulter, J., Saffiote, J., Barhanian, J., and Heinemann, S. F. (1998). "Altered cochlear efferent fiber innervation in alpha9 AChR subunit knockout mice," *Assoc. Res. Otolaryngol. Meeting Abstr.* **21**, 48.
- Veuillet, E., Collet, L., and Duclaux, R. (1991). "Effect of contralateral acoustic stimulation on active cochlear micromechanical properties in human subjects: Dependence on stimulus variables," *J. Neurophysiol.* **65**, 724–735.
- Warr, W. B., and Guinan, Jr., J. J. (1979). "Efferent innervation of the organ of Corti: two separate systems," *Brain Res.* **173**, 152–155.
- Warr, W. B. (1992). "Organization of olivocochlear efferent systems in mammals," in *The Mammalian Auditory Pathway: Neuroanatomy*, edited by D. B. Webster, A. N. Popper, and R. R. Fay (Springer-Verlag, New York), pp. 410–448.
- Warren, E. H., and Liberman, M. C. (1989). "Effects of contralateral sound on auditory-nerve response. I: Contributions of cochlear efferents," *Hearing Res.* **37**, 89–104.
- Williams, D. M., and Brown, A. M. (1995). "Contralateral and ipsilateral suppression of the 2f<sub>1</sub>-f<sub>2</sub> distortion product in human subjects," *J. Acoust. Soc. Am.* **97**, 1130–1140.
- Willott, J. F. (1986). "Effects of aging, hearing loss, and anatomical location on thresholds of inferior colliculus neurons in C57BL/6J and CBA mice," *J. Neurophysiol.* **56**, 391–408.
- Willott, J. F. (1991). *Aging and the Auditory System: Anatomy, Physiology and Psychophysics* (Singular Publishing, San Diego, CA).
- Zenner, H. P., Zimmermann, U., and Schmitt, U. (1985). "Reversible contraction of isolated mammalian cochlear hair cells," *Hearing Res.* **18**, 127–133.
- Zheng, Q. Y., Johnson, K. R., and Erway, L. C. (1999). "Assessment of hearing in 80 inbred strains of mice by ABR threshold analyses," *Hearing Res.* **130**, 94–107.



# The “inverse problem” solved for a three-dimensional model of the cochlea. III. Brushing-up the solution method

Egbert de Boer<sup>a)</sup>

Room D2-226, Academic Center, University of Amsterdam, Meibergdreef 9, 1105 AZ, Amsterdam, The Netherlands

Alfred L. Nuttall<sup>b)</sup>

Oregon Hearing Research Center, NRC04, Oregon Health Sciences University, 3181 SW Sam Jackson Park Road, Portland, Oregon 97201-3098 and Kresge Hearing Research Institute, University of Michigan, 1301 East Ann Street, Ann Arbor, Michigan 48109-0506

(Received 1 April 1998; revised 12 February 1999; accepted 19 March 1999)

In two earlier papers [de Boer, *J. Acoust. Soc. Am.* **98**, 896–903 and 904–910 (1995)] the inherent problems of the inverse-solution method in cochlear mechanics were described. The present paper shows results obtained with a more universal solution method. With the new method it is possible to construct a three-dimensional model of the cochlea producing a response that accurately simulates a measured mechanical basilar-membrane response. With earlier methods this could not be done. The inverse solution invariably yields that, with low stimulus levels, the model simulating a viable cochlea must be locally active. For the response of a dead animal a passive model is sufficient. Once more the inherent intricacies and problems of the inverse-solution method are discussed. Conservation of fluid volume leads to the concept of the “virtual stapes velocity.” For best results, the input signal to the inverse-solution procedure should be acquired in the form of a “composite cross-correlation spectrum.” Inverse analysis can, under certain circumstances, produce aberrant results. In this paper it is shown why the resulting impedance function is the most accurate in the region of the response peak. Therefore, it is unlikely that a passive model would exist of which the response simulates the data obtained from a healthy animal. © 1999 Acoustical Society of America. [S0001-4966(99)05406-5]

PACS numbers: 43.64.Kc, 43.64.Bt [RDF]

## INTRODUCTION

This paper continues the path followed in two earlier publications of the first author (de Boer, 1995a, 1995b), henceforth to be referred to as (INV-1) and (INV-2). The cited work described an “inverse solution” with which the main parameter function of a model of cochlear mechanics, the *basilar-membrane impedance function*, is retrieved from a measured response function. The inverse-solution method has many problems and pitfalls; therefore, the response computed from a cochlear model with the retrieved impedance—in (INV-2) termed the “resynthesized response”—will not automatically be identical to the measured response. Despite these problems it was shown that the resynthesized response had a major peak which resembled that of the measured response in several respects. In the meantime, a new set of experimental data on the response of the basilar membrane (BM) has become available (de Boer and Nuttall, 1997b—GLIDE), and the analysis technique has been further studied and improved (de Boer, 1998—NEW3D). Preliminary results were presented in de Boer and Nuttall (1997a—DAM) where a much better agreement between original and resynthesized responses was demonstrated than in (INV-2). In the present paper results are shown of application of the most fully developed technique

to date. With that technique the resynthesized response is (in almost all instances) identical to the original response—to within the accuracy of plotting.

With low-level stimuli, the recovered BM impedance shows a region where the BM enhances the power of the waves in the model. Thus to simulate a low-level response, the model has to be “locally active” (cf. de Boer, 1993, 1996). Strictly speaking, the use of a linear concept like “BM impedance” is not allowed for a nonlinear system as the cochlea. However, when only statistical concepts like “input–output cross-correlation functions” are utilized in defining the “response” of the cochlea, a straightforward interpretation of the results in terms of the theory of a linear system is possible; this is expressed by the EQ-NL theorem (de Boer, 1997b—CONNEX). All procedures and conclusions will then relate to that linear system. Therefore, we will only consider a linear model—as we did in (INV-1) and (INV-2). In the same vein, we will always express the response of the cochlea in terms of the input–output cross-correlation function (ccf) for wide-band noise stimulation. When the linear model is extended to include nonlinear elements, we tacitly assume that it remains within the class of models for which the EQ-NL theorem holds.<sup>1</sup> In this context it can be stated that with higher levels of stimulation the degree of activity of the cochlea diminishes (de Boer and Nuttall, 1997a—DAM). In responses measured post-mortem, activity has vanished.

The inverse solution method is often called an “ill-

<sup>a)</sup>Electronic mail: e.deboer@amc.uva.nl

<sup>b)</sup>Electronic mail: nuttall@ohsu.edu

posed problem” because small variations in the input data may cause large variations in the recovered BM impedance. It is shown in the present paper that this is not invariably true.

## I. SOLUTION METHODS

The cochlear model to be considered is the same “stylized” three-dimensional model as was used in (INV-1) and (INV-2). The main variables are:  $x$ , the coordinate along the longitudinal axis, and  $\omega$ , the radian frequency. The variable  $\omega$  should be considered a constant, and  $x$  functions as the independent variable. In the earlier work we applied an approximation of the model equation described by de Boer and van Bienema (1982). In the meantime, this approximation method has been improved considerably (de Boer, 1998–NEW3D). Still, the method is limited to a model of constant geometry along its length (a “stylized” model). We have also generalized the solution method to a form involving a matrix equation with a “full” matrix (cf. Allen, 1977; Mammamo and Nobili, 1993). That method can, if desired, be applied to a model of arbitrary shape. In the present paper we use the full-matrix method for the same stylized model as in the earlier papers.

In both of the new formulations internal wave reflections are handled much better than with the method used in (INV-1) and (INV-2). Such reflections may well arise since the recovered impedance function  $Z_{BM}(x, \omega)$  will be irregular due to random errors in the data. Appendix A summarizes the assumed geometrical and mechanical properties of the model, and presents the basic solution formulae. As a result of the improvement, the resynthesized response is identical (within plotting accuracy) to the response used as input to the inverse procedure (except in cases where the impedance function turns out to be exceedingly irregular).

## II. RESULTS I. THE SIMPLEST CASE

Data were collected on movements of the basilar membrane (BM) in the basal turn of the guinea pig cochlea using a laser velocimeter (cf. Nuttall *et al.*, 1990; Nuttall and Dolan, 1996). We recorded from a location that, in a viable cochlea, has its best frequency (BF) between 16 and 18 kHz. The method of data acquisition is described in de Boer and Nuttall (1997b–GLIDE). For all data used in the present paper *bands of noise* were used as acoustical stimuli, and the input–output cross-correlation function (ccf) was computed (see Appendix A of the last-cited paper). Fourier transformation leads to the input–output cross-correlation spectrum (to be called the ccf spectrum). In a similar way the wide-band ccf spectrum for the stapes velocity was obtained. The ccf spectrum of the BM response was divided by the ccf spectrum of the stapes response and corrected for differences in stimulus spectral density. As a result, *all response data in this paper represent the ratio of BM to stapes velocity*. Appendix B gives details (including a figure) on the response of the stapes and on signal processing.

Let us first consider *post-mortem* data (obtained approximately 20 min after the animal was killed by cardiac injection of sodium pentobarbital). In this case the stimulus signal

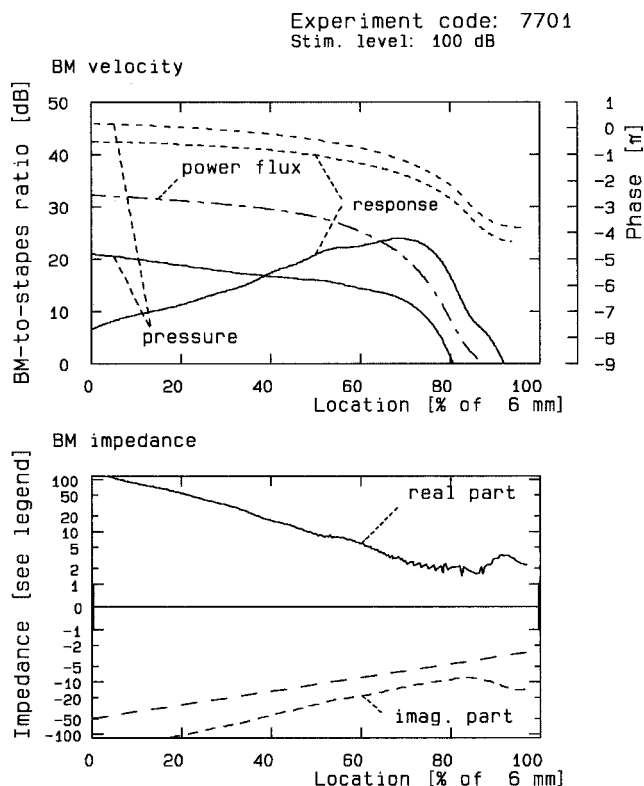


FIG. 1. Inverse solution based on post-mortem data. *Upper panel.* Response converted to the place ( $x$ -)domain (amplitude and phase). Animal code: “7701.” Actual best frequency (BF): 13.4 kHz; BF used in  $f$ -to- $x$  conversion: 16.8 kHz. Stimulus level (dB SPL over one octave) 100 dB. Maximum BM-to-stapes velocity ratio: 15.6. 0 dB on the amplitude scale corresponds to 1. Added functions: pressure near the BM (amplitude and phase) and power flux (dash-dot line). For the pressure 0 dB corresponds to 50 dyne  $\text{cm}^{-2}$  ( $5 \text{ kg m}^{-1} \text{ s}^{-2}$ ). *Lower panel.* Effective BM impedance recovered by inverse solution (real and imaginary parts). Three-dimensional model. The impedance scale is nonlinear: values from “-1” to “+1” are shown linearly, larger values are logarithmically compressed. The norm “ $l$ ” corresponds to  $200 \text{ g cm}^{-2} \text{ s}^{-1}$ , or  $2000 \text{ kg m}^{-2} \text{ s}^{-1}$ . Virtual stapes velocity: 2.68-0.75*i*. Coarsely dashed line: exponential function of  $x$ , see text.

was a wide band of noise covering frequencies from 1 to 40 kHz, presented at the highest available intensity from the sound system. In the first step of the inverse method the measured ccf spectrum has to be transformed from the *frequency* to the *place* domain. This amounts to going from the “local” to the “panoramic” view, in the terms of Patuzzi (1996), and use is made of the great similarity in shape of response functions in the two domains. Details about this transformation—which includes minor modifications of phase and amplitude—are given in Appendix C. In the transformation we used the following parameters:<sup>2</sup>

$$\alpha = 5 \text{ [cm}^{-1}\text{]}; \quad \omega_{\max} = 2\pi 60 \text{ [kHz]},$$

$$L = 0.6 \text{ [cm]}, \quad N = 256. \quad (1)$$

The length  $L$  of the model is divided into  $N$  sections,<sup>3</sup>  $\alpha$  and  $\omega_{\max}$  are parameters involved in the cochlear map [see Eq. (C1)]. The result is the BM velocity  $v_{BM}(x, \omega)$  as a function of location  $x$ , and this should be regarded as *the response of the chosen model to a tone with radian frequency  $\omega = 2\pi f_{BF}$  (where  $f_{BF}$  is the best frequency, or BF) applied to the stapes*. Figure 1, upper panel, shows a typical post-mortem BM velocity function  $v_{BM}(x, \omega)$  after this transfor-

mation plotted as a function of location  $x$ . Amplitude and phase are shown by a solid and a finely dashed line, respectively. In this particular case the BF location ( $x_{\text{BF}}$ ) was artificially displaced to the location corresponding to 16.8 kHz, which is the BF of the “live” cochlea of the same animal. The stimulus level indicated applies to one octave (10–20 kHz) of the noise spectrum, and corresponds to 106 dB SPL for a pure-tone stimulus.

Let us consider the *long-wave* model where the pressure  $p(x, \omega)$  is found by double integration of the velocity  $v_{\text{BM}}(x, \omega)$  [see Eq. (A4)]. We are left with two unknown integration constants. That is, to the obtained pressure solution  $p(x, \omega)$ , a linear combination  $p_e(x, \omega)$  like

$$p_e(x, \omega) = p_0 + p_1(x - L) \quad (2)$$

has to be added, where ( $x = L$ ) is the location of the helicotrema, and we must find suitable values for  $p_0$  and  $p_1$ . A similar problem was met in earlier work (INV-1 and INV-2) and solved by *setting both integration constants to zero*. Here we will discuss this problem in more general terms. When the integration has been started at the helicotrema and if it is assumed that at the helicotrema the pressure is zero, the constant  $p_0$  must clearly be zero. Similarly, setting  $p_1$  to zero implies that at  $x = L$  the derivative of the pressure  $p_e(x, \omega)$  is zero so that the extra term causes no fluid flow across the helicotrema. But what else does it imply?

Integrate Eq. (A4) once over all  $x$ ; it follows that at ( $x = 0$ ) the derivative of the pressure  $p(x, \omega)$  is  $(-i\omega\rho/h_{\text{eff}})$  times the integral of  $v_{\text{BM}}(x, \omega)$  from ( $x = 0$ ) to ( $x = L$ )—see Appendix A for the meaning of  $h_{\text{eff}}$ . Divided by  $(-i\omega\rho)$  this derivative yields a velocity which we have termed “virtual stapes velocity.” This, then, is the value we have to use for the stapes velocity in the inverse solution: the integral of  $v_{\text{BM}}(x, \omega)$  from ( $x = 0$ ) to ( $x = L$ ) divided by  $h_{\text{eff}}$ . *The result is that in each channel of the model the total volume of fluid remains constant.* The virtual stapes velocity will, in general, not be equal to the actual stapes velocity used in the experiments, but we hope that it will be nearly real and of order one.

Another, more general, way of expressing this adjustment of the stapes velocity to the integral of the BM velocity is the following. Suppose we have obtained a set of BM velocity data and have transformed these to the  $x$ -domain. We have assumed a certain shape of the model and imposed our transformed response function  $v_{\text{BM}}(x, \omega)$  upon it as its response. We do not know, however, whether our  $f$ -to- $x$  transformation is correct, or whether the geometries of the model and the BM are good approximations of the actual cochlea. It is, furthermore, not guaranteed that, in the  $x$ -domain, the total amount of fluid displaced by the BM equals the amount of fluid displaced by the stapes. We must adhere to the physical laws under which the model equation was derived, and one of these laws includes conservation of fluid volume. Therefore, we use the freedom we have to prescribe that the total fluid volume in one channel of the model remains constant. That means that the stapes velocity  $v_{\text{st}}$  cannot simply be made equal to unity but must be *adjusted* to the integral of  $v_{\text{BM}}(x, \omega)$  over all values of  $x$ , divided by  $h_{\text{eff}}$  (so that  $v_{\text{st}}$  will be equal to the “virtual stapes

velocity”). Because the reasoning is independent of the type of model involved, we will use the same procedure for the three-dimensional model and for the recordings from alive and dead animals.

Let us now turn to treating the post-mortem data with the *three-dimensional* model. The inverse solution was computed using the method described in Appendix A, with the following model parameters [see also Eq. (1)]:

$$\varepsilon = 0.2; \quad b = 0.1 \text{ [cm]}; \quad h = 0.1 \text{ [cm]}. \quad (3)$$

The parameter  $\varepsilon$  is the width of the BM divided by the width  $b$  of the model, and  $h$  is the height of one channel of the model.

The lower panel of Fig. 1 depicts the resulting BM impedance. The solid line shows the real part, and the finely dashed line the imaginary part. The ordinate scale is nonlinear, values with amplitudes smaller than “1” are plotted linearly (marked by the thick vertical bars on the sides), and values outside that range are logarithmically compressed. See the legend for the normalization constant, i.e., the impedance value corresponding to “1.”

Clearly, the *real part* of the BM impedance is positive from ( $x = 0$ ) to the end of the response peak. This means that the best-fitting model is a passive one. The *imaginary part* of the BM impedance is negative in the region to the left of the peak, and it remains negative throughout the region of the peak. From the fact that the imaginary part does not cross the zero line we infer that in this case resonance of the BM is not the primary cause of the roll-off of the amplitude. Therefore, it must be the *short-wave* nature of the wave that causes the amplitude ultimately to go down. We observe the phase in this region to vary by  $2\pi$  over a distance of the order of 1 mm which is fast enough to rule out the long-wave model. In the short-wave region the wave is extinguished before it reaches the location where the BM shows resonance. The physical property underlying this phenomenon has been described as “boundary-layer absorption” by Lighthill (1981). More about the physics of short waves in a cochlear model can be found in de Boer (1996), section 5.2. It follows that a long-wave model is not sufficient to describe the response of the “dead” cochlea.

The imaginary part of the BM impedance appears to vary like an exponential function of  $x$ . A coarsely dashed line is added to the figure that shows the function  $[C \exp(-\alpha x)]$  (with  $C$  equal to  $-10^4$ ). This line has a slightly smaller slope than the imaginary part of the impedance in the leftmost part of the figure indicating that the estimate of  $\alpha$  ( $5 \text{ cm}^{-1}$ ) has been somewhat too low (cf. Tsuji and Liberman, 1997).

The feature that the real part of  $Z_{\text{BM}}(x, \omega)$  is positive in the region of the response peak is robust, we found it in all 11 post-mortem experiments analyzed. The findings presented can be interpreted as corroborating the use of a simple standard passive model for the response of a dead animal.

Resynthesis is done by solving the forward problem with the obtained impedance function. Because complementary methods are used for inverse analysis and resynthesis, the computed response is indistinguishable from the input

response. In this case, resynthesis is “perfect” and there is no need to show the result.

The upper panel of Fig. 1 includes graphs for the *pressure*  $p(x, \omega)$  computed near the BM [Eq. (A1)]. The pressure’s amplitude function is seen to be gradually decreasing with increasing  $x$ . Its phase approximately varies as that of  $v_{\text{BM}}(x, \omega)$ . One more function is shown in the upper panel of Fig. 1, the *power flux* (dash-dot line). The latter function is computed by integrating the density of power dissipation by the BM over  $x$  (this density is equal to 0.5 times the product of the width of the BM, the pressure difference across the BM and the complex conjugate of the BM velocity). It is normalized to be zero at ( $x=L$ ). In a passive model the power flux obviously is a decreasing function of  $x$ .

### III. INTERLUDE I—THE COMPOSITE CCF SPECTRUM

Let us now turn to the case where a cochlea in good physiological condition is stimulated by a weak signal. To measure the response to a band of noise centered at the best frequency (BF) and presented at a very low level, the stimulus’ spectrum should have a bandwidth of one or two octaves. With a measuring time of approximately 60 s, the peak region in the ccf spectrum is well represented; but at levels 25–30 dB below the peak, the ccf spectrum becomes too noisy for the inverse-analysis procedure. Using a wide-band stimulus does not remove this defect because weak response components outside the passband are still lost in measurement noise.

To obtain an accurate record over a much wider frequency range, several records taken with stimuli with different noise bandwidths and presented at different levels were combined. First, the ccf spectrum in the BF region was measured by using a narrow-band stimulus centered at the BF and presented at a low level. Only the most central part of this ccf spectrum was kept; if we call the lowest crossover frequency  $f_a$ , the spectrum was stored from the frequency ( $f_a - f_0$ ) upward where  $f_0$  equals 450 Hz. Next, a stimulus containing more low-frequency components was presented, at a higher level, and the ccf’s spectral content in the frequency region from  $f_b - f_0$  to  $f_a + f_0$  was kept, where  $f_b$  is below  $f_a$ . This procedure was repeated a few times, with lower and lower bands of noise. Joining all partial ccf spectra, taking into account differences in stimulation level and spectral density of the stimuli used, produces what is called the *composite ccf spectrum*. In combining the various partial spectra, smooth transitions were created around the connection points  $f_a$ ,  $f_b$ , etc. with a window function extending from  $-f_0$  to  $+f_0$ . It is because the cochlear response is (nearly) linear at frequencies below the BF that such a combination is justified. The resulting composite spectrum can be considered as the ccf spectrum of a linear system, namely, of the linear model mentioned in the Introduction. Therefore, the composite ccf spectrum should be interpreted *as if it had been possible to measure the ccf spectrum in one experiment with a constant-amplitude weak stimulus over the entire frequency range from below 1 kHz to over 25 kHz*. The composite ccf spectrum was again made to express the ratio of BM to stapes velocity. At low stimulation levels four partial spectra were used; in measuring at very high levels two were

enough. For the data in dead animals and for the response of the stapes the composite-spectrum technique was not necessary (see Sec. II).

Figure 2(A), upper panel, presents a typical result, the composite ccf spectrum measured with a noise stimulus at a level of 20 dB (in the BF region), and shown here as a function of frequency. For the sake of completeness we repeat that the “stimulus level” is the dB SPL over one octave (10–20 kHz) of the noise spectrum. From the lowest frequencies to the BF (16.8 kHz) the composite ccf spectrum is observed to cover an amplitude range of over 50 dB, and for the frequency region *above* the BF the spectrum is accurate to within 25–30 dB down from the peak. In Fig. 2(A) the amplitude of the spectrum is normalized to 0 dB at its peak; the actual velocity ratio at the peak frequency (BF) is 404. In the leftmost region of the figure unwrapping of the phase has not been successful, and the phase curve is only shown from 2.8 kHz upward.

The lower panel of Fig. 2(A) shows the *ccf waveform function* that is associated with the composite ccf spectrum. When we assume that the cochlea is linear at this low level of stimulation, we can interpret the ccf waveform as *the impulse response of the BM*. Note that this waveform shows a “glide:” in the initial part of the waveform the frequency of oscillation is lower than in the later part (cf. de Boer and Nuttall, 1997b–GLIDE).

It is possible to estimate BM velocity and displacement at the point of maximum sensitivity. In Appendix B it is stated that the average stapes velocity for sinusoidal stimulation at 100 dB SPL is 0.877 mm/s. If we consider stimulation by a sinusoidal signal at 20 dB SPL, and assume the maximum BM-to-stapes ratio to be 500, the maximum BM velocity is 0.0439 mm/s and the BM displacement (at 17 kHz) is 0.410 nm.

Figure 2(B) illustrates the comparison of our data to those of related experiments reported in the literature (amplitude only). The smooth curves show a typical frequency response published by Sellick *et al.* (1983) and one from Robles *et al.* (1986). These curves are plotted from the smoothed data files produced by Diependaal as they have been described in (INV-2), section II. The Sellick *et al.* data, labeled “SYP 90,” reflect the ratio of BM and incus velocity for the *guinea pig*. For the Robles *et al.* data, labeled “RRR 328,” a very simple middle ear model was inserted to derive the BM-to-stapes ratio from the data since these were obtained with constant sound pressure at the tympanic membrane; see Diependaal *et al.* (1987) for more details. In the latter experiments the experimental animal was the *chinchilla*, and the data are shown as shifted upward in frequency by one octave. The more irregular curves are composite ccf spectra computed from our own (*guinea pig*) data. See for the amplitude normalization of the curves the legend to the figure. Of all responses the actual maximum values of the ratio shown are given in the legend. It is clear that our own data are close to the SYP 90 curve, over nearly the entire frequency range the deviations are limited to 10 dB. Variations in the maximum BM-to-stapes ratio cover a range of 16 dB. The (shifted) *chinchilla* data show about the same relative bandwidth in the peak, but have a “dip” in the region

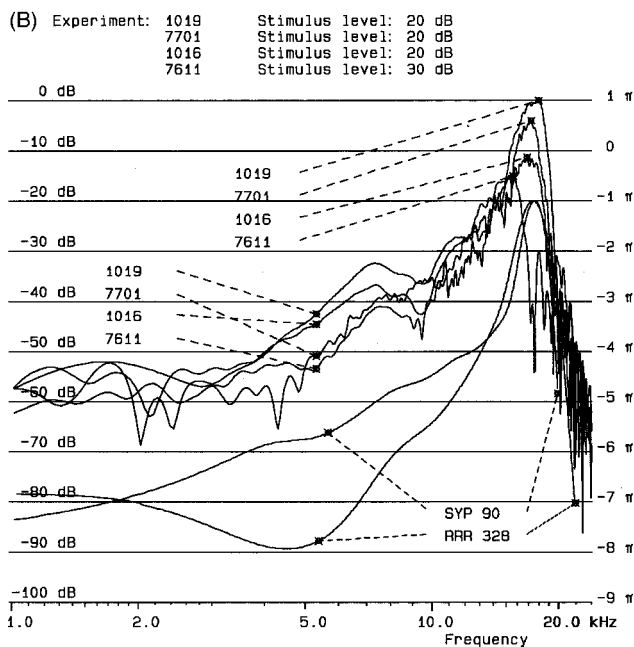
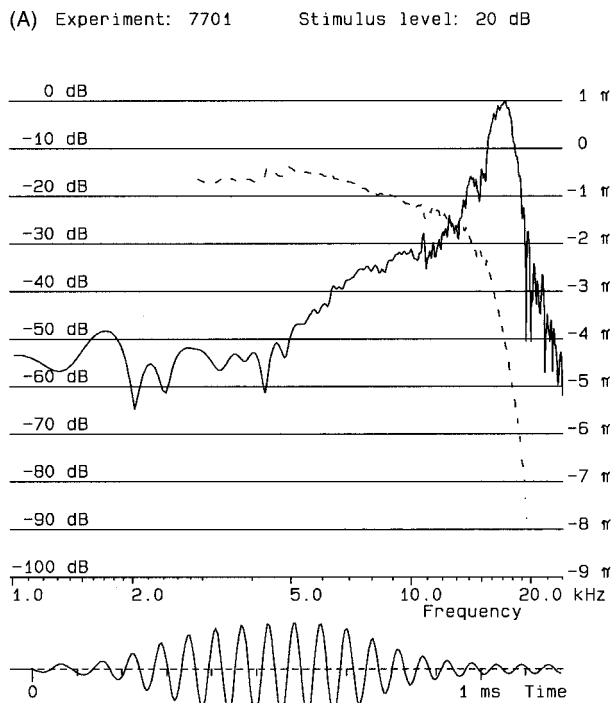


FIG. 2. (A) *Upper panel*. A typical composite ccf spectrum for low-level stimulation of a sensitive cochlea (see text). Animal code: “7701.” Stimulus level 20 dB. Maximum BM-to-stapes velocity (in the figure normalized to 0 dB): 404. Best frequency: 16.8 kHz. Connection points of partial spectra: 3, 5.8, 10, and 15 kHz. Overlap on both sides: 450 Hz. Note that measurement errors are larger to the right than to the left of the connection points (10 and 15 kHz). *Lower panel*: the corresponding impulse response. (B) Comparison of response curves (amplitude only). Labels have the following meaning. “SYP 90”: Sellick *et al.* (1983), guinea pig 90. “RRR 328”: Robles *et al.* (1986), chinchilla 328, these data are plotted one octave higher. Labels “1019,” etc. denote experiments as described in this paper. For “SYP 90” and “RRR 328” the amplitude is normalized to -20 dB at the peak. For our own data (composite frequency spectra) the first record (1019) is normalized to 0 dB at its peak, the other three have the same normalization. Stimulus level: 20 dB, except for 7611: 30 dB. Maximum BM-to-incus ratio for “SYP 90” is 663, the inferred maximum BM-to-stapes ratio for “RRR 328” is 2232. The maximum BM-to-stapes in our own data are: for 1019: 1059; for 7701: 404; for 1016: 287; for 7611: 188.

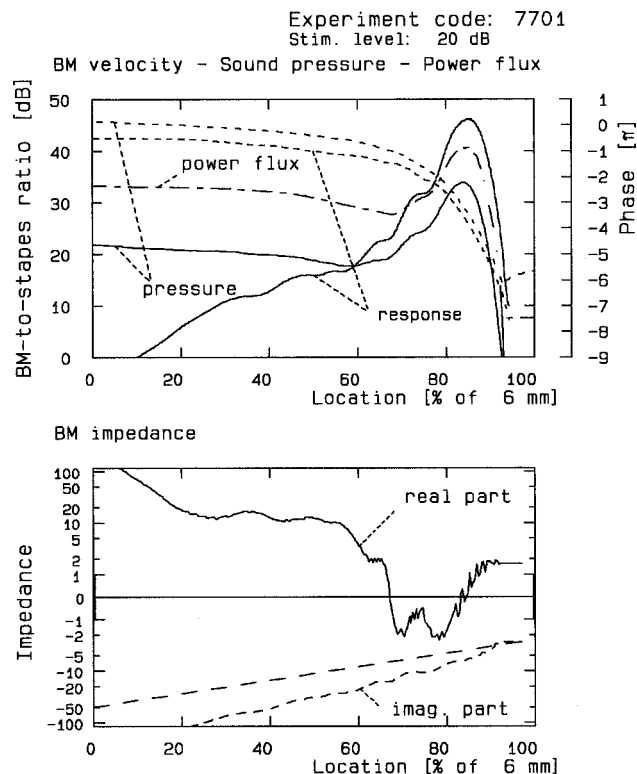


FIG. 3. *Upper panel*: response of Fig. 2(A) converted to the place domain—and slightly smoothed. Smoothing parameter (see Appendix B):  $\tau_1 = 612 \mu\text{s}$ . Layout as in Fig. 1, 0 dB on the amplitude scale corresponds to 1. Actual amplitude of BM velocity is shown reduced by 6 dB. Maximum BM-to-stapes velocity ratio: 404. *Lower panel*: effective BM impedance recovered by inverse solution. Animal code: “7701.” Three-dimensional model. Virtual stapes velocity: 2.61-0.85*i*. Coarsely dashed line: exponential function of  $x$ , the same as in Fig. 1.

around 5 kHz. For these data we cannot compare the maximum BM-to-ratio to our values because of the uncertainties in the middle-ear model employed. From the data in Ruggero *et al.* (1997), their Fig. 9, we infer that, for a sinusoidal stimulus at 20 dB SPL at the BF (10 kHz in this case) the maximum BM velocity is 0.16 mm/s.

#### IV. RESULTS II. THE EFFECTIVE BM IMPEDANCE

Figure 3 illustrates input to and output of the inverse procedure for the response data shown in Fig. 2(A). The upper panel of this figure shows the amplitude and phase of the BM response as functions of  $x$  just as in Fig. 1. The response is similar in shape to the composite ccf spectrum shown in Fig. 2(A) but the irregularities in the spectrum have been reduced by a moderate amount of smoothing (see Appendix B). The point ( $x=0$ ) in Fig. 3 corresponds to the frequency 4.68 kHz in Fig. 2(A), and is made to have a phase of  $-\frac{1}{2}\pi$  radians (see Appendix C). Using the parameters given by Eq. (1), the best frequency (16.8 kHz) was projected to the location  $x_{\text{BF}}$  corresponding to it.

The lower panel of Fig. 3 presents real and imaginary parts of the BM impedance  $Z_{\text{BM}}(x, \omega)$  as recovered by the inverse procedure for the three-dimensional model. See the legend for the value of the virtual stapes velocity. The *real part* of  $Z_{\text{BM}}(x, \omega)$  shows a clear excursion into the domain of negative values, and this occurs in the region where the re-

sponse amplitude is rising the most rapidly. Apparently, in this region the model should be “active” to produce the required response. This result confirms what has been predicted theoretically (de Boer, 1983) and found from actual data in earlier studies (INV-2; de Boer, 1993, 1996, Section 6.3; de Boer and Nuttall, 1997a–DAM). The highest point of the response peak lies where the real part of  $Z_{\text{BM}}(x, \omega)$  crosses the zero line from negative to positive. The *imaginary part* of the BM impedance remains negative over a large part of the  $x$ -axis including the peak region. Note that this imaginary part is not very different from the one shown in Fig. 1. It has been verified that the *input impedance* of the model (stapes pressure divided by virtual stapes velocity) has a positive real part, hence the model is stable for the frequency for which its response is shown (the BF). We can infer that it will also be stable for frequencies above and below it. We found similar results in 15 other experiments.

The response shown in Fig. 3 is assumed to be the response of the model to a tone with a frequency equal to the BF. For a tone with higher frequency the response pattern shifts to the left, and for a tone with lower frequency to the right. The BM impedance function will show corresponding shifts (in addition to being proportional to  $\omega$ ). The impedance function is recovered as a function of  $x$  for one frequency, the BF. For a higher frequency the impedance function will shift to the left (and will be multiplied by the frequency ratio), and for a lower frequency to the right. When viewed at a *fixed location*, the BM impedance should depend on (log) frequency in approximately the same way as shown in Fig. 3, lower panel, as a function of  $x$ . Then, for a location more to the left, the impedance function scales to a higher frequency, etc. This type of dependence of the BM impedance on frequency and place is common to almost all models of the cochlea. The simplest type of passive model shows it [de Boer, 1991, Eqs. (2.2.a,b,c)], see also Appendix C. Locally active models of the cochlea have been described of which the BM impedance has the same property. The most straightforward of these models has been published by Neely and Kim (1986).

In the upper panel of Fig. 3 two more functions are shown—the pressure near the BM (amplitude and phase) and the power flux. In a passive long-wave model the pressure is a decreasing function of  $x$  (Fig. 1). The peak that we observe in Fig. 3 reflects the power gain in the model. The power-flux function, again normalized to zero at ( $x=L$ ), shows a somewhat smaller rise than the pressure. In comparing the power-flux functions of Figs. 1 and 3, we observe that at the peak of the low-level response the power has been increased by approximately 35 dB over that of the passive system. Considered with respect to the value at the stapes location, the over-all power gain is much smaller, not more than 10 dB. Apparently much of the “activity” is used to compensate for dissipation.

There is no need to show a *resynthesis* result when the BM impedance resulting from the inverse solution is used because resynthesis is again perfect in this case. We conclude that we have succeeded in constructing a three-dimensional model of the cochlea that has a response that

corresponds very closely to the one measured. This is true for “dead” as well as “alive” cochleae.

## V. INTERLUDE II—A CHAIN OF TRANSFORMATIONS

We can consider the sequence from measured BM response, via the recovered BM impedance, to the resynthesized response, as a *sequence of transformations* in signal space. In this light we will briefly discuss the influence of measurement errors. The first step in the transformation is the projection from frequency to place, this involves a change of the independent variable and not much more. The second step is the inverse solution. This constitutes a *nonlinear transformation* because, according to Eqs. (A1) and (A2), the impedance  $Z_{\text{BM}}(x, \omega)$  is a *linear transform* of the BM velocity  $v_{\text{BM}}(x, \omega)$ , *divided* by the BM velocity itself. In general, Green’s functions are functions of two variables,  $x$  and  $x'$ . For small values of  $|x-x'|$  the Green’s function is dominated by the familiar logarithmic singularity (Sondhi, 1978; Allen and Sondhi, 1979) that is associated with short-wave behavior. Around the response peak, short-range effects of the Green’s function have the largest influence on the linear transform. This means that the transform will be affected mainly by errors in the data over a limited length. In the region of the stapes, however, both local effects and long-range effects of the Green’s function (originating from the much larger response in the peak region) contribute to the transform. It is clear that, after division by  $v_{\text{BM}}(x, \omega)$ , relative deviations in  $Z_{\text{BM}}(x, \omega)$  due to data errors will be small in the peak region but will increase toward the base of the model. Therefore, it is in the peak region where the result of the inverse procedure is the most reliable.<sup>4</sup> Moreover, in that region the relative error in  $Z_{\text{BM}}(x, \omega)$  is approximately proportional to the relative error in the data, so that no discontinuity or bifurcation is to be expected when the error size increases (up to a certain limit, of course). In other words, the inverse solution is a *continuous* transformation in signal space. In this view, it is extremely unlikely that a passive model could show a “correct” response in cases where the inverse-solution method indicates that a locally active model is required.

## VI. RESULTS III—THE EFFECT OF SMOOTHING

We will now illustrate the effect of *smoothing of the BM impedance*. Figure 4 shows, in an enlarged panel, the original response (solid line, amplitude, finely dashed line, phase) in the form directly derived from the composite ccf spectrum, without any additional smoothing, together with the recovered BM impedance [the data are from another animal than the one used in Figs. 2(A) and 3]. For greater clarity the abscissa is made to show only the right-half 50% of that of earlier figures. The BM impedance (solid lines) shows many irregularities, apparently linked with the irregularities in the response. The second impedance function (smooth dashed lines) results from severely smoothing the first impedance function (see Appendix B). Figure 5 illustrates resynthesized response functions. The resynthesized response corresponding to the unsmoothed impedance of Fig. 4 is identical to the original response—it is labeled “Or” in Fig. 5. The resyn-

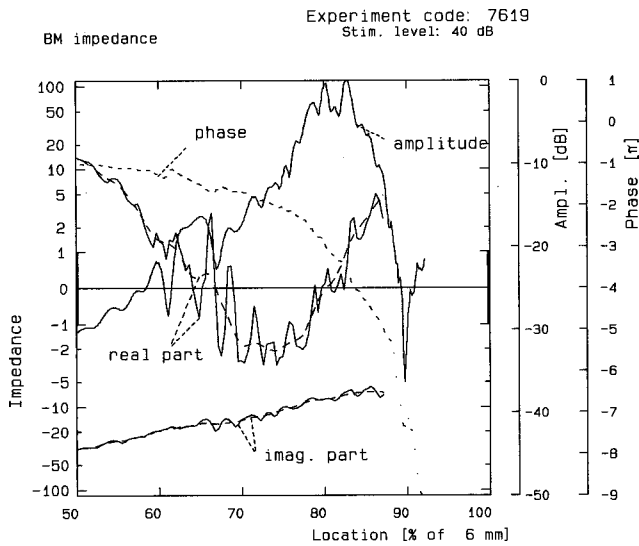


FIG. 4. Unsmoothed response functions (amplitude and phase, as indicated), and the effect of smoothing on the impedance function. Animal code: "7619." Response amplitude normalized to 0 dB at the peak. Maximum BM-to-stapes velocity ratio: 177. Virtual stapes velocity:  $1.98-0.44i$ . Solid lines: unsmoothed BM impedance function. Dashed lines: smoothed impedance function.

thesized response that corresponds to the smoothed impedance function is labeled "Rs." Only the amplitude function is shown because the phase function is almost indistinguishable from the original one. The resynthesized response "Rs" is a smooth function of  $x$ , but it appears to have lost about 5 dB of its peak. We can trace the origin of this effect as follows. The amplification that the wave in the model undergoes is a nonlinear (in fact, an expansive) function of the real part of the BM impedance. Therefore, if we smooth the impedance function, giving equal credit to positive and negative deviations from the "average" course of this function,

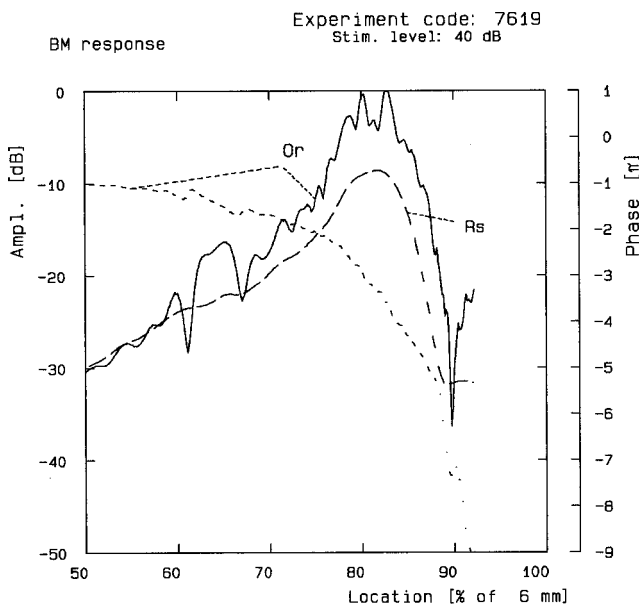


FIG. 5. Resynthesized response corresponding to Fig. 4 (animal code: "7619"). Solid line and finely dashed line: unsmoothed response ("Or"), amplitude and phase. Coarsely dashed line ("Rs"): amplitude of response corresponding to smoothed BM impedance function of Fig. 4 (the phase curve of this response is not shown, it overlaps the original phase curve).

we will always underestimate the gain. In the future, a novel smoothing routine for the BM impedance should be developed.

Conversely, we may consider *smoothing of the original response*. Observe the difference in appearance of the response curves in Figs. 2(A) and 3. The irregularities in the response function of Fig. 2(A) can be interpreted as being due to measurement errors, and in Fig. 3 it is tried to remove these errors by smoothing the response function. Then, the BM impedance function appears as being smoothed too. The way smoothing operates in this case can be derived from the reasoning in Sec. V, and it does not seem to contain unexpected pitfalls. In fact, the resynthesized response corresponding to the impedance shown in Fig. 3 is identical to the original response.

## VII. CONCLUSIONS AND REASONS FOR CAUTION

Obviously, we may conclude that the main goal of the inverse-solution procedure has been reached: *a three-dimensional model of the cochlea has been constructed which has a response that matches the original response very closely*. In order to simulate the cochlear response for low levels of stimulation, the model should have a locally active BM impedance. It is formally not allowed to apply the inverse method when the cochlea is definitely operating in the nonlinear range, such as is the case with stronger stimuli. However, the EQ-NL theorem (de Boer, 1997b-CONNEX) can be invoked; this states that such an application is justified as long as one operates only on input-output cross-correlation functions for wide-band stimuli. From results at higher stimulation levels it has been found that the degree of activity decreases with increasing stimulus level (de Boer and Nuttall, 1997a-DAM). Further results on the effect of stimulus level will be reported elsewhere. For a post-mortem response a passive model is sufficient (Sec. II). All these conclusions are correct, of course, within the class of models to which the EQ-NL theorem applies.

In Sec. V we found an argument to consider the inverse solution as "robust." However, results of inverse solution or resynthesis sometimes show unexpected deviations that are difficult to account for. In such cases one has to bear in mind the many simplifications that have been introduced in the procedure. The following items might be mentioned in this respect:

- (1) Long-wave behavior is assumed at and near ( $x=0$ ).
- (2) The response phase is fixed at  $-\frac{1}{2}\pi$  at ( $x=0$ ) which implies that the BM impedance at that point should be purely imaginary.
- (3) The model has a homogeneous cross section and the BM a homogeneous mode of vibration. The taper of cross-sectional area and BM width has not been taken into account (cf. Shera and Zweig, 1991).
- (4) A regular cochlear map is assumed.
- (5) The influence of errors in the input data may turn out appreciable in particular cases.
- (6) Unexpected internal reflections may be revealed, espe-

cially at locations where the real part of the BM impedance crosses the zero line very rapidly. In such a case  $N$  may have been chosen too small.

- (7) The stapes is driven as a velocity source and the coupling with the middle ear has not been taken into account.

It should be noted that items (1), (2), (3), (4), (5), and (7) have a direct bearing on the value of the virtual stapes velocity. In view of these uncertainties we are not surprised that this value can markedly deviate from (1) (other sources of error might reside in the mode of vibration of the stapes and the location of the bead on the footplate).

One question remains: Can, on the basis of the recovered BM impedance, the model be brought into a form where each component represents a realistic part of the actual cochlear mechanism? At present it is impossible to answer this question. Specific problems in cochlear modeling [like that about the causality of  $Z_{\text{BM}}(x, \omega)$ , see de Boer and Nuttall (1997a–DAM)] have to be solved first. The inverse-solution method is often called an “ill-posed problem” because small variations in the input data may cause large variations in the BM impedance. We have shown that this is not necessarily always true.

## ACKNOWLEDGMENTS

The authors are indebted to Jont Allen and Christopher Shera (USA) and to Fabio Mammano and Renato Nobili (Italy) for their contributions, which included generous help and instruction in programming model solutions in MATLAB<sup>®</sup>. Thanks are due to (in alphabetical order) Gary Dootz, Meng He Guo, Bob Masta, Darren Miller, Irina Omelchenko, and Edward Porsov (USA) for their help in performing the experiments. Discussions with Jan van Dijk (The Netherlands) are gratefully acknowledged, they led to better understanding of the mechanism involved in the Green’s function method. Critical remarks by three anonymous JASA reviewers have led to considerable improvements of the text and the figures. The work has been supported by NIH Contract No. NIDCD-DC-00141, U.S. Dept. of Education Contract No. H 235K3001, and NWO Contract No. SLW 01.011.

## APPENDIX A: THREE-DIMENSIONAL MODEL

In a three-dimensional model of the cochlea the BM forms a part, the movable part, of the partition that divides the model into an upper and a lower channel. The remainder of the partition is stationary. We assume that over its own width the BM moves as a half-period of a sine function (as if it is hinged at its inner and outer boundaries). We will use the following symbols:  $x$  is the longitudinal coordinate,  $\omega$  the radian frequency,  $\rho$  the density of the fluid,  $v_{\text{BM}}(x, \omega)$  the BM velocity, and  $Z_{\text{BM}}(x, \omega)$ <sup>5</sup> is the BM impedance. Height and width of the model are  $h$  and  $b$ , respectively. The parameter  $\varepsilon$  is the width of the BM divided by the width  $b$  of the model. In the general theory of long-wave as well as two- and three-dimensional models of the cochlea, the hydrodynamics of the fluid can be described by the function  $Q(k)$ , where the (real) variable  $k$  is the wave number (de Boer, 1981, 1991, Section 2.3; 1996, Sections 5.1–5.3). For this

concept to be definable it is necessary that the model is of *uniform geometry* along its entire (infinite) length and that the transversal mode of vibration of the BM—as described above—is also uniform over the length. Therefore, we only consider a *stylized* or *homogeneous* model in which  $h$ ,  $b$  and  $\varepsilon$  are the same for all  $x$ . The function  $Q(k)$  takes into account three-dimensional effects, (i) that the BM is moving in a particular mode over only a fraction ( $\varepsilon$ ) of its width, (ii) that fluid can move in three dimensions, and (iii) that long as well as short waves are possible. Note that  $Q(k)$  describes *only* properties of the fluid (cf. de Boer, 1997a–CLASS). The model is made of finite length by prescribing appropriate boundary conditions at two locations, ( $x=0$ ) and ( $x=L$ );  $L$  then becomes the length of the model. This technique of handling a finite-length model has been introduced by Siebert (1974). We will refer to this method as the “ $k$ -domain method.” An improved method for forward and inverse solutions for this model has been described in (de Boer, 1998–NEW3D).

To prepare for greater generality we have reformulated the solution in the way described by Allen (1977), Allen and Sondhi (1979), and Mammano and Nobili (1993). First, divide the  $x$ -axis into  $N$  discrete points  $x_i$  ( $i=1, \dots, N$ ). The pressure in the fluid  $p(x, \omega)$ —close to the BM in the upper channel—and the BM velocity  $v_{\text{BM}}(x, \omega)$  are represented by column vectors  $\mathbf{p}$  and  $\mathbf{v}$ , both of length  $N$ . The *hydrodynamics of the fluid* in the model is described by a matrix  $\mathbf{G}$  of size  $N \times N$  which represents the Green’s function. The pressure  $\mathbf{p}$  can then be expressed as:

$$\mathbf{p} = i\omega\rho(\mathbf{G}\mathbf{v} + s v_{\text{st}}). \quad (\text{A1})$$

Here  $\mathbf{s}$  is a column vector which represents the way the stapes boundary condition expresses itself in the pressure  $p(x, \omega)$ , and  $v_{\text{st}}$  is the stapes velocity. For a model of given geometry the matrix  $\mathbf{G}$  and the “stapes propagator”  $\mathbf{s}$  are known. In this approach,  $\mathbf{G}$  and  $\mathbf{s}$  express the hydrodynamics of the fluid constrained as it is by the geometry of the model [Eq. (A1) contains the factor  $i\omega\rho$ ]. Furthermore,  $\mathbf{G}$  and  $\mathbf{s}$  incorporate the boundary conditions mentioned above in connection with the  $k$ -domain method. Assume further that pressure and velocity are related via the BM impedance  $Z_{\text{BM}}(x, \omega)$ . Write this relation as<sup>6</sup>

$$\mathbf{p} = -\frac{1}{2}\mathbf{Z}\mathbf{v}, \quad (\text{A2})$$

where  $\mathbf{Z}$  is a diagonal  $N \times N$  matrix which has  $Z_{\text{BM}}(x_i, \omega)$  ( $i=1, \dots, N$ ) in its main diagonal. Substitute Eq. (A2) in Eq. (A1):

$$(i\omega\rho\mathbf{G} + \frac{1}{2}\mathbf{Z})\mathbf{v} = -i\omega\rho s v_{\text{st}}. \quad (\text{A3})$$

In the *forward* solution Eq. (A3) is solved for  $\mathbf{v}$  with given  $\mathbf{Z}$  and  $v_{\text{st}}$ . In the *inverse* solution the pressure  $\mathbf{p}$  is computed directly from Eq. (A1),  $\mathbf{v}$  and  $v_{\text{st}}$  being given. Next, Eq. (A2) is applied to find the impedance. By the nature of the problem formulation the matrix  $\mathbf{G}$  is a “full” matrix. We have derived  $\mathbf{G}$  and  $\mathbf{s}$  in the manner described by Mammano and Nobili (1993) but have given our “stylized” model a constant width  $b$ , a constant BM width  $\varepsilon b$  ( $\varepsilon=0.2$ ) and a constant height  $h$ . For simplicity, we have made  $b$  equal to  $h$ .



It has been verified that solutions computed with the  $k$ -domain method lead to results that are almost indistinguishable from those obtained by the full-matrix method. Because the  $k$ -domain method involves a “sparse” matrix, it is considerably faster than the full-matrix method. Furthermore, it allows a larger number of divisions of the abscissa. The computations for this paper have been made by the full-matrix method for the stylized three-dimensional model as described above. The main findings were verified by the  $k$ -domain method, with the length divided into 700 sections.

A much simpler solution exists in the *long-wave* approximation (see de Boer, 1996, Sections 4.1–4.5). In this case all operations are carried out with a second-order differential equation:

$$\frac{d^2}{dx^2} p(x, \omega) = \frac{-i\omega\rho}{h_{\text{eff}}} v_{\text{BM}}(x, \omega). \quad (\text{A4})$$

The parameter  $h_{\text{eff}}$  is equal to  $h/\varepsilon$ . This equation means that the inverse solution is obtained via double integration of the (given) BM velocity function. In this case two integration constants appear and have to be handled separately (see Sec. II).

## APPENDIX B: MEASUREMENT—DATA PROCESSING

Data were collected as records of the average BM velocity  $v_{\text{BM}}(x, t)$ , the clock sampling frequency was 208 kHz and one period of the pseudo-random noise signal consisted of 4096 data, thus producing a period duration close to 20 ms. The noise was presented continuously. Averaging was done over more than 3000 noise periods. The BM velocity data were used in the form of input–output ccf spectra. See for further information de Boer and Nuttall (1997b–GLIDE, Appendix B).

The response of the stapes was measured using a noise signal with a bandwidth of 6 or 8 octaves. A collection of raw stapes–response data is shown in Fig. B1. The abscissa covers the frequency range from 5 to 24 kHz. The solid lines show the amplitude functions (the noise stimulus has constant spectral density and random phase) for nine experiments. It is attempted to label the individual curves but this has only been partly successful. To avoid further clogging of the figure, the phase is shown for only one experiment (7611). The dashed line shows the phase function of the input–output ccf spectrum (the raw data have random phases), and the curve is displaced by an integer number of  $2\pi$  rad (see legend). In the region from 10 to 20 kHz these phase data indicate an average time delay of 175  $\mu\text{s}$ . We stress that no windowing has been applied to the stapes data shown (apart from averaging over at least 3000 noise periods). For further analysis the input–output ccf spectrum of the stapes data was used, this was processed in the same way as the BM data. Finally, the ccf spectrum of the BM response was divided by that of the stapes response and corrected for the difference in stimulus spectral density so that the data all represent the ratio of BM-to-stapes response.

The stimulus spectrum contains 200 components in the octave from 10 to 20 kHz (the period of the pseudo-random noise signal is 20 ms). The average spectral value, approxi-

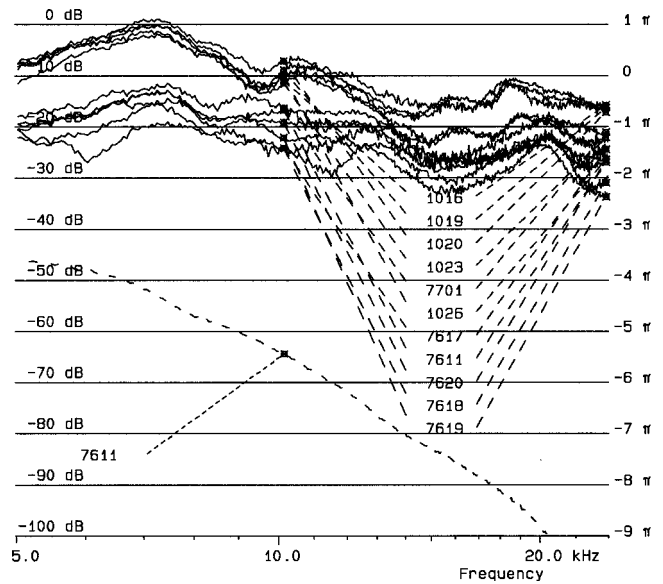


FIG. B1. Velocity response of the stapes in 11 experiments. Stimulus level 100 dB (referred to a bandwidth of 1 octave). 0 dB corresponds to 0.781 mm/s taken over a 50 Hz bandwidth. The phase is shown for one experiment (7611), phase unwrapping started at 5 kHz, and the phase is shown displaced by  $-4\pi$  rad. The phase curves of other experiments are almost identical.

mately  $-22$  dB in the figure, corresponds to the velocity 0.0620 mm/s taken over a 50 Hz bandwidth (0 dB corresponds to 0.781 mm/s). If, instead of a noise signal, a sinusoid in the 15 kHz region were presented at 100 dB SPL, the average stapes velocity would be  $\sqrt{200} \cdot 0.0620$  mm/s = 0.877 mm/s. This is the value we use in Sec. III to estimate the BM velocity.

In the processing of the raw ccf spectra the following modifications have been introduced. First, frequency components above 20 kHz and below 1 kHz were reduced in magnitude (see Appendix B of de Boer and Nuttall, 1997b–GLIDE). The resulting spectrum was (inversely) Fourier transformed and the smoothed ccf waveform was windowed in the  $\tau$  domain, i.e., when  $\tau_0$  is the point of maximum amplitude, the amplitude was gradually and smoothly reduced from the point  $4\tau_0$  on, to reach zero at eight times  $\tau_0$  and to remain zero for larger values of  $\tau$ . After another (forward) Fourier transformation, frequency components above 20 kHz and below 1 kHz were reduced again. As a result, both the spectrum and the waveform have become reasonably smooth functions of their variables. The extra smoothing of the composite ccf spectrum between Figs. 2 and 3 was done by way of windowing in the time domain, this time with a Gaussian weighing function  $\exp(-\frac{1}{2}\tau^2/\tau_1^2)$ ; the parameter  $\tau_1$  is given in the legend to Fig. 3.

In Fig. 4 the BM impedance function was smoothed by spectral windowing. First, the function was inverted to put more emphasis on the peak region where the impedance magnitude is small. The original function, an array of  $N$  complex values, was extended to a length of 4096 values and Fourier transformed. The amplitude at high values of  $k$  was reduced by a Gaussian function with a standard deviation of 720 (this corresponds to 75  $\text{cm}^{-1}$ ) and an inverse Fourier transformation was executed. The resulting function was in-

verted again. The first  $N$  points of the output array contain the smoothed BM impedance function.

## APPENDIX C: CORRECTIONS FOR PHASE AND AMPLITUDE

The transformation of the response function from the frequency to the  $x$ -domain is carried out with a regular ‘‘cochlear map.’’ Assume (radian) frequency  $\omega$  and location  $x$  to be related as:

$$\omega = \omega_{\max} \exp(-\frac{1}{2}\alpha x), \quad (\text{C1})$$

where  $\omega_{\max}$  is the resonance radian frequency at the location of the stapes ( $x=0$ ) and  $\alpha$  is a constant. Underlying this concept is the assumption that, as  $x$  varies, the BM stiffness varies exponentially with the space constant  $\alpha$  and the BM mass stays constant. In the transformation the best frequency (BF) was projected to the location  $x_{\text{BF}}$  corresponding to it, *lower* frequencies were scaled to the *left* and *higher* frequencies to the *right* of it.

When relation Eq. (C1) is used for the transformation, the phase of the response at ( $x=0$ ) will be that of the original frequency response at the frequency corresponding to ( $x=0$ )—for Fig. 3 this frequency is 4.68 kHz. It will not have the value required at the stapes location of the model. In the long-wave approximation the phase at ( $x=0$ ) must be  $-\frac{1}{2}\pi$  radians if the BM impedance is purely imaginary at that point. In Diependaal *et al.* (1987) it is derived how to correct for this defect in a short-wave model. For the long-wave case the same type of reasoning can be applied, and it leads to a slightly different relation. Assume that the model is stiffness controlled, and let the BM impedance vary with  $x$  as

$$Z_{\text{BM}}(x, \omega) = \frac{C_0}{i\omega} \exp(-\alpha x), \quad (\text{C2})$$

where  $C_0$  is the stiffness at ( $x=0$ ). In the long-wave model the local wave number  $k(x)$  then is given by

$$k(x) = \omega \sqrt{\frac{2\rho}{h_{\text{eff}} C_0}} \exp(-\frac{1}{2}\alpha x), \quad (\text{C3})$$

where  $h_{\text{eff}}$  is the effective height of the model ( $h_{\text{eff}}=h/\varepsilon$ )—see Appendix A. The accumulated phase from ( $x=0$ ) to location  $x$  is

$$\varphi(x, \omega) = \varphi_0 + \omega \sqrt{\frac{2\rho}{h_{\text{eff}} C_0}} \frac{2}{\alpha} [\exp(-\frac{1}{2}\alpha x) - 1]. \quad (\text{C4})$$

We equate the integration constant  $\varphi_0$  to  $-\frac{1}{2}\pi$ . The data yield  $\varphi(x_0, \omega)$  as a function of  $\omega$  for fixed location  $x_0$  and we desire to know  $\varphi(x, \omega_{\text{BF}})$  as a function of  $x$ , where  $\omega_{\text{BF}}$  is  $2\pi$  times BF. Abbreviating  $\varphi(x, \omega_{\text{BF}})$  by  $\varphi(x)$  and  $\varphi(x_0, \omega)$  by  $\varphi(\omega)$  we find from Eq. (C4) the following relation:

$$\varphi(x) + \frac{1}{2}\pi = [\varphi(\omega) + \frac{1}{2}\pi] \frac{\omega_{\text{BF}}}{\omega} \frac{(e^{\frac{1}{2}\alpha x} - 1)}{(e^{\frac{1}{2}\alpha x_0} - 1)} \quad (\text{C5})$$

for computing  $\varphi(x)$  from the measured phase  $\varphi(\omega)$ , where  $x_0$  is the location on the  $x$  scale corresponding to the BF. To apply this relation the phase  $\varphi(\omega)$  should first be unwrapped.

Furthermore, it should be checked (afterward) that the BM impedance is nearly purely imaginary at ( $x=0$ ).

One additional step has to be taken. The amplitude must be scaled because in the measurements the BM velocity at the stapes is (for low frequencies) proportional to frequency, and in the model it should be (near the stapes) proportional to  $\exp(\frac{3}{4}\alpha x)$ . The amplitude scaling needed to correct for this is accurately described in (INV-2), and it amounts to the amplitude of the response being multiplied by  $(\omega/\omega_{\text{BF}})^{1/2}$  before it is transformed to the  $x$ -domain. This correction is also based on the long-wave approximation.

<sup>1</sup>One not-directly-obvious implication is that in the experiments spontaneous otoacoustic emissions are assumed to be either suppressed or not to affect the response of the cochlea.

<sup>2</sup>The value of  $\alpha$  was selected to give agreement with Greenwood’s (1990) mapping function for the guinea pig in the appropriate frequency range. The value of  $\omega_{\max}$  was chosen higher than required in order to create a larger region near the stapes location where the long-wave approximation can be applied. The length  $L$  was chosen long enough to accommodate the best frequencies (BFs) involved in the experiments and not much more.

<sup>3</sup>The number  $N$  of sections could be increased to 512 without much effect.

<sup>4</sup>This is why, with a more ‘‘primitive’’ inverse-solution technique, ‘‘activity’’ was always correctly predicted and resynthesis invariably yielded an acceptable match in the peak region (de Boer, 1993, 1995b, 1996).

<sup>5</sup>In a ‘‘classical’’ cochlear model  $Z_{\text{BM}}(x, \omega)$  must be a realizable point-impedance function. In a ‘‘nonclassical’’ model the *effective BM impedance*  $Z_{\text{BM}}^{\text{eff}}(x, \omega)$  can be found after the response has been computed. Because a classical model with its impedance  $Z_{\text{BM}}(x, \omega)$  equal to  $Z_{\text{BM}}^{\text{eff}}(x, \omega)$  has the same response (de Boer, 1997a—CLASS), we will not distinguish between  $Z_{\text{BM}}^{\text{eff}}(x, \omega)$  and  $Z_{\text{BM}}(x, \omega)$  in this paper.

<sup>6</sup>The minus sign and the factor  $\frac{1}{2}$  are due to sign conventions. Upward movement of the BM is counted positive,  $\mathbf{p}$  is the pressure in the upper channel and  $-\mathbf{p}$  in the lower channel. Velocity and pressure should be interpreted as averages over the width of the BM.

Allen, J. B. (1977). ‘‘Two-dimensional cochlear fluid model: New results,’’ *J. Acoust. Soc. Am.* **61**, 110–119.

Allen, J. B., and Sondhi, M. M. (1979). ‘‘Cochlear macromechanics—time domain solutions,’’ *J. Acoust. Soc. Am.* **66**, 123–132.

de Boer, E. (1981). ‘‘Short waves in three-dimensional cochlea models: Solution for a ‘block’ model,’’ *Hearing Res.* **4**, 53–77.

de Boer, E. (1983). ‘‘No sharpening? A challenge for cochlear mechanics,’’ *J. Acoust. Soc. Am.* **73**, 567–573.

de Boer, E. (1991). ‘‘Auditory physics. Physical principles in hearing theory. III,’’ *Phys. Rep.* **203**, 127–229.

de Boer, E. (1993). ‘‘Some like it active,’’ in *Biophysics of Hair Cell Sensory Systems*, edited by H. Duifhuis, J. W. Horst, P. van Dijk, and S. M. van Netten (World Scientific, Singapore), pp. 3–21.

de Boer, E. (1995a). ‘‘The ‘inverse problem’ solved for a three-dimensional model of the cochlea. I. Analysis,’’ *J. Acoust. Soc. Am.* **98**, 896–903 (INV-1).

de Boer, E. (1995b). ‘‘The ‘inverse problem’ solved for a three-dimensional model of the cochlea. II. Application to experimental data sets,’’ *J. Acoust. Soc. Am.* **98**, 904–910 (INV-2).

de Boer, E. (1996). ‘‘Mechanics of the cochlea: Modeling efforts,’’ in *The Cochlea*, edited by P. Dallos, A. N. Popper, and R. R. Fay (Springer-Verlag, New York), pp. 258–317.

de Boer, E. (1997a—CLASS). ‘‘Classical and nonclassical models of the cochlea,’’ *J. Acoust. Soc. Am.* **101**, 2148–2150.

de Boer, E. (1997b—CONNEX). ‘‘Connecting frequency selectivity and nonlinearity for models of the cochlea,’’ *Aud. Neurosci.* **3**, 377–388.

de Boer, E. (1998—NEW3D). ‘‘A method for forward and inverse solutions of a three-dimensional model of the cochlea,’’ *J. Acoust. Soc. Am.* **103**, 3725–3728.

de Boer, E., and van Bienema, E. (1982). ‘‘Solving cochlear mechanics problems with higher-order differential equations,’’ *J. Acoust. Soc. Am.* **72**, 1427–1434.

de Boer, E., and Nuttall, A. L. (1997a—DAM). ‘‘On cochlear cross-correlation functions: connecting nonlinearity and ‘activity’,’’ in *Diversity*

- in *Auditory Mechanisms*, edited by E. R. Lewis, G. R. Long, R. F. Lyon, P. M. Narins, C. R. Steele, and E. Hecht-Poinar (World Scientific, Singapore), pp. 291–297.
- de Boer, E., and Nuttall, A. L. (1997b)–GLIDE). “The mechanical waveform of the basilar membrane. I. Frequency modulations (“glides”) in impulse responses and cross-correlation functions,” *J. Acoust. Soc. Am.* **101**, 3583–3592.
- Diependaal, R. J., de Boer, E., and Viergever, M. A. (1987). “Cochlear power flux as an indicator of mechanical activity,” *J. Acoust. Soc. Am.* **81**, 184–186.
- Greenwood, D. D. (1990). “A cochlear frequency-position function for several species—29 years later,” *J. Acoust. Soc. Am.* **87**, 2592–2605.
- Lighthill, M. J. (1981). “Energy flow in the cochlea,” *J. Fluid Mech.* **106**, 149–213.
- Mammano, F., and Nobili, R. (1993). “Biophysics of the cochlea: Linear approximation,” *J. Acoust. Soc. Am.* **93**, 3320–3332.
- Neely, S. T., and Kim, D. O. (1986). “A model for active elements in cochlear biomechanics,” *J. Acoust. Soc. Am.* **79**, 1472–1480.
- Nuttall, A. L., and Dolan, D. F. (1996). “Steady-state sinusoidal velocity responses of the basilar membrane in guinea pig,” *J. Acoust. Soc. Am.* **99**, 1556–1565.
- Nuttall, A. L., Dolan, D. F., and Avinash, G. (1990). “Measurements of basilar membrane tuning and distortion with laser Doppler velocimetry,” in *The Mechanics and Biophysics of Hearing*, edited by P. Dallos, C. D. Geisler, J. W. Matthews, M. A. Ruggero, and C. R. Steele (Springer-Verlag, Berlin), pp. 288–295.
- Patuzzi, R. (1996). “Cochlear micromechanics and macromechanics,” in *The Cochlea*, edited by P. Dallos, A. N. Popper, and R. R. Fay (Springer-Verlag, New York), pp. 186–257.
- Robles, L., Ruggero, M. A., and Rich, N. C. (1986). “Basilar membrane mechanics at the base of the chinchilla cochlea. I. Input–output functions, tuning curves, and response phases,” *J. Acoust. Soc. Am.* **80**, 1364–1374.
- Ruggero, M. A., Rich, N. C., Recio, A., Narayan, S. S., and Robles, R. (1997). “Basilar-membrane responses to tones at the base of the chinchilla cochlea,” *J. Acoust. Soc. Am.* **101**, 2151–2163.
- Sellick, P. M., Yates, G. K., and Patuzzi, R. (1983). “The influence of Mössbauer source size and position on phase and amplitude measurements of the guinea pig basilar membrane,” *Hearing Res.* **10**, 101–108.
- Siebert, W. M. (1974). “Ranke revisited—A simple short-wave cochlear model,” *J. Acoust. Soc. Am.* **56**, 594–600.
- Shera, C. A., and Zweig, G. (1991). “A symmetry suppresses the cochlear catastrophe,” *J. Acoust. Soc. Am.* **89**, 1276–1289.
- Sondhi, M. M. (1978). “Method for computing motion in a two-dimensional cochlear model,” *J. Acoust. Soc. Am.* **63**, 1468–1477.
- Tsuji, J., and Liberman, M. C. (1997). “Intracellular labeling of auditory nerve fibers in the guinea pig: Central and peripheral connections,” *J. Comp. Neurol.* **381**, 188–202.

# Dolphin hearing: Relative sensitivity as a function of point of application of a contact sound source in the jaw and head region

B. Møhl

*Department of Zoophysiology, Institute of Biological Sciences, University of Aarhus, DK-8000 Aarhus C, Denmark*

W. W. L. Au, J. Pawloski, and P. E. Nachtigall

*Hawaii Institute of Marine Biology, University of Hawaii, P.O. Box 1106, Kailua, Hawaii 96734*

(Received 20 April 1998; revised 16 February 1999; accepted 19 March 1999)

The auditory input area of the dolphin head was investigated in an unrestrained animal trained to beach itself and to accept noninvasive electroencephalograph (EEG) electrodes for the recording of the auditory brain-stem response (ABR). The stimulus was a synthetic dolphin click, transmitted from a piezo-electric transducer and coupled to the skin via a small volume of water. The results conform with earlier experiments on acute preparations that show best auditory sensitivity at the middle of the lower jaw. Minimum latency was found at the rear of the lower jaw. A shaded receiver configuration for the dolphin ear is proposed. © 1999 Acoustical Society of America.

[S0001-4966(99)05606-4]

PACS numbers: 43.64.Tk, 43.64.Ri [RDF]

## INTRODUCTION

Dolphins are acoustically orienting mammals, adapted to the marine environment. Profound changes were required to transform an efficient ear for aerial sound, as present in the terrestrial ancestors of cetaceans, into one of equal or greater efficiency underwater. Mammalian soft tissues behave much like water acoustically. Consequently, underwater sound is not being reflected at the boundary of a body and is not easily channeled, as it is in, e.g., the auditory meatus of terrestrial mammals. In fact, a submerged mammalian body is quite transparent to sound (cf. ultrasonic scanning), with exceptions for volumes containing gas, or dense bone.

From anatomical observations, it is therefore difficult to identify a route of sound from the water medium to the inner ear in cetaceans. A variety of theories has been proposed, spanning from the completely sound-transparent head (Reyssenbach de Haan, 1957), over meatal transmission (Purves, 1966) to the lower-jaw acoustic window and waveguide idea (Norris, 1964, 1980). Even hypotheses dispensing with the inner ear as the mechano-neuronal transducer have been proposed (Goodson and Klinowska, 1990).

Since the nature, shape, and size of the receiver are major determinants of the performance of any sonar, it is not surprising that a number of experiments have been designed to provide evidence of this matter. An early, classical paper by Bullock *et al.* (1968) used acoustically evoked potentials, recorded from the midbrain of a number of dolphins in acute preparations. Best sensitivity was found to sound applied to the front half of the lower jaw, in fair agreement with the jaw-hearing hypothesis of Norris (1964). Using cochlear microphonics in similar preparations, McCormick *et al.* (1970) obtained similar results. The jaw-hearing hypothesis was also supported by a psychophysical experiment by Brill *et al.* (1988). In this experiment, the target-detection ability of a trained dolphin, with and without sound-blocking material

applied to the mandible, was found to be influenced by the blocking. Recently, Popov and Supin (1990) used a sound source in the far field and the auditory brain-stem response (ABR) with a time-delay, difference-based triangulation technique to determine the auditory input location of three dolphin species. This technique dispenses with surgical procedures, allowing physiological measurements from a lightly restrained, alert animal. They concluded that sound entered the ear in the meatal/bulla region and not via the mandible.

From knowledge on the directionality of the dolphin ear (Au and Moore, 1984) it can be inferred that the effective area of sound collection, the equivalent of the pinna of terrestrial mammals, is about 8 cm<sup>2</sup>. The theories of the fully transparent head, and the meatal route, both appear inconsistent with this result. However, as the effective area consists of the sum of areas shaded in an unknown way, the directional evidence is not suitable for testing the theories.

Thus, the interpretation of the various experiments is not consistent, and even contradictory (Norris, 1964; Popov and Supin, 1990). Different approaches ranging from measurements on cadavers, over acute, sedated preparations, mildly restrained alert animals, and to fully cooperating, trained animals have been used. Some experiments were performed in air, others with submerged animals, some with far-field stimulation, others with point or contact stimulation. These varied approaches simply do not produce readily comparable results. The auditory input area in dolphins thus continues to be incompletely defined.

The approach of this work was to use a new combination of the above approaches. Stimuli were applied by a contact hydrophone in a suction cup to a female dolphin (*Tursiops truncatus*) trained to beach itself on a mat and remain motionless for minutes. Employing noninvasive electrodes, the ABR technique was used to quantify the response. The aim was to define the zone of entrance for sound by combin-

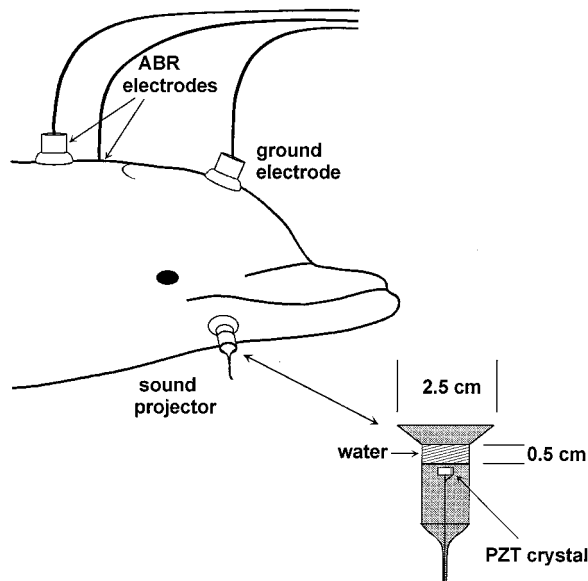


FIG. 1. Sketch of recording geometry electrodes, and of the sound projector.

ing knowledge of place-of-point stimulation with amplitude and latency of response.

## I. MATERIAL AND METHODS

An Atlantic bottle-nosed dolphin (*Tursiops truncatus*, adult female, named Kolohe) with no suspected auditory deficiencies, was trained at the Hawaii Institute of Marine Biology facility, Coconut Island, Hawaii to beach itself on a foam mattress, and to accept application of the sound transducer and three skin electrodes. A series of white dots (zinc ointment), spaced 10 cm apart, was applied for each session along a line from the right flipper to the tip of the lower jaw. The reference point was 5 cm below the angle of the mouth. The dots served as landmarks for positioning the sound transducer (see Fig. 1).

The transducer was made from a disc of PZT piezoelectric ceramics (diameter 6.3 mm, thickness 2 mm), backed by corprene. This assembly was placed in a 25-mm-diameter suction cup mold and potted with degassed Uralite 3138 (Hexcel). This compound has a rho-c value (specific acoustic impedance, the product of density and sound velocity) close to that of water (Moffett *et al.*, 1986). On application, the cavity of the cup was filled with water, creating a sound-conducting bridge between the transducing element and the skin. Only the right side of the animal was stimulated.

The ABR electrodes were modified tin/lead alloy plate-shaped electrodes (Dantec 13L71), mounted in 45-mm-diameter suction cups. They were placed, respectively, on the vertex, below the contralateral meatus, and on the melon, using conductive electrode gel to create contact with the skin.

The ABR signal was amplified differentially and band-pass filtered (0.3 to 3 kHz) by a Grass P15 preamplifier before being led through an additional gainblock and a Precision Filters 88A, 130-dB/octave low-pass filter to a Tucker-Davis 16-bit sampling and averaging device. Sampling rate

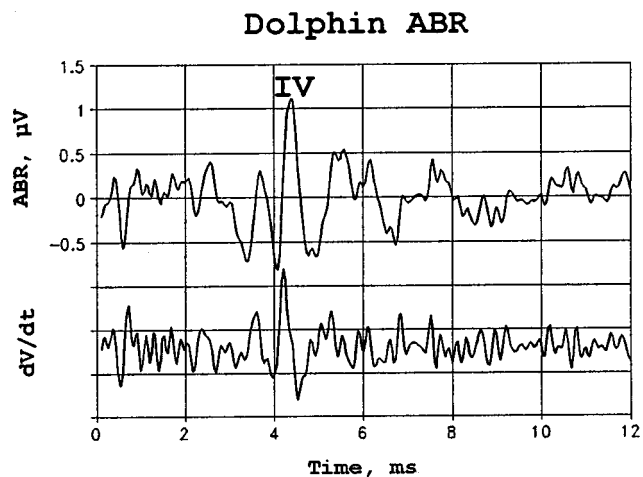


FIG. 2. Example of the ABR signal (top) and its differentiated representation (bottom). Zero on the time axis denotes onset of sound stimulation. Average of 512 stimulations.

was 25 kHz. An on-line artifact-rejecting algorithm was used to prevent signals with excessive noise from entering the average.

The stimulus was a broadband, synthetic, generic porpoise click (Au, 1993), which when filtered by the transmitting response of the transducer had peak energy at 53 kHz and was 3 dB down at 46 and 62 kHz. Repetition rate was 10 clicks per s.

The free-field peak-to-peak level at 1 cm was about 140 dB *re* 1  $\mu\text{Pa/V}$ . This signal was generated by a Datel 1200 card, synchronized with the Tucker-Davis unit. Each trial consisted of either 512 or 1024 stimulations, the responses to which were averaged for a period of 12 ms after stimulus onset.

## II. RESULTS

A typical recorded ABR waveform is shown in Fig. 2. It is similar to that recorded from submerged dolphins by Ridgeway *et al.* (1981), using far-field stimulation. The ABR signal is a compound, multi-peaked signal, generated by auditory neurons in the brainstem, tens of centimeters away from the recording electrodes. The various peaks are believed to reflect the summated activity in the individual neurons of the auditory nerve and brainstem nuclei. The amplitude with our conditions is on the order of microvolts. The most prominent peak (IV in Ridgeway *et al.*, 1981) occurs with a latency of about 4 ms after onset of stimulation.

The input/output functions, when plotted as ABR peak-amplitude versus stimulus intensity, have slopes around 1  $\mu\text{V}$  per 15-dB increase in stimulation. The best dynamic range obtained was 50 dB.

The ABR amplitude depended on point of application. Figure 3 shows a map of interpolated attenuation values required to produce an ABR signal of 1  $\mu\text{V}$ . Best sensitivity was found to sound applied just forward of the pan bone, about 25 cm behind the tip of the lower jaw. Moderate sensitivity was found to sounds applied on the forward portion of the lower jaw and in a posterior region of the pan bone. Stimulation within the mouth, on the internal, lateral side of

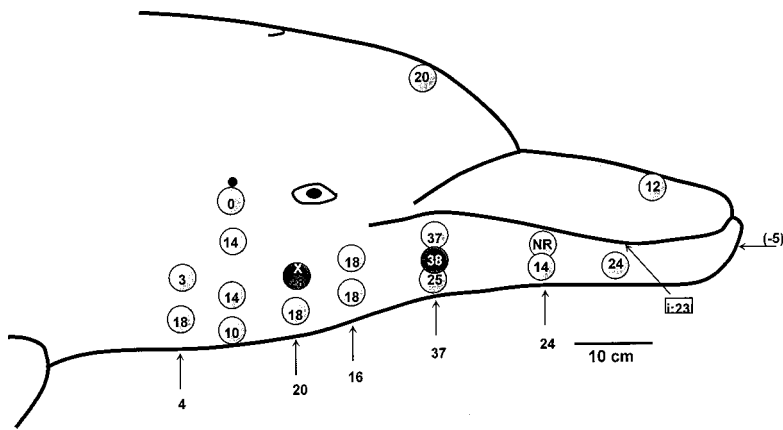


FIG. 3. Points of stimulation (filled circles) with attenuation (dB *re*: fixed reference) required for an ABR-IV yield of  $1 \mu\text{Vpp}$ . The point marked "X" is the position of minimum delay. The point marked "38" (in white) is the point of maximum sensitivity. Numbers below arrows signify attenuation for stimulation on the ventral midline, while the square label shows attenuation for stimulation on the inside of the lower jaw. No ABR could be detected at the point marked "NR."

the jaw about 10 cm caudal of the tip and 2 cm below the row of teeth, showed sensitivity equal to that of the outside.

A latency measure was obtained by differentiation of the ABR signal, establishing the point in time of maximal change in neuronal firing. This point is on the rising front of the most prominent peak (IV). The increment used was the clock period ( $40 \mu\text{s}$ ). The reason for performing this operation is to obtain a measure of latency with a smaller ambiguity than can be obtained from peak determination of the ABR. The sharpening of the response in the time domain of the differentiated waveforms as compared with that of the mother function is illustrated in Fig. 2. Latencies decreased with increased stimulation within a range of 6 to  $8.5 \mu\text{s}$  per dB.

The shortest latency was found for stimulation in the area below the eye. When plotted as relative delay against distance from this point (marked X in Fig. 3), delays are consistently larger than predicted from a linear model of straight-line propagation and the lowest value for velocity of sound propagation in fatty tissues of the dolphin head (Norris and Harvey, 1974).

### III. DISCUSSION

The ABR signal is rather variable in waveform, amplitude, and latency. The sources of this variability are many, and their results are cumulative. Simple stochastic ones are to some extent taken care of by the averaging process, but others such as transient noises, differences in transducer coupling, and muscle-tone changes in the animal, contribute in ways that cannot be evaluated, and thus limit the interpretative strength of the data.

Having the experiment done in air helped to ensure that the stimulus was localized to the spot where the suction cup was applied, and not propagated to the animal via some unknown path. Also, point-contact stimulation of an animal outside the medium for which its peripheral auditory system is adapted is questionable. Is it sound or vibration stimulation? What is the effective stimulation level? And what are the consequences of essentially "loading" the input impedance of the ear with the very low rho-c value of air instead of the high one of water?

The sensitivity distribution of Fig. 3 is not substantially different from that obtained by Bullock *et al.* (1968) in acute preparations with electrodes implanted directly in the brain

stem and stimulation with hand-held transducers, pressed against the surface. Both studies point to the front half of the lower jaw as being the most sensitive area with sound applied via contact transducers. This area appears to be extended somewhat forward relative to the classical acoustic window proposal of Norris, but essentially this study joins those that support the jaw-hearing hypothesis. The best ABR sensitivity was found to sound applied 25 cm behind the tip of the jaw.

The finding of good sensitivity to sound applied to the inner side of the lower jaw is noteworthy for two reasons. One is that it counterindicates a sound-reflective property of the inner side of the jaw. The other reason is that dolphins occasionally are observed to investigate objects by sonar with open mouth. This would indeed seem to require acoustic sensitivity to the inside of the mouth, as confirmed here.

The differences in latency presented in Fig. 4 are likely to be due to differences in sound path, and we thus prefer to use the term delay to describe this effect. The shortest delay is found in an area about 10 cm ventral of the eye, in a region close to, if not identical to, the meatal/bulla region of sound entrance of Popov and Supin (1990). However, this area is at the rear limit of that of best sensitivity. This discrepancy in area location for the two parameters may indicate the presence of a shaded receiving transducer, where input is

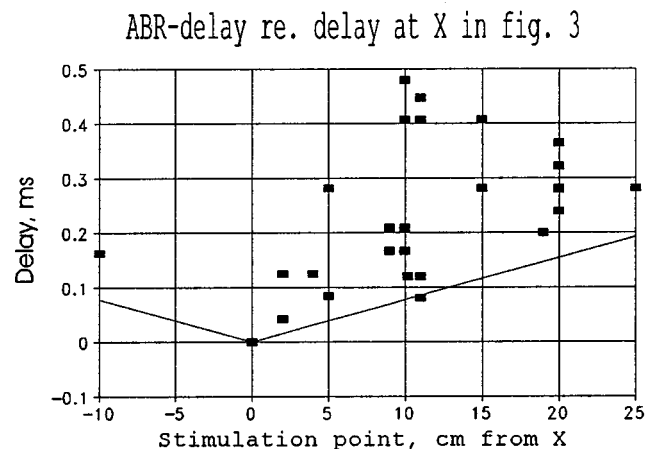


FIG. 4. Delay of the ABR-IV component as a function of distance from point "X" in Fig. 3. Stimulation at fixed reference value. The lines are graphical representations of the hypothesis of delay being a linear function of distance from point X in Fig. 3, using a velocity of sound of 130 cm/ms.

weighted according to direction of arrival and position across the aperture of the receiver. In this way, the directional characteristics of the ear can be shaped. Note that delay is increased for stimulation behind point *D*, indicating sound to be routed forward of this point.

Figure 4 demonstrates that delay is consistently larger than expected assuming a constant, low velocity of propagation along a linear path (the jaw). Part of this may be explained by the reference point (at the surface) being displaced from the common point of entry to the auditory pathway. However, unrealistically large values of this displacement (20 cm) are required to fit the data. Also, the assumed propagation velocity of 130 cm/ms, the lowest found in fatty tissues of the dolphin head by Norris and Harvey (1974), may be off, but again only unrealistically low values (<80 cm/ms) can explain the data from this hypothesis. The excessive delays remain at present unexplained.

When evaluating the models, it should be borne in mind that the cross section of the mandible is on the same order of magnitude as the dominant wavelength of *p*-waves (compressional or longitudinal waves, as opposed to transversal and shear waves) in water and soft tissues at 50 kHz (3 cm). A logical consequence of this observation is that models inspired from optical analogies (reflections, refraction, etc.) are problematic, as they require structures that are considerably larger than the wavelength. However, in mixed media with solid components, other sorts of waves than compressional, longitudinal ones can be realized.

#### IV. CONCLUSION

Good sensitivity to point-applied sound is found in the entire region of the lower jaw, and is best in the middle half of the jaw. The inside of the jaw is as sensitive as is the outside. The question, "Where is sound entering the auditory system of dolphins?" may have to be rephrased into two questions, one dealing with the area of minimum delay and one dealing with the area of minimum attenuation. As these areas seem to differ, a shaded receiver configuration is proposed.

#### ACKNOWLEDGMENTS

This research was jointly supported by Office of Naval Research (Grant No. N00014-95-1-0605) and the Danish National Research Foundation. This is Hawaii Institute of Marine Biology Contribution No. 1041.

- Au, W. W. L. (1993). *The Sonar of Dolphins* (Springer, New York), p. 184.
- Au, W. W. L., and Moore, P. W. B. (1984). "Receiving beam patterns and directivity indices of the bottlenose dolphin *Tursiops truncatus*," *J. Acoust. Soc. Am.* **75**, 255–262.
- Brill, R. L., Sevenich, M. L., Sullivan, T. J., Sustman, J. D., and Witt, R. E. (1988). "Behavioural evidence for hearing through the lower jaw by an echolocating dolphin (*Tursiops truncatus*)," *Mar. Mamm. Science* **4**, 223–230.
- Bullock, T. H. (1968). "Electrophysiological studies of central auditory mechanisms in cetaceans," *Z. F. Vergl. Physiol.* **59**, 117–156.
- Goodson, D., and Klinowska, M. (1990). "A proposed echolocation receptor for the bottlenose dolphin (*Tursiops truncatus*): Modelling the receive directivity from tooth and lower jaw geometry," in *Sensory Abilities of Cetaceans: Laboratory and Field Evidence*, edited by J. A. Thomas and R. A. Kastelein (Plenum, New York), pp. 255–268.
- McCormick, J. G. (1970). "Sound conduction in the dolphin ear," *J. Acoust. Soc. Am.* **48**, 1418–1428.
- Moffett, M. B., Powers, J. M., and McGrath, J. C. (1986). "A rho-c hydrophone," *J. Acoust. Soc. Am.* **80**, 375–381.
- Norris, K. S. (1964). "Some problems of echolocation in cetaceans," in *Marine Bioacoustics*, edited by W. N. Tavolga (Pergamon, New York), pp. 316–336.
- Norris, K. S. (1980). "Peripheral sound processing in odontocetes," in *Animal Sonar Systems*, edited by R. G. Busnel and J. F. Fish (Plenum, New York), pp. 495–509.
- Norris, K. S., and Harvey, G. W. (1974). "Sound transmission in the porpoise head," *J. Acoust. Soc. Am.* **56**, 659–664.
- Popov, V., and Supin, A. (1990). "Localization of the acoustic window at the dolphin's head," in *Sensory Abilities of Cetaceans*, edited by J. A. Thomas and R. A. Kastelein (Plenum, New York), pp. 417–426.
- Purves, P. E. (1966). "The anatomy and function of the outer and middle ear in cetaceans," in *Whales, Dolphins and Porpoises*, edited by K. S. Norris (University of California Press, Berkeley), p. 324.
- Reysenbach de Haan, F. W. (1957). "Hearing in whales," *Acta Oto-Laryngol. Suppl.* **134**, 114.
- Ridgway, S. H., Bullock, T. H., Carder, D. A., Seeley, R. L., Woods, D., and Galambos, R. (1981). "Auditory brainstem response in dolphins," *Proc. Natl. Acad. Sci. USA* **78**, 1943–1947.

# Intensity discrimination of Gaussian-windowed tones: Indications for the shape of the auditory frequency-time window

Nicolle H. van Schijndel, Tammo Houtgast, and Joost M. Festen  
*Department of Otolaryngology, University Hospital VU, P.O. Box 7057,  
1007 MB Amsterdam, The Netherlands*

(Received 22 August 1997; revised 21 September 1998; accepted 23 February 1999)

The just-noticeable difference in intensity  $jnd(I)$  was measured for 1-kHz tones with a Gaussian-shaped envelope as a function of their spectro-temporal shape. The stimuli, with constant energy and a constant product of bandwidth and duration, ranged from a long-duration narrow-band “tone” to a short-duration broadband “click.” The  $jnd(I)$  was measured in three normal-hearing listeners at sensation levels of 0, 10, 20, and 30 dB in 35 dB(A) SPL pink noise. At intermediate sensation levels,  $jnd(I)$  depends on the spectro-temporal shape: at the extreme shapes (tones and clicks), intensity discrimination performance is best, whereas at intermediate shapes the  $jnd(I)$  is larger. Similar results are observed at a higher overall sound level, and at a higher carrier frequency. The maximum  $jnd(I)$  is observed for stimuli with an effective bandwidth of about 1/3 octave and an effective duration of 4 ms at 1 kHz (1 ms at 4 kHz). A generalized multiple-window model is proposed that assumes that the spectro-temporal domain is partitioned into “internal” auditory frequency-time windows. The model predicts that intensity discrimination thresholds depend upon the number of windows excited by a signal:  $jnd(I)$  is largest for stimuli covering one window.  
© 1999 Acoustical Society of America. [S0001-4966(99)00306-9]

PACS numbers: 43.66.Ba, 43.66.Fe [RVS]

## INTRODUCTION

This study addresses a fundamental psychoacoustical question: how does the auditory system extract spectro-temporal information while processing complex sounds? To obtain information about both the spectral and the temporal structure of a signal, the auditory system performs a frequency-time (f-t) analysis. The result of an f-t analysis is a spectrogram, showing the distribution of signal energy over frequency and time. In the spectrogram, the signal energy at a given point is determined by an integration over a specific frequency-time window. The shape of this f-t analysis window determines which characteristics of the sound are displayed. It is commonly assumed that the peripheral auditory system carries out an f-t analysis with its own specific f-t window. This study seeks to characterize the auditory f-t window.

An f-t analysis window cannot be restricted arbitrarily both in time and in frequency. The minimum area of an f-t window is unity if this area is defined as the product of the effective bandwidth and the effective duration (see Stewart, 1931; Gabor, 1947). The lower bound is attained by the Gaussian function (see Gabor, 1947). Given this restriction, the shape of an f-t analysis window can vary. Different f-t analysis windows will display different details in the f-t representation.

In this study we use a psychophysical approach to gain an insight into the shape of the f-t window underlying auditory sound analysis. Before explaining the experimental paradigm, we will briefly review some current ideas on spectral and temporal resolution in the auditory system and their relation to the auditory f-t window.

In psychoacoustics, the partition of the frequency axis into critical bands is a basic concept. Roughly, the auditory

filters have a constant relative bandwidth of 1/3 octave (see, e.g., Scharf and Buus, 1986). This suggests that the spectral width of the auditory f-t window is about 1/3 octave.

In the time domain, however, the picture is less clear. Recall that, for a given bandwidth, the *smallest possible* temporal width is defined by the minimum window area. Thus the temporal width of the auditory f-t window must be at least as long as the minimum temporal width possible, given a specific spectral width. Taking into account psychoacoustical arguments for defining a temporal width, there is no complete consensus. Plack and Moore (1990) discuss the problem of describing the temporal resolution of the auditory system with a single value. They note that the integration time constant appears to decrease with increasing frequency (see also Gerken *et al.*, 1990; Florentine *et al.*, 1988). This suggests that the temporal width of the auditory f-t window decreases with increasing frequency. Viemeister and Wakefield (1991) are interested in the “resolution-integration” paradox: Models describing temporal resolution use short time constants, whereas models describing the improvement in detection and discrimination with increasing signal duration are based on a process of long-term temporal integration. Their conclusions favor the short time constants (roughly 3 ms for 1-kHz sinusoids). Although temporal integration data (time-intensity trade in detection) can be explained easily by an energy-detector model (single look) with an adjustable time window matched to the signal duration (see, for example, Dai and Wright, 1995), the multiple-look model of Viemeister and Wakefield with short time constants can account for both the data on temporal resolution and temporal integration. In general, temporal resolution experiments suggest that the temporal width of the auditory f-t window is about 3 ms at 1 kHz and smaller at higher



frequencies. This is very close to the minimum duration possible if the bandwidth of the auditory f-t window is 1/3 octave.

The present research evaluates the auditory f-t window by assessing just-noticeable differences in intensity for stimuli with different spectro-temporal shapes. By varying the spectro-temporal shape, the number of “internal” (auditory) f-t windows excited by the signal can be varied. Our hypothesis is that this variation affects the just-noticeable difference in intensity. The basis for this hypothesis can be found in the existing models with respect to intensity discrimination.

An important model that describes intensity discrimination performance is the so-called multiband excitation-pattern model (see, e.g., Florentine and Buus, 1981; Durlach *et al.*, 1986; Buus, 1990; Buus and Florentine, 1994). This model operates in the spectral domain. The idea is that the excitation pattern induced by the signal is divided into several spectral bands, and the content of each band is processed individually. Information can be combined across bands to come to an overall percept. Alternatively, psychoacoustical data with respect to temporal mechanisms can be accounted for, at least qualitatively, by the multiple-look model (see Florentine, 1986; Viemeister and Wakefield, 1991). The multiple-look model divides the signal into short-duration segments. As in the multiband model, the information in different segments or “looks” is considered statistically independent. A combination of different looks will result in more information and therefore in lower thresholds. Conceptually, of course, the multiband model and the multiple-look model are very similar, one operating in the spectral domain and the other in the temporal domain. Durlach *et al.* (1986) note that the frequency channels in their model for discrimination of broadband signals can refer to time intervals as well.

The rationale of the hypothesis of this study is a generalization of these “multi-channel models” (multiband and multiple-look models), which in this paper will be called the “multiple-window model” in the f-t domain. Our hypothesis is the following: the auditory spectro-temporal domain is partitioned into “internal” auditory frequency-time windows. So, the “internal” f-t representation of a signal can be characterized by the number of f-t windows covered by the signal. As in the multiband excitation-pattern model (Florentine and Buus, 1981), the multiple-window model assumes that the discriminability within a window is independent of excitation level (Weber’s Law). Thus intensity discriminability for a given signal depends on the number of independent auditory f-t windows covered by the signal: the just-noticeable difference in intensity  $jnd(I)$  will be smaller if more windows are involved.

Evaluation of intensity discrimination thresholds for a range of stimuli with well-defined variations in spectro-temporal shape may reveal the shape of the auditory f-t window. By manipulation of the spectro-temporal shape of the stimuli the number of auditory f-t windows covered by the signal can be varied. We are looking for the signal for which the “internal” auditory representation best matches the auditory f-t window. For that purpose we used sinusoids with a

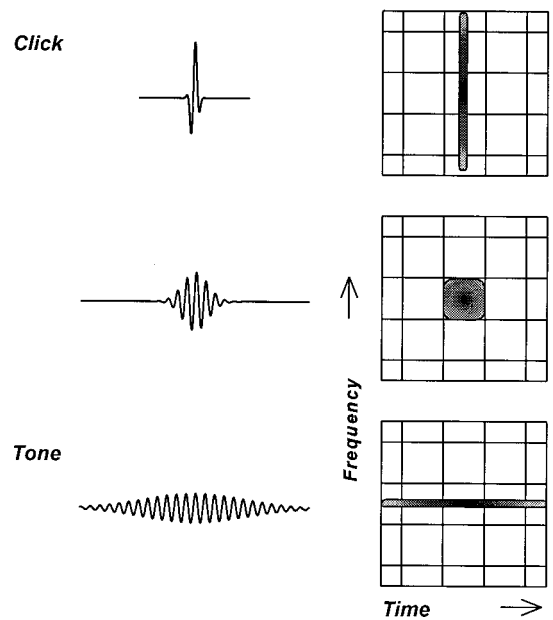


FIG. 1. A schematic representation in the f-t plane of three stimuli with different shape factors on a grid of f-t windows.

Gaussian-shaped temporal envelope. Consequently, it can be shown that the spectral envelope is Gaussian shaped as well. A Gaussian-windowed signal was chosen because of its minimum effective f-t area. Another appealing property is its symmetry in frequency and time. A series of amplitude discrimination experiments was performed for a range of these stimuli in which only one variable was changed, the so-called shape factor, which determines the effective<sup>1</sup> number of periods included under the Gaussian envelope. By varying the shape factor, the representation of the signal in the f-t plane was systematically varied while keeping its area constant, ranging from a long-duration narrow-band “tone” to a short-duration broadband “click” (see Fig. 1). A tone, corresponding to a small shape factor, will excite many f-t windows along the time axis; a click, corresponding to a large shape factor, will excite many f-t windows along the frequency axis; somewhere between tone and click fewer windows will be excited. Thus the number of auditory f-t windows excited by the signal varies as a function of the shape factor according to an U-shaped curve.

The multiple-window idea states that  $jnd(I)$  varies with the number of f-t windows involved in the discrimination task. The more elementary f-t windows which are involved, the smaller the  $jnd(I)$  will be. This implies that the signal with the largest  $jnd(I)$  covers the minimum number of windows. The shape factor corresponding to this signal will be called the “critical” shape factor. Thus the signal with the critical shape factor is most successful in exciting only the minimum number of f-t windows; the “internal” representation of this signal is most closely related to the elementary f-t window. Within the context of the multiple-window idea, the shape of the f-t representation of that signal best matches the shape of the elementary f-t window in (peripheral) auditory coding. The aim of the experiments is to test the multiple-window hypothesis by examining whether  $jnd(I)$  varies with a varying spectro-temporal shape of the stimuli.

If  $\text{jnd}(I)$  varies as a function of shape factor, the critical shape factor gives some insight into the auditory f-t window.

In the first experiment, the relation between intensity discrimination and shape factor was determined for 1-kHz sinusoids at various sensation levels (0, 10, 20, 30 dB SL) in 35-dB(A) SPL pink noise. Low sensation levels were used to avoid spread of excitation as much as possible. In the second experiment, intensity discrimination performance was measured at 4 kHz. Finally, in the third experiment, intensity discrimination performance was measured at a 20-dB higher level for both noise and signal.

## I. METHOD

### A. Stimuli

The stimuli  $s(t)$  consist of Gaussian-windowed tones, defined by

$$s(t) = A \sqrt{\alpha f_0} \sin\left(2\pi f_0 t + \frac{\pi}{4}\right) \exp(-\pi(\alpha f_0 t)^2). \quad (1)$$

These are sinusoids with carrier frequency  $f_0$  and a gradual onset and offset (see Fig. 1). The shape factor  $\alpha$  determines the effective number of sinusoidal periods, equal to  $1/\alpha$ , contained within the Gaussian envelope. If  $\alpha$  is small, the number of periods is large (tone). If  $\alpha$  is large, the number of periods is small (click). Throughout the experiments the independent variable is the shape factor  $\alpha$  (0.0375, 0.075, 0.15, 0.3, 0.6, and 1.2). The effective duration of the Gaussian signal is  $\Delta_t = 1/(\alpha f_0)$ . The effective bandwidth is  $\Delta_f = \alpha f_0$ .

The amplitude of the signal is defined by  $A \sqrt{\alpha f_0}$ . The amplitude difference is produced by increasing the amplitude constant  $A$  from  $A_0$  to  $A_0 + \Delta A$ . By introducing the phase factor  $\pi/4$  the energy of the signal is independent of  $\alpha$  and  $f_0$ . As a result, the total energy  $E$  of the signal is  $(\sqrt{2}/4)A^2$ , only depending on the amplitude constant  $A$ .

As already mentioned in the Introduction, in this study the temporal and spectral domain are investigated in combination. Therefore, the set of stimuli in this study consists of stimuli that cover more critical bands along the frequency axis and only one look (about 3 ms at 1 kHz) along the time axis, stimuli that cover only one critical band and more time looks, and stimuli in between. In Table I, the bandwidth, duration, and effective number of periods of the stimuli used in the experiments can be found. The column labeled “# f-t windows” gives the estimated number of f-t windows covered by the 1-kHz tone, assuming a Gaussian auditory f-t window with a shape factor of 0.23, corresponding to a bandwidth of 1/3 octave and a duration of 4 ms at 1 kHz and a duration of 1 ms at 4 kHz.

### B. Apparatus

Stimuli were generated digitally at a sampling frequency of 40 kHz and were played out over TDT (Tucker Davis Technologies System II) hardware. Because Gaussian-windowed signals do not have compact support,<sup>3</sup> the signals were cutoff at frequencies corresponding to their 60-dB down points. A Wandel und Goltermann RG-1 analog noise

TABLE I. The effective duration  $\Delta_t$ , the effective bandwidth  $\Delta_f$ , the effective number of periods, and the estimated number of f-t windows for stimuli with different shape factor  $\alpha$  and carrier frequency  $f_0$  as used in the experiments.

$\alpha$	$f_0$	$\Delta_t$	$\Delta_f$	# periods	# f-t windows
0.0375	1000 Hz	27 ms	37.5 Hz	27	7
	4000 Hz	6.7 ms	150 Hz	27	7
0.075	1000 Hz	13 ms	75 Hz	13	3
	4000 Hz	3.3 ms	300 Hz	13	3
0.15	1000 Hz	6.7 ms	150 Hz	6.7	2
	4000 Hz	1.7 ms	600 Hz	6.7	2
0.3	1000 Hz	3.3 ms	300 Hz	3.3	1
	4000 Hz	0.83 ms	1200 Hz	3.3	1
0.6	1000 Hz	1.7 ms	600 Hz	1.7	3
	4000 Hz	0.42 ms	2400 Hz	1.7	3
1.2	1000 Hz	0.83 ms	1200 Hz	0.83	6
	4000 Hz	0.21 ms	4800 Hz	0.83	6

generator produced the continuous pink noise. Signals and noise were attenuated (TDT PA4) separately, and subsequently summed (TDT SM3). The stimuli were presented monaurally through Sony MDR-CD999 headphones. Masking noise levels were measured on a Brüel & Kjær type 4152 artificial ear with a flat-plate adapter. The entire experiment was controlled by an IBM PC-compatible computer. Subjects were tested individually in a soundproof room.

### C. Procedure

Intensity discrimination performance was measured using an adaptive, three-interval, three-alternative forced-choice paradigm (3I, 3AFC). Each trial consisted of three observation intervals. The time between the onset of the three stimuli was always 500 ms, but the duration of the stimuli differed with different shape factors. Taking into account the cutoff at 60 dB below the top, the total duration of the longest signal was 80 ms. Two intervals contained the reference signal (with amplitude constant  $A_0$ ) and one interval contained the incremented signal (with amplitude constant  $A_0 + \Delta A$ ). The incremented signal occurred randomly in one of the three observation intervals. Each observation interval was marked by a visual display. The onset of the stimuli coincided with the onset of the display. The noise was presented continuously. The subject’s task was to indicate the interval that contained the incremented signal by pushing the appropriate button on a PC keyboard. There was no response time limit. Immediately after the response, feedback was provided. After the response, 500 ms elapsed before a following trial started.

In obtaining a threshold estimate, the adaptive procedure was started at an increment amplitude  $\Delta A$ , several steps larger than the anticipated threshold. In the adaptive procedure, the transition from increasing to decreasing difficulty, and vice versa, defined a turnaround. Adaptive thresholds were determined with a one-down/one-up procedure followed by a two-down/one-up procedure after four turnarounds. The steps in the amplitude increment were accom-

plished by multiplication or division of  $\Delta A$  by a factor  $\mu$  ( $\mu < 1$ ):  $A_{\text{new}} = A_0 + \mu \Delta A_{\text{old}}$  or  $A_{\text{new}} = A_0 + (1/\mu) \Delta A_{\text{old}}$ , respectively. As a result, the amplitude step in dB gets smaller as the difference in amplitude  $\Delta A$  between reference signal and incremented signal gets smaller. For the initial steps,  $\mu$  was 0.66; after four turnarounds,  $\mu$  was set to 0.8. A run was ended after 24 turnarounds and the geometric mean of the  $\Delta A$  values of the last 16 turnarounds was used to estimate the threshold  $\Delta A_{\text{jnd}}$ , theoretically equivalent to 70.7% correct (Levitt, 1971). Assuming unbiased responses, the threshold in this paradigm corresponds to a sensitivity  $d'$  of about 1.265 (see, e.g., Versfeld *et al.*, 1996). For each subject each condition was repeated six times. The test order of the conditions was balanced according to a Latin square.

Discrimination thresholds were expressed as the just-noticeable difference in intensity,  $\text{jnd}(I)$  in decibels:

$$\text{jnd}(I) = 20 \log_{10} \frac{A_0 + \Delta A_{\text{jnd}}}{A_0}, \quad (2)$$

where  $\Delta A_{\text{jnd}}$  indicated the amplitude increment yielding 70.7% correct responses.

Beforehand, to set sensation levels for individual subjects and conditions, masked detection thresholds were determined in a similar manner as described above (3 AFC adaptive procedure). Thus the detection threshold was defined as the threshold at which 70.7% of the stimuli was detected correctly by the listener.

#### D. Subjects

Three subjects (23–25 years), including the first author, participated in the experiments. All had normal hearing (absolute thresholds better than 15 dB HL at octave frequencies from 125 to 4 kHz and at 6 kHz). Subjects were given practice to stabilize their performance. On the average this took 30 min of practice for five successive days. As a result, practice effects were negligible during the actual experiment.

#### E. Data analysis

The data analysis was performed on the logarithms of the  $\text{jnd}(I)$  to make sure that the variance was approximately independent of the size of the  $\text{jnd}(I)$  (see Florentine, 1983; Florentine *et al.*, 1987). Therefore, the average  $\text{jnd}(I)$  was calculated as the geometric mean of the individual data in decibels. An analysis of variance (ANOVA) for repeated measures was used to examine the statistical significance of the effects. Differences were considered significant when the tests indicated a probability less than 0.05.

## II. EXPERIMENTS

### A. Experiment I: Intensity discrimination as a function of shape factor and sensation level

The carrier frequency was 1 kHz. The level of the pink masking noise was set at 35 dB(A) SPL. The masked detection threshold of the stimuli was essentially constant as a function of shape factor (see the Appendix and Fig. A3b for further discussion). The sensation levels of the stimuli were varied from 0 to 30 dB, in 10-dB steps.

Figure 2 shows the discrimination threshold  $\text{jnd}(I)$  as a function of shape factor and sensation level for the individual subjects and the averaged discrimination thresholds across subjects. Error bars indicate the standard error of the mean.

The three listeners show similar behavior. At intermediate levels, i.e., at 10 dB SL for all subjects and at 20 dB SL for subjects JK and NS,  $\text{jnd}(I)$  varies as a function of the shape factor. When the shape factor is increased from 0.0375 to 0.15, intensity discrimination performance deteriorates (higher thresholds): at 10 dB SL,  $\text{jnd}(I)$  increases by a factor of 1.7 when the shape factor is quadrupled. When the shape factor is changed from 0.3 to 1.2, intensity discrimination performance improves (lower thresholds): at 10 dB SL,  $\text{jnd}(I)$  decreases by a factor of 1.4 when the shape factor is quadrupled. The maximum  $\text{jnd}(I)$  (poorest performance) occurs at 10 dB SL for shape factors of 0.15 and 0.3. At lower and higher levels (0 and 30 dB, respectively) the  $\text{jnd}(I)$  does not vary with the shape factor. When the sensation level is increased from 0 to 10 dB, an increase in  $\text{jnd}(I)$  of about 1 dB is observed for a shape factor of 0.15 and 0.3.

The trends shown in Fig. 2 are supported by the statistical analysis. A three-way repeated measures ANOVA [sensation level (4)  $\times$  shape factor (6)  $\times$  subject (3)] on the individual data shows a significant effect of both the sensation level [ $F(3,6) = 5.58$ ;  $p = 0.036$ ] and the shape factor [ $F(5,10) = 11.33$ ;  $p < 0.001$ ]. Also the interaction between level and shape factor is significant [ $F(15,30) = 2.78$ ;  $p < 0.01$ ]. The latter result is probably introduced because, at 10 and 20 dB SL,  $\text{jnd}(I)$  reaches a maximum at a shape factor of 0.15 or 0.3, while at 0 and 30 dB SL  $\text{jnd}(I)$  does not vary systematically as a function of the shape factor.

Two additional experiments (II and III) were conducted to investigate in more detail how the threshold behavior varies with the shape factor.

### B. Experiment II: Intensity discrimination at 4 kHz

To examine whether the 10 dB SL maximum is also present at other carrier frequencies, the 10 dB SL condition was repeated with a carrier frequency of 4 kHz. The results are displayed in Fig. 3. Again, a maximum  $\text{jnd}(I)$  is reached at a shape factor of 0.15 or 0.3.  $\text{jnd}(I)$  increases by a factor of 1.7 when the shape factor is quadrupled from 0.0375 to 0.15, and  $\text{jnd}(I)$  decreases by a factor of 1.7 when the shape factor is quadrupled from 0.3 to 1.2.

The trends are confirmed by a three-way repeated measures ANOVA [carrier frequency (2)  $\times$  shape factor (6)  $\times$  subject (3)] on the individual data from this experiment (4 kHz) combined with the 10 dB SL results from the first experiment (1 kHz). The analysis shows a significant main effect of the shape factor [ $F(5,10) = 19.29$ ;  $p < 0.0001$ ], but no significant effect of carrier frequency nor a significant interaction of frequency and shape factor.

### C. Experiment III: Intensity discrimination for a higher overall level [pink noise: 55 dB(A) SPL]

To examine whether the observed trends really depend on sensation level and not on overall level, we increased the

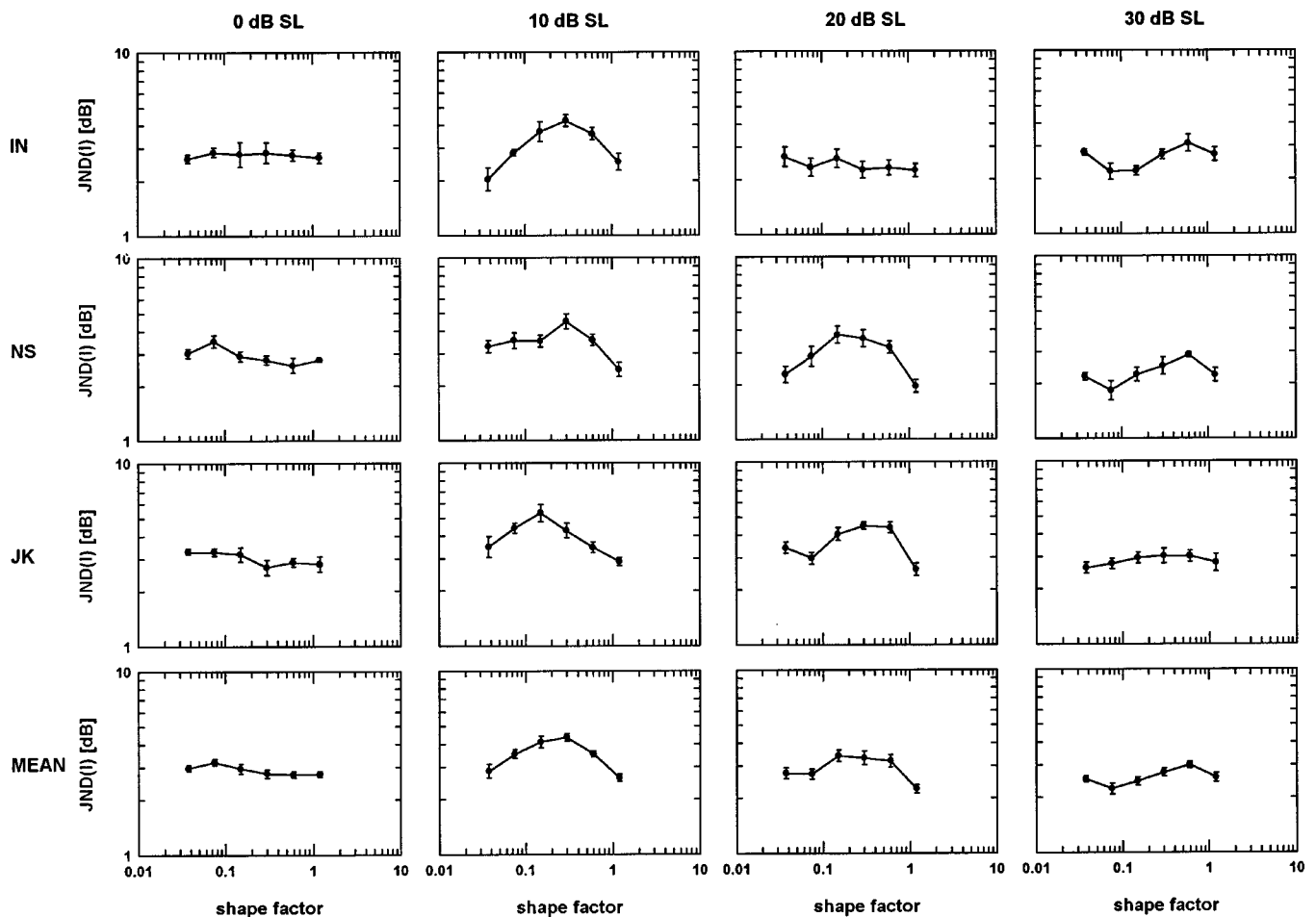


FIG. 2.  $jnd(I)$  for Gaussian-windowed 1-kHz tones plotted as a function of the shape factor, for the individual subjects (Rows 1–3) and averaged across subjects (Row 4). From the left to the right the sensation level is varied: 0, 10, 20, or 30 dB SL. Each point is the mean of six threshold estimates for one observer. The error bars show the standard error of the mean.

background noise level to 55 dB(A) SPL and repeated the experiment with Gaussian-windowed tones of 1 kHz at 10 dB SL.

For the higher overall level, the discrimination thresholds obtained for each subject and for the mean of the three listeners are shown in Fig. 4. Again a maximum was reached for a shape factor of 0.15 or 0.3.  $jnd(I)$  increases by a factor of 1.8 when the shape factor is quadrupled from 0.0375 to 0.15 and decreases by a factor of 1.7 when the shape factor is quadrupled from 0.3 to 1.2.

A three-way repeated measures ANOVA [overall level (2)  $\times$  shape factor (6)  $\times$  subject (3)] on the individual data of this experiment [55 dB(A) SPL] combined with the 10 dB SL data of experiment I [35 dB(A) SPL] shows a significant main effect of the shape factor [ $F(5,10)=33.51$ ;  $p < 0.00001$ ]. The effect of overall level and the interaction between shape factor and overall level are not significant.

### III. DISCUSSION

The results show that the just-noticeable difference in intensity  $jnd(I)$  of Gaussian-windowed tones may vary as a function of the shape factor. For 1-kHz tones at sensation levels of 10 and 20 dB SL in 35 dB(A) SPL pink noise,  $jnd(I)$  reaches a maximum at a critical shape factor of 0.15

or 0.3 (see Fig. 2). At both lower and higher sensation levels,  $jnd(I)$  is relatively constant for different shape factors. For a 4-kHz carrier frequency, a similar variation in  $jnd(I)$  with the signal shape is obtained: again, at a shape factor of 0.15 or 0.3 a maximum is observed (see Fig. 3). Also, after increasing the overall level [noise level: 55 dB(A) SPL] the variation in  $jnd(I)$  persists (see Fig. 4).

In this study the spectro-temporal shape of the stimuli ranged from a relatively long-duration tone to a very short-duration click. As the signals vary from tone to click, two processes occur: temporal shortening and spectral widening. First, as the signal decreases in duration and increases in bandwidth, the primary effect is temporal shortening. This causes an increase in the  $jnd(I)$  until the bandwidth reaches the critical band. At this point the second process, the increase in bandwidth, becomes important, serving to reduce the  $jnd(I)$ . These two processes are addressed more or less separately in literature.

With respect to the effect of temporal shortening, Florentine (1986) and Buus and Florentine (1992) have done extensive research measuring intensity discrimination for pure tones as a function of duration. They found that intensity discrimination improves with increasing duration. Our results show the same behavior: at low levels (10 and 20 dB SL)  $jnd(I)$  decreases toward smaller shape factors, corre-

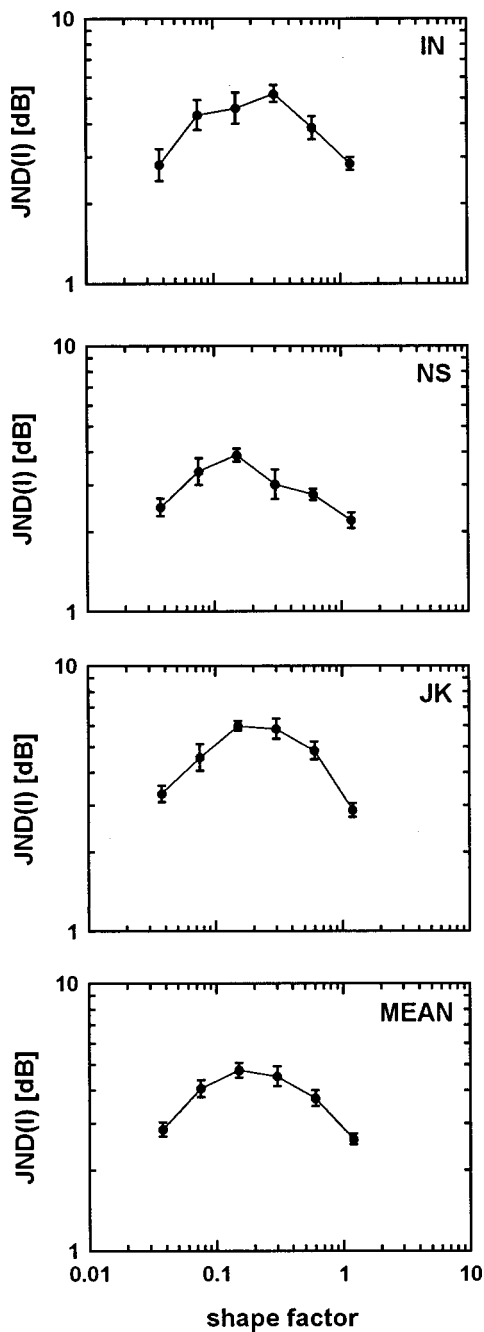


FIG. 3.  $jnd(I)$  for Gaussian-windowed 4-kHz tones plotted as a function of the shape factor. Sensation level: 10 dB; noise level: 35 dB(A) SPL. The upper three panels show the  $jnd(I)$  for the listeners separately. The lowest panel shows the mean result of the three listeners. Other details are the same as in Fig. 2.

sponding to longer durations. Also, quantitatively, the rate of improvement measured in this study agrees with the rate found by Florentine (1986) and Buus and Florentine (1992).

At some point, separating the effect of temporal shortening and the effect of the increase in bandwidth is not possible. Studying  $jnd(I)$  as a function of duration, Florentine (1986) omitted durations of 4 ms and less from the fitting procedure because these data deviated from a linear function [in a double logarithmic plot of  $jnd(I)$  versus duration]. She noted that this may have been due to the spectral splatter. Our data also show this effect, a flattening of the curve for

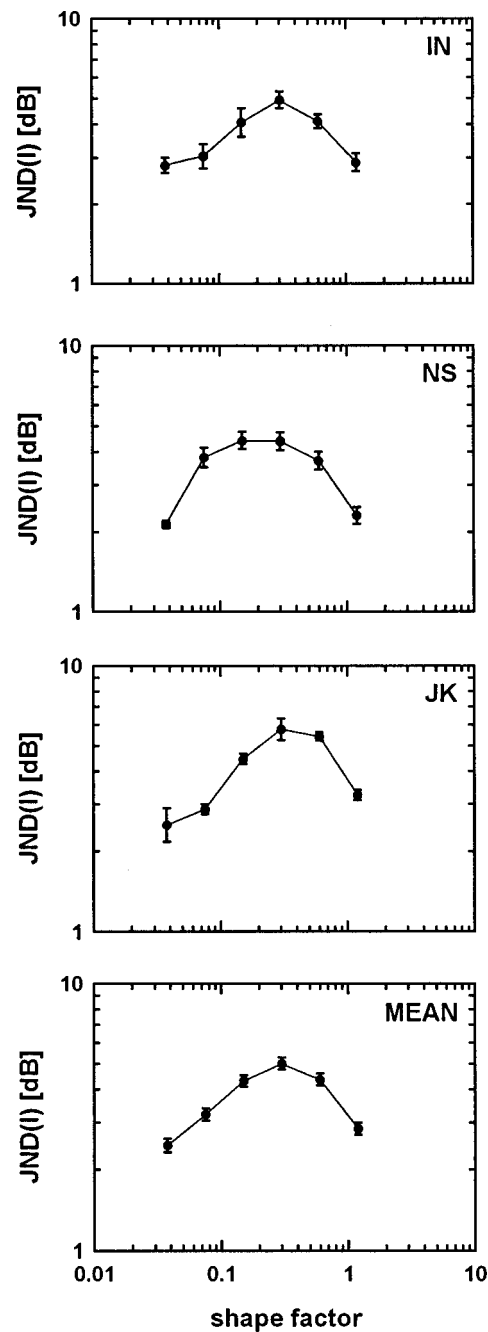


FIG. 4.  $jnd(I)$  for Gaussian-windowed 1-kHz tones plotted as a function of the shape factor. Sensation level: 10 dB; noise level: 55 dB(A) SPL. The upper three panels show the  $jnd(I)$  for the listeners separately. The lowest panel shows the mean result of the three listeners. Other details are the same as in Fig. 2.

small durations, at a shape factor between 0.15 and 0.3. Our explanation is, analogous to Florentine's remark, that at this point the bandwidth of the Gaussian-windowed signal exceeds the width of the auditory filter. From this point on, the *spectral* width of the signal determines the discrimination threshold, i.e., the process of increasing bandwidth becomes important.

With respect to the effect of the increase in bandwidth, Buus (1990) measured intensity discrimination as a function of bandwidth. He found that  $jnd(I)$  is independent of bandwidth when the stimulus bandwidth is less than the width of

the auditory filter. For larger bandwidths, at low levels, a decrease in  $jnd(I)$  with increasing bandwidth was found. Our data also show this trend:  $jnd(I)$  decreases for shape factors larger than 0.3, at sensation levels of 10 dB and 20 dB.

In quantitative terms, an optimum detector predicts a decrease of  $jnd(I)$  by a factor of 2 (in decibels) when the bandwidth or the duration is quadrupled. In the multiple-window model the two processes of the variation in bandwidth and the variation in duration are combined in the variation of the number of f-t windows covered by the signal. Then, in the multiple-window model  $jnd(I)$  is expected to decrease by a factor of 2 when the number of f-t windows is quadrupled. We found a decrease by a factor of 1.7 instead of 2, somewhat less than predicted by the multiple-window model. Florentine (1986) also found smaller improvements (a factor 1.5 when the duration was quadrupled). Possible explanations for this small deviation from the model improvement predictions are a reduced discriminability in the individual f-t windows as the number of f-t windows increases, a suboptimal combination of the information of the different windows (see also Buus and Florentine, 1992) or that the windows are not totally statistically independent.

The most important finding of this study is that, for intermediate sensation levels, the data qualitatively agree with the generalized multiple-window hypothesis put forward in the introduction. As a result, we can identify a “critical” shape factor, for which intensity discrimination performance is worst. This “critical” shape factor has a value between 0.15 and 0.3, both at a carrier frequency of 1 kHz and 4 kHz, at an overall level of 35 dB(A) SPL and 55 dB(A) SPL. So, in the proposed auditory spectro-temporal representation, a Gaussian-windowed sinusoid with a bandwidth of about 1/3 octave and an effective duration of about 4 ms at 1 kHz and 1 ms at 4 kHz (including effectively about four sinusoidal periods) can be considered an approximation of the “elementary” f-t window of the perceptually relevant auditory spectrogram. These values are in line with the idea of the critical band of the multiband excitation-pattern model (Florentine, 1986) and with the temporal constants of the multiple-look model, i.e., about 3 ms for 1-kHz tones and decreasing toward higher center frequencies (Viemeister and Wakefield, 1991).

Having discussed the main issue of the paper, i.e., the relation between  $jnd(I)$  and the shape factor, as observed at 10 or 20 dB SL, in terms of the multiple-window model, a few aspects of the data deserve some further discussion. (1) The masked detection threshold is virtually constant as a function of shape factor. (2) At 0 dB SL,  $jnd(I)$  does not depend on shape factor. (3) For the critical shape factor,  $jnd(I)$  increases about 1 dB when the sensation level increases from 0 to 10 dB SL. (4) At higher sensation levels (30 dB SL), intensity discrimination again is a constant as a function of the shape factor. First, the role of the internal noise versus the external noise in the multiple-window model will be addressed. This will help to clarify points 1 and 2. Then, points 3 and 4 will be discussed.

The noisy representation of intensities in the auditory system that underlies the observed intensity discrimination thresholds is formed as the sum of external and internal vari-

ance. The external variance is mediated by the external background noise added to the signal in the experimental procedure. The internal variance is introduced in the auditory system itself, for example resulting from the variance in the neural coding process. If the signal-to-external-noise ratio is not too low, the internal noise dominates intensity discrimination performance. Following Weber's Law it is assumed that the variance due to internal noise is proportional to the signal energy. Thus the signal-to-internal-noise ratio is independent of the excitation level. Therefore, when the energy of the signal is distributed over several windows rather than concentrated within a single window, the “quality” within each individual window in terms of signal-to-internal noise ratio does not change. As a result, the combination of several windows will yield a better performance, for the internal noise is independent between windows. This forms the basis for the improvement predicted by the multiple-window model. These predictions are consistent with the results found at 10 and 20 dB SL.

However, at very low signal-to-external noise ratios the external noise dominates. Thus to clarify point 1 (masked detection thresholds) and point 2 [ $jnd(I)$  at 0 dB SL], the role of the external noise needs to be discussed. Contrary to the internal noise, the external noise in each window is signal independent. Therefore, when the energy of the signal is distributed over several f-t windows, the signal-to-external-noise ratio in each window decreases. This poorer quality in each individual window is counterbalanced by the combination of the information across several windows. The net effect is a constant threshold as a function of the number of f-t windows for the optimum detector. Thus the explanation for the constant masked thresholds (point 1) and the constant  $jnd(I)$  at 0 dB SL (point 2) is a trade off between the increase in the number of f-t windows covered by the signal and the decrease in the signal-to-external-noise ratio in each individual window. This might imply that masked detection thresholds cannot be used to assess the shape of the auditory f-t window.

The third point to be addressed is the observed increase in  $jnd(I)$  when the sensation level is increased from 0 to 10 dB. We believe that this is due to a two-stage strategy listeners will use in an intensity discrimination task at 0 dB SL: a detection stage followed by a discrimination stage. At 0 dB SL not all stimuli are detectable; the stimulus with the incremented amplitude has a higher probability of being detected. Detecting a stimulus in a particular interval is a one-interval process; discriminating among the stimuli is a three-interval process. Thus due to the difference in memory load (Durlach and Braida, 1969), assuming that listeners benefit from the detection cue at 0 dB SL seems reasonable (see also the Appendix). As a result  $jnd(I)$  is smaller at 0 dB SL than at 10 dB SL. Because most studies regarding  $jnd(I)$  as a function of level report a decreasing  $jnd(I)$  as a function of level, our results, the increase in  $jnd(I)$  when the sensation level increases from 0 to 10 dB, might seem a little unexpected. However, in most studies (see, e.g., Jesteadt *et al.*, 1977; Florentine, 1983; Florentine *et al.*, 1987; Ozimek and Zwillocki, 1996) the lowest sensation level at which the  $jnd(I)$  is measured is 5 or 10 dB SL; At this level the effect of the

detection strategy has probably disappeared. The only study known by the authors that measured the  $jnd(I)$  at 0 dB SL was a study by Hanna *et al.* (1986). Unfortunately, their results can neither confirm nor disprove our results.

Regarding the last point (i.e., 4), spread of excitation is important. In the spectral domain, it is well known that spread of excitation, i.e., the growth of the excitation pattern with increasing level, occurs. As a result,  $jnd(I)$  is independent of bandwidth at high sensation levels. This effect was found by, for instance, Buus (1990) and is also accounted for by the multiband excitation-pattern model (Florentine and Buus, 1981; Buus and Florentine, 1994). In the multiple-window approach, spread of excitation is anticipated both in the temporal and in the spectral domain: The higher the sensation level, the larger the area on the f-t plane excited by the signal. Therefore, at 30 dB SL, probably even for the critical shape factor the internal signal representation may already cover many elementary f-t windows. This may explain why  $jnd(I)$  becomes independent of the shape factor at higher levels.

To substantiate the qualitative arguments of the multiple-window idea, and the role of external and internal noise as described in the preceding paragraphs, a simple detection and discrimination model was developed. We refer to the Appendix for a description of the model. The aim of the model is to simulate the trends observed in the data: the dependence of the discrimination threshold on the shape factor at 10 dB SL, whereas at 0 dB SL the discrimination threshold is a constant as a function of shape factor; the constant detection threshold as a function of shape factor; the slight increase in  $jnd(I)$  at a shape factor of 0.3 when the sensation level increases from 0 to 10 dB SL. The simulated trends (see the Appendix) agree with the observed trends in the data.

The results of this study point to an auditory f-t window with a constant relative bandwidth and a duration inversely related to frequency: at low frequencies the spectral width of the f-t windows is small and the duration long, whereas at high frequencies the spectral width of the f-t windows is broad and the duration short. This perceptually relevant tiling of the f-t plane may be considered a form of scale analysis, suggesting that the signal analysis performed by the auditory system can be modeled by a type of wavelet analysis (see, e.g., Rioul and Vetterli, 1991).

#### IV. CONCLUSIONS

For Gaussian-windowed tones at intermediate sensation levels, intensity discrimination performance varies non-monotonically as a function of the shape factor. The performance is best for extreme shape factors (tones and clicks), and poor (higher thresholds) for intermediate shape factors. The shape factor with the maximum  $jnd(I)$  is defined as the critical shape factor. We propose a generalized multiple-window model to interpret these results. This is based on the idea of an "internal" auditory f-t window in sound coding. The Gaussian-shaped tone with the critical shape factor gives an indication for the shape of the auditory f-t window. The spectral width of this elementary window is roughly a critical band. For the frequencies studied, the temporal width is ap-

proximately four periods: 4 ms at 1 kHz and 1 ms at 4 kHz. A generalized multiple-window model can explain the observed phenomena qualitatively. Further, it is observed that, at very low and high sensation levels,  $jnd(I)$  does not depend on the shape factor. The effects of sensation level on the relative role of external and internal noise, and on spread of excitation provide an explanation for these findings.

#### ACKNOWLEDGMENTS

This work was supported by The Netherlands Organization for Scientific Research (NWO). Thanks are due to two anonymous reviewers for their comments and suggestions. We wish to thank Niek Versfeld for the fruitful and stimulating discussions.

#### APPENDIX: DETECTION AND DISCRIMINATION MODEL

A detection and discrimination model was developed to simulate the trends observed in the data and to substantiate the qualitative arguments of the multiple-window idea, and the role of external and internal noise. After the description of the model, the most important finding, viz. the dependence of the discrimination threshold on the shape factor at 10 dB SL will be simulated. In the simulation this translates into the dependence of the discrimination threshold on the number of elementary f-t windows covered by the signal. Then, analogous to the points addressed in Sec. III, the following trends will be simulated: (1) the detection threshold as a function of the number of f-t windows covered by the signal; (2) the discrimination threshold as a function of the number of windows at 0 dB SL; (3) for one f-t window, the discrimination threshold at 0 and 10 dB SL.

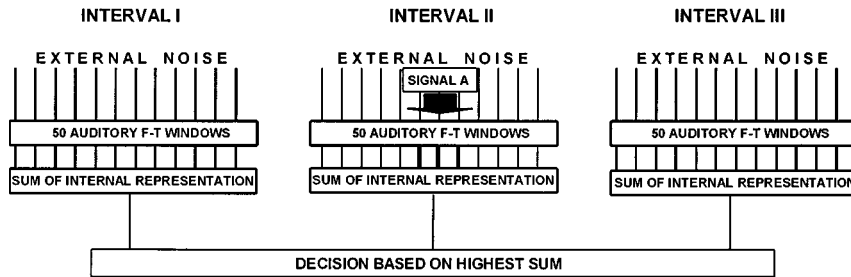
In the simulations the 3 AFC two-down one-up adaptive procedure (see Sec. IC) is adopted: in each trial three intervals are presented; a decision algorithm decides which interval contains the signal in case of detection or the incremented signal in case of discrimination. Thus in the model a human observer is mimicked and the simulated thresholds can be compared directly to the experimental data.

All stimuli have total energy  $E$ . If a signal covers just one f-t window of the auditory system, this f-t window contains the total energy  $E$ . If a signal extends over a number of  $N$  f-t windows, the  $N$  f-t windows contain each  $1/N$  part of  $E$ . In Table I a rough estimate of the number of f-t windows corresponding to the stimuli used in the experiments can be found. This estimate is based on an f-t window with a shape factor of 0.23 (about the "critical" shape factor), corresponding to a Gaussian-windowed stimulus with a bandwidth of  $1/3$  octave. The external noise that enters each f-t window is modeled as Gaussian noise with spectral density  $N_0$ . This noise having a random phase and an amplitude taken from a Rayleigh distribution is added to the signal. In the model, the external noise of the different f-t windows is assumed to be uncorrelated.

#### Detection

In the simulated detection experiments, one interval contains the signal plus external noise and the other two contain

## A) DETECTION MODEL



## B) DISCRIMINATION MODEL

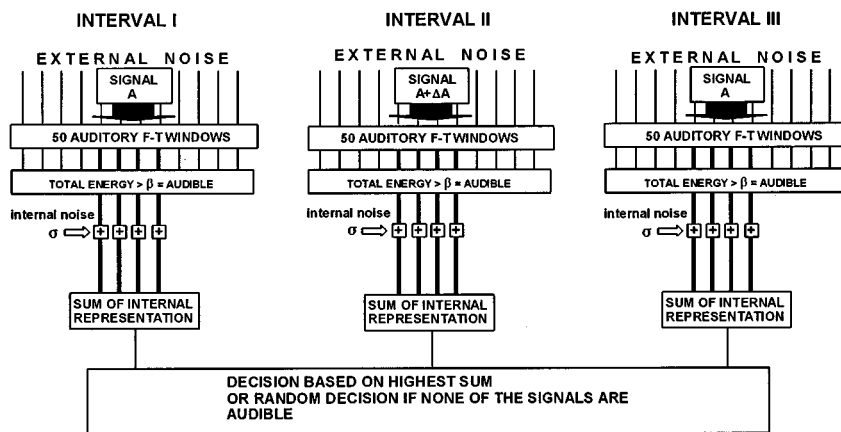


FIG. A1. A schematic representation of the detection (A) and discrimination (B) model.

only external noise. In Fig. A1a a scheme of the detection model is plotted. If the signal covers more than one f-t window (in case of small and large shape factors), the energy of the signal within an interval is divided over the proper number of f-t windows. In a detection task where stimuli are not always audible, assuming that the auditory system is unable to focus exactly on the f-t windows covered by the signal seem reasonable. Therefore, 50 f-t windows are considered for all shape factors, comparable to, for example, an integration time of 200 ms (50 times 4 ms). On this internal auditory representation, detection decisions are based. Detection performance is limited by the external noise  $N_0$  and the total number of f-t windows (set at 50) considered. This classical decision algorithm “picks out” the interval containing the highest sum of the internal representation of stimulus level of the f-t windows.

### Discrimination

In the simulated discrimination experiments, two intervals contain the reference signal and one interval contains the signal with the incremented amplitude. In Fig. A1b a scheme of the discrimination model is plotted. In some of our experimental conditions the stimuli are very close to the detection threshold (0 or 10 dB SL), and as a result the

stimuli are not always audible. Therefore, the decision strategy for the discrimination experiment is divided into two stages: a detection stage followed by a discrimination stage. In the detection stage the decision is made whether the signal is audible or not. Only, if an interval contains an audible signal, this is forwarded to the discrimination stage. Finally, the decision has to be made which of the audible stimuli is the one with the incremented amplitude. This two-stage approach agreed with the experience of the listeners at low sensation levels in the discrimination experiment: the listeners’ strategy was to select only between audible stimuli. According to the listeners’ experience, the decision strategy in the simulations was as follows: If two or all of the stimuli were audible, the interval containing the highest sum of the internal representation of stimulus level over the f-t windows was chosen; if only one interval contained an audible signal, this interval was chosen; if none of the stimuli was audible, randomly one of the three intervals was picked.

In the detection stage, audibility of the signal is defined with respect to the energy distribution of the external noise. In the model, a signal in noise is audible (detectable) if the sum of the internal representation is higher than  $\beta$ . The constant  $\beta$  is chosen such that the probability that noise alone will have a total energy higher than  $\beta$  is 1%. As in the



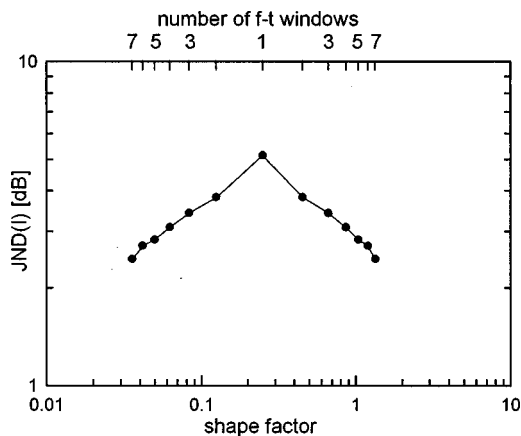


FIG. A2.  $jnd(I)$  as a function of the shape factor at 10 dB SL. An estimate of the number of f-t windows corresponding to the shape factors is shown on the top axis.

detection simulations, in the detection stage the total of 50 f-t windows is considered.

In the discrimination stage where the signal is always audible, assuming that the listener can focus exactly on the f-t windows covered by the signal seems reasonable. Therefore, only the f-t windows containing the signal are considered. ‘‘Coding’’ noise is added to the internal representation. From the literature (see, e.g., Buus and Florentine, 1991) it is known that, in discrimination tasks, the sensitivity  $d'$  is roughly proportional to the difference limen in intensity:  $d' = k * \log_{10}((E + \Delta E)/E)$ . Therefore, the variance of the internal ‘‘coding’’ noise component in the discrimination stage was taken to be proportional to the energy of the signal (constant variance in dB, Weber’s Law): The noise was taken from a Gaussian distribution with a standard deviation  $\sigma$ . Considering the range of the  $jnd(I)$  of our results,  $\sigma = 4$  dB was taken as a reasonable value. The internal noise is uncorrelated across the f-t windows [see Durlach *et al.* (1986)].

### Results of the simulations

The simulated discrimination threshold as a function of the shape factor at 10 dB SL is plotted in Fig. A2. On the top axis of Fig. A2 the number of f-t windows used to simulate the different shape factors is shown. The shape factors and the corresponding estimate of the number of f-t windows can also be found in Table I. At 10 dB SL,  $jnd(I)$  has a maximum for the critical shape factor or, alternatively, for one f-t window.  $jnd(I)$  decreases for smaller and larger shape factors, or, alternatively, as the number of f-t windows increases. These trends are also observed in the data (see Figs. 2, 3, and 4).

The simulated detection threshold as a function of the number of f-t windows covered by the signal is plotted in Fig. A3a. The figure shows that the detection threshold  $E/N_0$  is independent of the number of f-t windows. In Fig. A3b, the mean of the informal detection threshold data at 1 kHz is plotted as a function of shape factor. The data are expressed in decibels *re*: an arbitrary reference. The data show a slight increase in the detection threshold as the shape factor and, as a result, the bandwidth increases. This trend was also observed by Van den Brink and Houtgast (1990) for signals

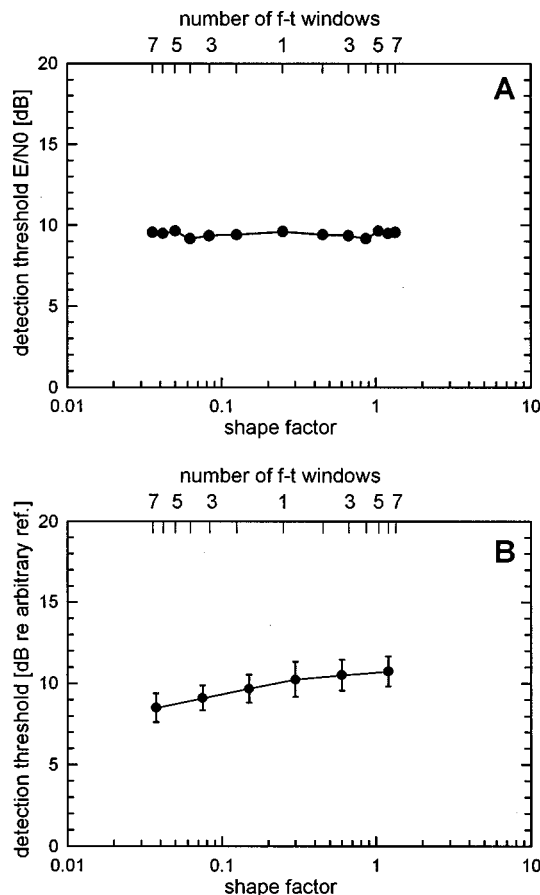


FIG. A3. (A) The model-predicted detection threshold as a function of the shape factor. An estimate of the number of f-t windows corresponding to the shape factors is shown on the top axis. (B) The mean of the informal detection threshold data at 1 kHz as a function of shape factor. The data are expressed in decibels *re*: an arbitrary reference.

with constant spectro-temporal area. Because no maximum (nor minimum) can be observed in our data, it is concluded that, essentially, the detection threshold does not depend on the number of f-t windows covered by the signal. Experimentally observed and simulated trends agree.

The simulated discrimination threshold as a function of the shape factor at 0 dB SL is shown in Fig. A4. On the top

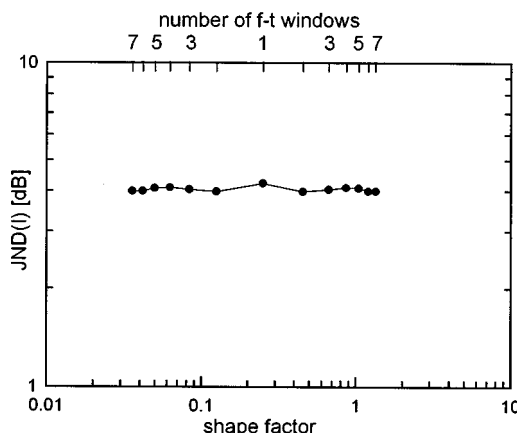


FIG. A4.  $jnd(I)$  as a function of shape factor at 0 dB SL. The corresponding number of f-t windows is shown on the top axis.

axis of the figure the number of f-t windows used to simulate the different shape factors is shown. Please see also Table I. The  $jnd(I)$  at 0 dB SL does not depend on the shape factor, or, alternatively, the number of f-t windows. The simulated trends agree with the data (see Fig. 2). Comparing Fig. A2 (10 dB SL) and Fig. A4 (0 dB SL), it can be seen that for the “critical” shape factor (or one f-t window) the  $jnd(I)$  increases with about 1 dB when the sensation level increases from 0 to 10 dB SL. This trend was also observed in the data (see Fig. 2).

<sup>1</sup>The effective number of periods is defined as the effective duration divided by the period of the carrier frequency of the Gaussian-windowed sinusoid. This is equal to the reciprocal of the shape factor of the signal ( $1/\alpha$ ).

<sup>2</sup>The intensity discrimination task at 0 dB SL is *not* equal to a detection task, because each interval contains a signal. However, if the reference stimuli are presented at 0 dB SL, the signals are not always audible. When, in a trial, one or two of the stimuli are not audible this is perceived by the subject as a mixture between an amplitude discrimination task and a detection task. Hanna *et al.* (1986) also measured  $jnd(I)$  at 0 dB SL.

<sup>3</sup>A function  $f(t)$  has compact support if it is zero outside the interval  $T_0 < t < T_0 + \Delta T$ .

Brink, W. A. C., van den, and Houtgast, T. (1990). “Spectro-temporal integration in signal detection,” *J. Acoust. Soc. Am.* **88**, 1703–1711.

Buus, S. (1990). “Level discrimination of frozen and random noise,” *J. Acoust. Soc. Am.* **87**, 2643–2654.

Buus, S., and Florentine, M. (1991). “Psychometric functions for level discrimination,” *J. Acoust. Soc. Am.* **90**, 1371–1380.

Buus, S., and Florentine, M. (1992). “Possible relation of auditory-nerve adaptation to slow improvement in level discrimination with increasing duration,” in *Auditory Physiology and Perception*, edited by Y. Cazals, L. Démany, and K. Horner (Pergamon, New York), pp. 279–288.

Buus, S., and Florentine, M. (1994). “Sensitivity to excitation-level differences within a fixed number of channels as a function of level and frequency,” in *Advances in Hearing Research*, edited by G. A. Manley, G. M. Klump, C. Köppl, H. Fastl, and H. Oeckinghaus (World Scientific, Singapore), pp. 401–414.

Dai, H., and Wright, B. A. (1995). “Detecting signals of unexpected or uncertain durations,” *J. Acoust. Soc. Am.* **98**, 798–806.

Durlach, N. I., and Braida, L. D. (1969). “Intensity perception. I. Preliminary theory of intensity resolution,” *J. Acoust. Soc. Am.* **46**, 372–383.

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.

Florentine, M. (1983). “Intensity discrimination as a function of level and frequency and its relation to high-frequency hearing,” *J. Acoust. Soc. Am.* **74**, 1375–1379.

Florentine, M. (1986). “Level discrimination of tones as a function of duration,” *J. Acoust. Soc. Am.* **79**, 792–798.

Florentine, M., and Buus, S. (1981). “An excitation-pattern model for intensity discrimination,” *J. Acoust. Soc. Am.* **70**, 1646–1654.

Florentine, M., Buus, S., and Mason, C. R. (1987). “Level discrimination of tones as a function of level and frequency from 0.25 to 16 kHz,” *J. Acoust. Soc. Am.* **81**, 1528–1541.

Florentine, M., Fastl, H., and Buus, S. (1988). “Temporal integration in normal hearing, cochlear impairment, and impairment simulated by masking,” *J. Acoust. Soc. Am.* **84**, 195–203.

Gabor, D. (1947). “Acoustical quanta and the theory of hearing,” *Nature (London)* **159**, 591–594.

Gerken, G. M., Bhat, V. K. H., and Hutchison-Clutter, M. H. (1990). “Auditory temporal integration and the power-function model,” *J. Acoust. Soc. Am.* **88**, 767–778.

Hanna, Th. E., von Gierke, S. M., and Green, D. M. (1986). “Detection and intensity discrimination of a sinusoid,” *J. Acoust. Soc. Am.* **80**, 1335–1340.

Jesteadt, W., Wier, C. C., and Green, D. M. (1977). “Intensity discrimination as a function of frequency and sensation level,” *J. Acoust. Soc. Am.* **61**, 169–177.

Levitt, H. (1971). “Transformed up-down methods in psychoacoustics,” *J. Acoust. Soc. Am.* **49**, 1519–1527.

Ozimek, E., and Zwillocki, J. J. (1996). “Relationships of intensity discrimination to sensation and loudness levels: Dependence on sound frequency,” *J. Acoust. Soc. Am.* **100**, 3304–3320.

Plack, C. J., and Moore, B. C. J. (1990). “Temporal window shape as a function of frequency and level,” *J. Acoust. Soc. Am.* **87**, 2178–2187.

Rioul, O., and Vetterli, M. 1991. “Wavelets and signal processing,” *IEEE Signal Proc. Mag.*, October, 14–38.

Scharf, B., and Buus, S. (1986). “Audition I: Detection and discrimination,” in *Handbook of Perception and Human Performance*, edited by K. Boff (Wiley, New York).

Stewart, G. W. (1931). “Problems suggested by an uncertainty principle in acoustics,” *J. Acoust. Soc. Am.* **2**, 325–329.

Versfeld, N. J., Dai, H., and Green, D. M. (1996). “The optimum decision rules for the oddity task,” *Percept. Psychophys.* **58**, 10–21.

Viemeister, N. F., and Wakefield, G. H. (1991). “Temporal integration and multiple looks,” *J. Acoust. Soc. Am.* **90**, 858–865.

# Speech intelligibility and localization in a multi-source environment<sup>a)</sup>

Monica L. Hawley,<sup>b)</sup> Ruth Y. Litovsky, and H. Steven Colburn  
*Hearing Research Center and Department of Biomedical Engineering, Boston University,  
Boston, Massachusetts 02215*

(Received 7 April 1998; revised 28 December 1998; accepted 23 February 1999)

Natural environments typically contain sound sources other than the source of interest that may interfere with the ability of listeners to extract information about the primary source. Studies of speech intelligibility and localization by normal-hearing listeners in the presence of competing speech are reported on in this work. One, two or three competing sentences [IEEE Trans. Audio Electroacoust. **17**(3), 225–246 (1969)] were presented from various locations in the horizontal plane in several spatial configurations relative to a target sentence. Target and competing sentences were spoken by the same male talker and at the same level. All experiments were conducted both in an actual sound field and in a virtual sound field. In the virtual sound field, both binaural and monaural conditions were tested. In the speech intelligibility experiment, there were significant improvements in performance when the target and competing sentences were spatially separated. Performance was similar in the actual sound-field and virtual sound-field binaural listening conditions for speech intelligibility. Although most of these improvements are evident monaurally when using the better ear, binaural listening was necessary for large improvements in some situations. In the localization experiment, target source identification was measured in a seven-alternative absolute identification paradigm with the same competing sentence configurations as for the speech study. Performance in the localization experiment was significantly better in the actual sound-field than in the virtual sound-field binaural listening conditions. Under binaural conditions, localization performance was very good, even in the presence of three competing sentences. Under monaural conditions, performance was much worse. For the localization experiment, there was no significant effect of the number or configuration of the competing sentences tested. For these experiments, the performance in the speech intelligibility experiment was not limited by localization ability. © 1999 Acoustical Society of America. [S0001-4966(99)00606-2]

PACS numbers: 43.66.Pn, 43.71.Gv, 43.66.Qp [DWG]

## INTRODUCTION

In everyday environments listeners are faced with complex arrays of signals arriving from multiple locations. The auditory system has a remarkable ability to separate out individual sources and to extract information from those sources. Historically, this ability has been referred to as the “cocktail party effect” (Cherry, 1953; Pollack and Pickett, 1958). More recently it has been described as a problem of “sound source determination” (cf. Yost, 1992, 1997) or “sound source segregation” (Bregman, 1990). Considerations of these phenomena lead to the fundamental question: how does the auditory system perform these functions in real-world daily environments, such as a crowded room, a subway station or a social gathering? This question not only is important for understanding how the normal auditory system functions, but has significant implications towards understanding how impaired auditory systems process complex signals in everyday environments.

Early studies on this topic (e.g., Cherry, 1953; Pollack and Pickett, 1958) employed a two-channel competition paradigm, whereby multiple speech signals were presented

over headphones simultaneously to the two ears, such that each ear received completely different stimuli. These studies showed that under certain conditions listeners are able to ignore information presented to one ear and focus on the information presented to the other ear. In addition, intelligibility of speech was markedly worse if both signals were presented to a single ear. Based on these findings and other related work it was proposed that the ability to understand speech in noisy environments improves when the target speech and the competing sounds are spatially separated (e.g., Hirsh, 1950; Dirks and Wilson, 1969).

In order to test the hypothesis that spatial separation enhances speech intelligibility, one must manipulate the relative locations of the target and competing sources. Such manipulations have been conducted in several recent studies which employed virtual-acoustic simulation of environments under headphones (Bronkhorst and Plomp, 1992; Nilsson and Soli, 1994; Koehnke and Besing, 1996; Yost *et al.*, 1996; Peissig and Kollmeier, 1997), as well as actual stimuli in a sound-deadened room (Bronkhorst and Plomp, 1990; Yost *et al.*, 1996). Most studies measured speech intelligibility in the presence of masking noise with the target signal in front and with the direction of the interfering noise as a variable (e.g., Plomp and Mimpen, 1981; Peissig and Kollmeier, 1997). Not surprisingly, a consistent finding across these

<sup>a)</sup>Portions of this paper were presented at the 131st meeting of the Acoustical Society of America, May 1996.

<sup>b)</sup>Electronic mail: MLH@bu.edu

studies is that intelligibility improves when the target and competing sources are spatially separated. However, little is known about the effect of spatial separation for targets located on the sides and with multiple competing sounds that are placed either on the same side or on opposite sides.

The nature of the interfering source(s) is another important consideration. In most studies the competing sources are not other speech tokens, but flat-spectrum noise (e.g., Dirks and Wilson, 1969), speech-spectrum-shaped noise (e.g., MacKeith and Coles, 1971; Plomp and Mimpen, 1979; Koehnke and Besing, 1996), multi-token babble (called speech-simulating noise) (Peissig and Kollmeier, 1997), or noise that is modulated with the envelope fluctuations of speech (Festen and Plomp, 1990; Bronkhorst and Plomp, 1992). A few studies have used speech signals as both targets and competing sounds that were either sentences (e.g., Jerger *et al.*, 1961; Plomp, 1976; Duquesnoy, 1983; Festen and Plomp, 1990; Peissig and Kollmeier, 1997) or single words (Yost *et al.*, 1996); however, typically there was only one competing sound. To the authors' knowledge, very few studies have been conducted with multiple competing sentences (Abouchacra *et al.*, 1997; Ericson and McKinley, 1997; Peissig and Kollmeier, 1997).

Previous studies have shown that speech intelligibility is markedly reduced as the number of competing sources is increased (Bronkhorst and Plomp, 1992; Yost *et al.*, 1996; Abouchacra *et al.*, 1997; Peissig and Kollmeier, 1997). In the present study, both the target and up to three competing sources were presented from several locations in the frontal hemifield in various configurations in an attempt to separate the effects of increasing the number of competing sources from the effect of spatial locations. Each sentence was played from a speaker at a common level so the changes in intelligibility or localizability of the target sound are consequences of the spatial locations of the target and competing sources. Presentation of target and competing sentences spoken by the same talker minimized nonspatial cues such as pitch differences between talkers (cf. Yost, 1997).

The ability of a person to tell where a sound is located in space may also be affected by other sounds in the environment. While many studies have been performed to measure a normal-hearing listener's ability to localize either natural sources (see Blauert, 1997 for review) or simulated sources (e.g., Wightman and Kistler, 1989b; Besing and Koehnke, 1995), relatively few have explored the effect of competing sources (Good, 1994; Wightman and Kistler, 1997b) or reverberation (Giguere and Abel, 1993; Besing and Koehnke, 1995) on localization accuracy. Most localization studies have used noise or click train stimuli; however, in our study the stimuli are sentences. Gilkey and Anderson (1995) showed that anechoic localization performance without competing sources was similar for speech sounds and click trains in the azimuthal dimension.

In the present study we measured both speech intelligibility and localization of speech signals in the presence of up to three competing sentences. In both experiments the location of the competing sources included locations that were close, intermediate and far relative to the target's location. Finally, the experiments were conducted under two condi-

tions: (1) a sound-field condition in which sounds were presented from loudspeakers in a sound-deadened room without head movements and (2) a virtual condition in which sounds were prerecorded through the ears of KEMAR (Knowles Electronic Mannequin for Acoustical Research) from the same loudspeakers in the sound field and presented over headphones in a sound-booth.

## I. METHODS

Each experiment (speech intelligibility and localization) was repeated with different subjects using the two methods of presentation (sound-field and virtual). The stimuli and recording techniques are common to both experiments.

### A. Sound-field room

The sound-field room used for both virtual recording and sound-field testing is a large laboratory room (approximately 30 by 20 ft). Part of the room is separated by office partitions into a testing space approximately 9 by 13 ft. Egg-crate foam sheets line the walls and close the space above the partitions to the ceiling. The testing space was measured to have a reverberation time  $T_{60}$  (time required for level to decrease by 60 dB) of roughly 200 ms for wideband click stimuli. (The value of  $T_{60}$  was between 150 and 300 ms for one-third octave narrow-band noises centered between 250 and 4000 Hz.) The ambient noise level in the room was approximately 50 dBA. Seven loudspeakers were positioned along a 180° arc in the frontal hemifield at 30° increments at a distance of 5 ft from the listener, who was seated at a desk-chair. Head movement was minimized with a modified head/neck rest (Soft-2 Head Support, Whitmeyer Biomechanix). The computer was placed in a nearby IAC sound-booth to reduce fan noise. The experimenter's monitor and keyboard were on a desk in the same room, but outside the testing space.

### B. Listeners

A total of 12 paid listeners (9 females and 3 males) 18–21 years old participated; all were native speakers of English with hearing thresholds at or below 15 dB HL between 250 and 8000 Hz. Three listeners were tested binaurally in the sound field (denoted S1, S2 and S3). Nine listeners were tested using the virtual stimuli: three listeners were tested binaurally (denoted as V1, V2 and V3) and six were tested monaurally, each listening with either their left or right ear (denoted as VL1, VL2, VL3, VR1, VR2, VR3). Each listener participated in both the intelligibility and the localization experiments under the same listening conditions. Listeners had no prior experience in any psychoacoustic experiments.

### C. Stimuli

The speech tokens were sentences from the Harvard IEEE corpus (IEEE, 1969), which consists of 72 phonetically balanced lists of 10 sentences each. A subset of the sentences was recorded by each of two male talkers (45 lists by talker DA and 27 lists by talker CW). The spectrum of each talker

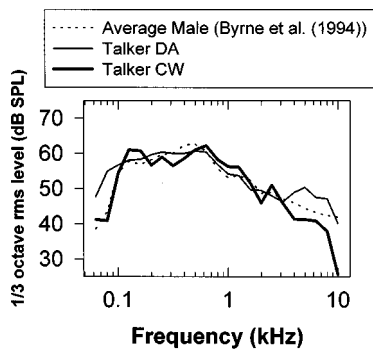


FIG. 1. Spectra for two male talkers in this study compared with average spectrum, normalized to 70 dB SPL overall level for each talker, for American English male talkers from Byrne *et al.* (1994).

is comparable to the average male talker from Byrne *et al.* (1994) as shown in Fig. 1. Each sentence was scaled to the same root-mean-square value, which corresponded to approximately 62 dBA when played from the front loudspeaker, and recorded by a single microphone (Bruel & Kjaer Type 4192) at approximately the position of the listener's head.

In the sound-field experiments, stimuli were played through loudspeakers (Radio Shack Optimus 7) controlled by a personal computer (486DX, Gateway 2000) using Tucker Davis Technologies (TDT) hardware with a dedicated channel for each loudspeaker. Each channel includes a digital-to-analog (D/A) converter (TDT DD3-8), a filter with cutoff frequency of 20 kHz (TDT FT5), an attenuator (TDT PA4) and a Tascam power amplifier (PA-20 MKII). Each speaker was equalized from 200 to 15 000 Hz to have a flat spectrum ( $\pm 1$  dB) over this range by prefiltering each loudspeaker according to its impulse response measured in an anechoic chamber (cf. Kulkarni, 1997). For the sound-field experiments, when a target and competing sentence were presented from the same loudspeaker, the stimuli were digitally mixed prior to the D/A conversion.

For the virtual-source experiments, each of the 720 sentences was recorded when played from each of the seven positions in the room using the same loudspeaker. These recordings were made binaurally through a KEMAR mannequin (Etymotic ER-11 microphone system, with ear canal resonance removed) onto Digital Audio Tape (DAT). The appropriate combinations of sentences recorded from various positions were subsequently digitally mixed (using a Silicon Graphics Inc. Indigo) and recorded back onto a DAT tape. During testing, stimuli were played back (Sony DTC-8 or Sony DTC-700 DAT players) over headphones (Sennheiser HD 520 II) in an IAC sound-treated booth with no compensation for the headphone transfer functions.<sup>1</sup> Subjects were tested individually. For monaural testing, only one headphone was stimulated.

## II. SPEECH INTELLIGIBILITY EXPERIMENT

In the speech intelligibility experiment, the subject's task was to identify the content of an unknown sentence presented from a known location that was fixed throughout a block of trials. Prior to each block, the listener was familiar-

ized with the location of the target sentence, the number of competitors to expect, and the content of each competitor, which were all printed on the answer sheet. However, the location of each competitor was not known and varied from trial to trial. Prior to each trial, a reference stimulus (500-ms, 250-Hz tone) was presented from the target loudspeaker. After each stimulus presentation, the listener was given approximately 16 s to write down the target sentence. *Post hoc* scoring consisted of tallying the number of key words out of five per sentence that had been incorrectly reported. As an example, the key words in the following sentence are italicized: The *streets* are *narrow* and *full* of *sharp* turns. Scoring was strict so that all errors of any kind were counted. An analysis of the results from the binaural virtual condition showed no significant difference between scoring strictly and allowing for minor errors including addition or deletion of suffixes (e.g., played/plays/play) and spelling mistakes (e.g., cloth/clothe) (Dunton and Jones, 1996). Listeners were allowed only one presentation of each target sentence, rather than being permitted an unlimited number of presentations as in the Yost *et al.* (1996) study.

Each sentence in the corpus was designated as either a "target" sentence or "competing" (nontarget) sentence based on its duration; all target sentences were shorter than every competing sentence. While sentences spoken by both talkers were used, in any given block the same talker spoke both the target and the competing sentences. Within a trial, all sentences had nearly synchronous onsets, but asynchronous offsets. The level of each competing sentence was equal to that for each target sentence; hence, the effective level of the overall competing sound increased as the number of competing sentences was increased.

### A. Design

The two parameters of primary interest in this study were the number of competing sentences (one, two or three) and the relative locations of the target and competing sentences. The target sentence was played from one of three locations: left ( $-90^\circ$ ), front ( $0^\circ$ ) or right ( $+90^\circ$ ). Competing sentences were played simultaneously from various combinations of positions categorized as either *close*, *intermediate* or *far* for each target location as listed in Table I. Notice that multiple competitors always originate from separate locations. In the *close* configurations, one competing sentence was at the same position as the target. In the *intermediate* configurations, the nearest competing sentence was  $30^\circ$  to  $90^\circ$  from the target. In the *far* configurations, the nearest competing sentence was more than  $90^\circ$  from the target. The corresponding overall signal-to-noise ratios at the better monaural ear for the various competing sentence configurations are given in Table II.

A third parameter of interest was the effect of sound-field or virtual listening mode. A between-subjects design for listening mode was used in order to accommodate the large design of the experiment without repeating speech material. Therefore, the entire design was repeated for each subject in only one listening mode.

Each target-competing sentence configuration was re-

TABLE I. Categories for competing sentence configurations (– for left, + for right of midline) used in speech intelligibility experiment. Locations presented simultaneously are enclosed by parentheses.  $\Delta_{\min}$  is the minimum absolute separation between target location and location of any competing sentence.

Target location	No. of competing sources	Competing sentence configuration		
		Close $\Delta_{\min}=0^\circ$	Intermediate $30^\circ \leq \Delta_{\min} \leq 90^\circ$	Far $\Delta_{\min} > 90^\circ$
–90° 0° +90°	1	–90° 0° +90°	–60°, –30°, 0° –90°, –60°, –30°, +30°, +60°, +90° +60°, +30°, 0°	+30°, +60°, +90° –90°, –60°, –30°
–90° 0° +90°	2	(–90° –60°), (–90° +90°) (0° +30°) (+60° +90°), (–90° +90°)	(0° +30°) (–90° –60°), (–90° +90°) (0° +30°)	(+60° +90°) (–90° –60°)
–90° 0° +90°	3	(–90° –60° –30°), (–90° 0° +90°) (–30° 0° +30°), (–90° 0° +90°) (+30° +60° +90°), (–90° 0° +90°)	(–30° 0° +30°) (–90° –60° –30°) (–30° 0° +30°)	(+30° +60° +90°) (–90° –60° –30°)

peated between 10 and 25 times (with different target and competing sentences), hence performance is based on scoring between 50 and 125 key words per condition. Different numbers of trials were used for each condition so that there was a similar number of blocks for each target location. Thirty-seven blocks were tested, each containing between 14 and 21 trials. Within a block, the location of the target and the number (and content) of the competing sentences remained constant while the location of the competing sentence(s) was randomized from trial to trial so that each competing sentence configuration was repeated at least twice per block. The order of the blocks was randomized and all listeners performed the experiment in the same order.<sup>2</sup> Due to an error in the preparation of the tape for virtual listening, two conditions (target at –90° or +90° and the competing sentences at –30°, 0° and 30°) were not randomized along with the other conditions. Therefore, for the virtual listening subjects only, these conditions were tested separately after the speech intelligibility and localization experiments were completed.

Listeners received the same, minimal training at the beginning of each testing session. First, a known sentence was presented twice from each of the seven possible locations from –90° to +90°, in order left to right. Then examples of listening to a known sentence in the presence of one, two and then three known competing sentences were played from se-

lected locations. Finally, listeners were given five practice trials for each of one, two and three competing sounds (from only a subset of the locations actually tested), which were not scored. The speech intelligibility experiment contained a total of 600 trials requiring approximately 6.5 h of testing per listener over multiple sessions.

## B. Results

Performance was scored separately for each listener as the error rate, the percentage of key words that were incorrectly identified in each condition. The average and standard deviation across listeners for the two binaural testing groups (sound-field or virtual listening) are shown in Fig. 2. Results for the one, two and three competing sentence cases are plotted in separate panels. The virtual listening subjects (*Vs*) are denoted by open symbols and the sound-field listening subjects (*Ss*) are denoted by filled symbols. The symbol shape corresponds to the location of the target, which was either at –90° (circle), 0° (square) or +90° (triangle). The different symbols are offset slightly for clarity.

Figure 2 shows several important findings. First, for each fixed number of competing sentences, configurations with the largest errors are those in which the competing sentences are close to the target location (i.e., one competing source has the same location as the target). Second, the maxi-

TABLE II. Signal-to-noise ratio at the better ear for all configurations tested. Values were computed using the Shaw and Vaillancourt (1985) source-to-eardrum transfer functions. When there were multiple conditions contributing to the average, the average for all conditions in a category for a given number of competitors is given first and the value for each individual condition is given in parentheses in the same order as reported in Table I. Only the conditions on the left side are shown, the symmetric cases on the right side are equivalent.

Target location	No. of competing sources	Competing sentence configuration		
		Close $\Delta_{\min}=0^\circ$	Intermediate $30^\circ \leq \Delta_{\min} \leq 90^\circ$	Far $\Delta_{\min} > 90^\circ$
–90° 0°	1	–2.5 –0.7	–1.0 (–3.3, –3.0, +3.3) +7.0 (+9.4, +8.6, +3.0)	+10.9 (+7.1, +12.6, +13.0)
–90° 0°	2	–5.7 (–7.6, –3.8) –5.1	–1.1 +0.35 (+4.3, –3.6)	+8.1
–90° 0°	3	–5.4 (–7.7, –3.2) –6.7 (–7.5, –5.8)	–4.6 +1.8	+5.1

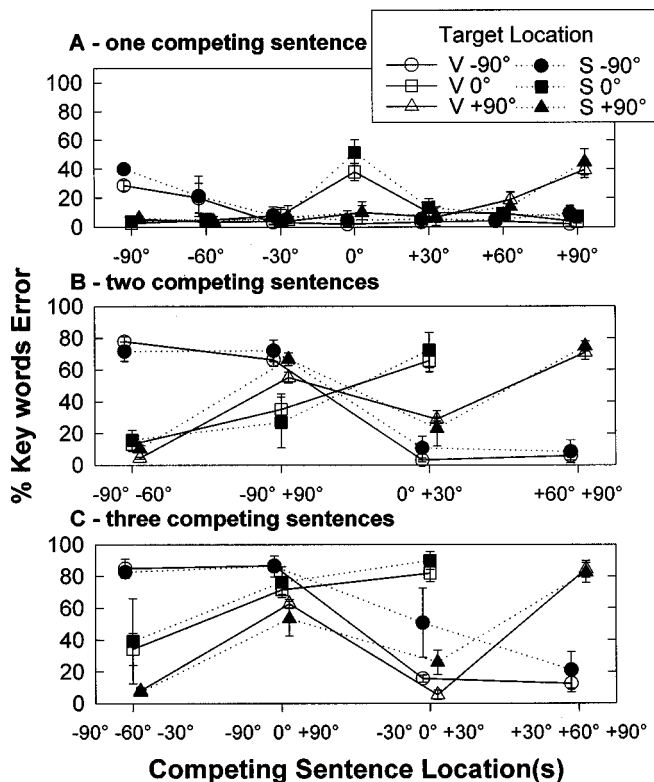


FIG. 2. Speech intelligibility error rates for binaural listening with one- (panel A), two- (panel B) and three- (panel C) competing sentences. Average values and standard deviation over three listeners in each group (virtual-listening and sound-field) are plotted for each competing sentence configuration. Virtual-listening results are shown using open symbols and sound-field listening results are shown with closed symbols. Target location is indicated by symbol: circles for  $-90^\circ$ , squares for  $0^\circ$ , and triangles for  $+90^\circ$ .

imum error rates observed generally increase as the number of competing sentences in the environment is increased (although in some cases, ceiling and floor effects lessen the difference). Third, similar error rates are observed for the sound-field and virtual listening conditions. These observations are consistent with analysis of variance (ANOVA) results. Specifically, a two-way mixed design ANOVA [binaural testing condition (sound-field or virtual listening)  $\times$  competing sentence location(s)] was performed for each target position ( $-90^\circ$ ,  $0^\circ$  or  $+90^\circ$ ) and number of competing sentences (one, two or three) separately (a total of nine ANOVAs). The binaural testing condition was a between-subjects factor and number of competing sentences was a within-subjects factor. For all analyses, significance was considered at a level of 0.05, using Scheffe's correction for multiple comparisons when necessary. For all nine ANOVAs, the results yielded a significant effect of competing sentence location. The configurations when the competing sentences were *close* (one competing sentence at the same location as the target) with respect to the target position yielded significantly more errors than the configurations which were *intermediate* or *far*. There was no significant effect of testing condition (sound-field versus virtual listening); however, an interaction was observed between the testing condition and the competing sentence location for the target at  $-90^\circ$  and  $+90^\circ$  and the three-competing sentences condition, with the

difference only for the competitor locations of ( $-30^\circ$ ,  $0^\circ$  and  $+30^\circ$ ). The error rates were lower when the condition was tested separately (virtual listening) than when tested randomized with other conditions (sound-field listening). Further studies are needed to determine if this difference is due to the particular competitor location or due to the (lack of) randomization.

The ANOVA analyses were repeated using an arcsin transformed version of the error rate which takes into account the limitation of the measurement range (Walker and Lev, 1953; used by Nabelek and Robinson, 1982). The only difference in the results was that there was a significant interaction between the testing condition and three-competing sentence location only for the target at  $+90^\circ$  and no longer for the target at  $-90^\circ$ .

A further validation of the equivalence of sound-field and virtual presentations (indicated by the ANOVA results) is the high correlation between the average error rate for the two binaural testing conditions ( $r=0.96$ ,  $p<0.01$ ); the regression line is not statistically different from the unity line [correlation coefficient:  $t(41)=0.82$ ,  $p>0.05$ ; intercept:  $t(41)=1.40$ ,  $p>0.05$ ]. Thus, speech intelligibility results are essentially equivalent when tested in the sound field or recorded in the sound field and played back over headphones.

Figure 3 shows performance for virtual conditions sepa-

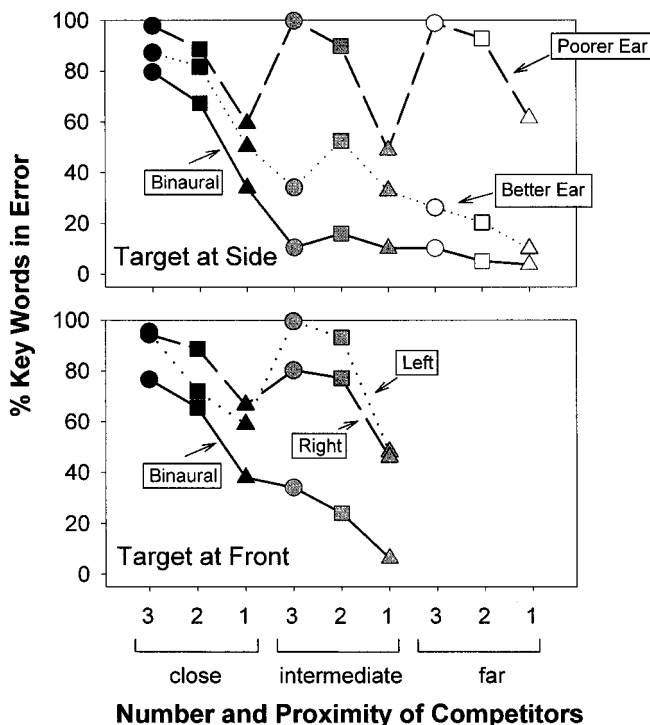


FIG. 3. Speech intelligibility error rates for virtual-listening with the target at side ( $-90^\circ$  and  $+90^\circ$  averaged) and at front ( $0^\circ$ ). Average values are plotted for each group (binaural and each monaural ear) and each competing sentence configuration. Values for binaural virtual-listening are connected with solid lines, those for monaural left virtual-listening (or better monaural) are connected with dotted lines, and those for monaural right virtual-listening (or poorer monaural) are connected with dashed lines. Symbol shading denotes the proximity of the competitors to the target: *close* (black), *intermediate* (gray) or *far* (white) and the number of competing sentences is denoted by the symbol shape *three* (circle), *two* (square) or *one* (triangle).

rated into panels by target position: side (average of symmetrical conditions for target at  $-90^\circ$  and  $+90^\circ$ , top panel), front ( $0^\circ$ , bottom panel). Each panel shows performance under binaural (connected by solid lines) and monaural (connected by dotted and dashed lines) conditions. The proximity of the competing sentences to the target location is denoted by the symbol shading with black for *close* (one competing sound in the same location as the target), gray for *intermediate* (closest competing sound  $30^\circ$  to  $90^\circ$  from the target location) and white for *far* (closest competing sound more than  $90^\circ$  from the target location) competing configurations. The symbol shape denotes the number of competing sentences: circle for three, square for two and triangle for one. The symmetrical conditions, in which the relative location of the target and competing sentences were mirror-images of each other, were averaged since the symmetrical conditions differed by an average of only 2.4% key word errors.

As already noted in Fig. 2, the binaural listening condition results (symbols connected by solid lines in Fig. 3) show that the largest error rates for each target location are consistently seen for the *close* conditions, regardless of the number of competing sentences. The *intermediate* and *far* conditions show similar error rates. It is clear that the proximity of the competing sentences is more influential on the error rate than is the number of competing sentences. Our results in the single competing source case are consistent with previous studies that have found a large decrease in threshold when the competing sound was separated  $90^\circ$  from the target location (e.g., Dirks and Wilson, 1969; MacKeith and Coles, 1971; Plomp and Mimpen, 1981; Bronkhorst and Plomp, 1992; Peissig and Kollmeier, 1997).

When listening monaurally, there is often a large difference in performance depending on which ear is being stimulated (dotted lines and dashed lines in Fig. 3) and the spatial relation of this ear to both the target and the competing sentence locations. The “better monaural ear” is defined as the ear that produces the lowest error rate (which also typically has the higher signal-to-noise ratio<sup>3</sup>) for a particular competing configuration. When the target and competing sounds arise from different spatial locations, the effective signal-to-noise ratio is usually higher at one ear than the other due to the acoustical shadow of the head.<sup>4</sup>

The performance in binaural conditions (Fig. 3) is always better than “better monaural ear” performance. Bronkhorst and Plomp (1992) found an average of 3 dB better performance (corresponds to roughly 30% fewer errors in our technique) for spatially separated competitors in binaural versus “better monaural ear” performance for environments including up to six competitors. Therefore, the lack of large improvement in many environments for the binaural listeners over the better ear monaural listeners is likely due to ceiling and floor limitations of our measurement technique. There are several situations, however, which show a large difference (maximal expected binaural advantage) between binaural and better monaural ear listening in *intermediate* configurations in which the target was at the front ( $0^\circ$ ) in the presence of two or three competing sentences. This suggests that in real-world situations (like the ones we are simulat-

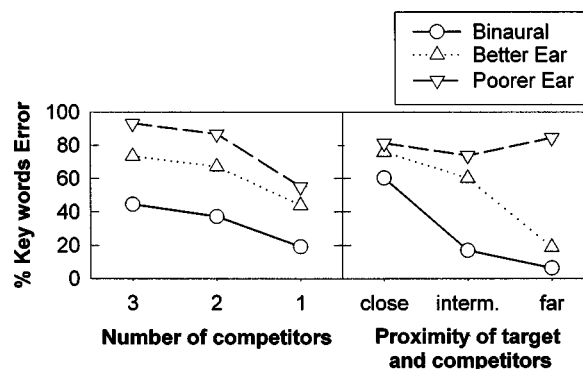


FIG. 4. Error rates in virtual-listening speech intelligibility experiments as a function of the number of competitors and the proximity of the competitors to the target location for binaural (circle), better monaural ear (up-pointing triangle) and poorer monaural ear (down-pointing triangle).

ing), the actual advantage of binaural listening is large only for particular, but typical situations.

### C. Discussion

This study measured speech intelligibility in the presence of multiple competing sentences in which the competing sounds are varied in both their number and proximity to the target location. The major findings are: (1) results are similar in the sound-field and virtual listening conditions; (2) for binaural and better-monaural-ear conditions, the proximity of the target and competing sentences has more influence on the error rate than the number of competing sentences; (3) when the monaural ear is in an unfavorable position, then the number of competing sentences has more influence on the error rate than does proximity of target and competing sentences; (4) there is a large difference between binaural and best-monaural performance only for particular configurations.

There was no evidence that testing in a sound-field without head movements yields different results from testing in a virtual listening experiment, even though individualized ear recordings were not made. This view is consistent with early studies such as Dirks and Wilson (1969) and confirm that the virtual presentation of speech sounds is reasonable (e.g., Bronkhorst and Plomp, 1992; Koehnke and Besing, 1996) for research and clinic applications. Yost *et al.* (1996) showed a difference between sound-field and virtual conditions; however, the listeners in the sound-field were allowed head movements while virtual stimuli were presented independent of head position. Our results suggest that, even though the listeners did not hear the sources in their correct elevations (listeners V1 and V2 reported elevated sources), they were able to use the azimuthal separation of virtual sources to increase the intelligibility of the target.

Figure 4 summarizes the effect of number of competitors and proximity of competitors to the target location separately. For all listening modes, the effect of number of competitors is similar with decreases of 25%–38% as the number of competitors is decreased from three to one. The effect of proximity of target and competitors is only apparent for the binaural and better-monaural listening modes with fewer errors (changes of approximately 55%) when the competitor



locations are moved from close to far. No consistent change in error rates was seen for the poorer-monaural-ear condition as a function of proximity; all data points are near 100% error rate. Therefore, for the binaural and better-monaural-ear listening modes, the effect of proximity of target and competitors have more influence on the error rate than does the number of competitors. This finding is consistent with a previous study by Yost *et al.* (1996). Having both the target and competing sentences be tokens from the same talker may have emphasized this spatial aspect. The poorer-monaural-ear listening mode is more affected by the number of competitors simply because there is no effect of the proximity of the target and competitors. The listeners could not consistently use the spatial separation in the sources to better understand the target under monaural conditions (Yost *et al.*, 1996), which is perhaps related to poorer localization ability under monaural conditions (discussed in Sec. III C).

These data show that monaurally hearing listeners perform much worse than listeners with binaural hearing in many situations; however, they may perform nearly as well in specific conditions when the target is situated on the side of their good ear. Previous comparisons between binaural and monaural listening have generally been made for one competing sound (e.g., Dirks and Wilson, 1969; MacKeith and Coles, 1971) with the exception of Bronkhorst and Plomp (1992) who studied environments containing up to six competitors. While there appear to be small differences between binaural and better-monaural-ear conditions in the single competing sentence situation (triangles in Fig. 3), the largest advantages were observed for configurations with multiple competing sources in intermediate proximity. Our results further suggest that the large binaural advantages are also present in asymmetrical competing configurations when either two or three competing sentences are on the side of the nonfunctioning ear. Bronkhorst and Plomp (1992) determined that the binaural advantage was 2.4 to 4.3 dB depending on the number and location of competing sources. Therefore, the benefit of the binaural system is available regardless of the environment; however, it does not always improve intelligibility.

A similar pattern of binaural advantage (decrease in error rate from better monaural to binaural performance) for the various environments tested is obtained when the target and competitor location(s) are processed through a modified version<sup>5</sup> of the model by Zurek (1993) (Fig. 5). In this model, the intelligibility of the speech is predicted from an Articulation Index calculation based on the signal-to-noise ratio at the better monaural ear and binaural interaction in individual frequency bands weighted by their importance to understanding the speech. This analysis shows that although the signal-to-noise ratio is improved for binaural listening over better monaural listening conditions, the actual intelligibility is not always improved. This lack of consistent improvement is due to the ceiling and floor effects inherent to speech intelligibility.

The advantage of binaural listening over better monaural ear listening is not the most relevant comparison, however, since the monaural listener does not have the luxury to choose the one ear that is “better” for a given situation. A

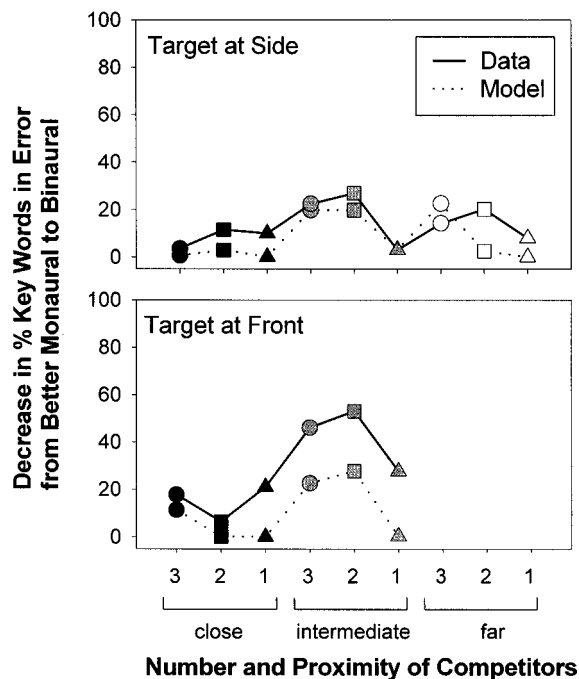


FIG. 5. Decrease in error rates for binaural listening over better monaural listening mode for speech intelligibility experiments as a function of the number of competitors (symbol shape) and the proximity of the competitors to the target location (symbol color). The data are connected by solid lines and the model prediction [modified version of the model by Zurek (1993)] for the same environments are connected by dotted lines.

truly monaural listener would be at a significant disadvantage some of the time. It is interesting to note that the configurations with the target sentence in front and the competing sounds toward the nonstimulated ear (environment with the largest binaural advantage) would be a common environment encountered by a unilaterally deaf person trying to utilize speech-reading.<sup>6</sup> The monaural listener has to either employ speech-reading or to turn his or her head such that the target is at the side and the competing sentences are in front; by reorienting his or her head, the listener would obtain a 40% decrease in error rate for monaural listening. MacKeith and Coles (1971) noted that “if the listener has only one usable ear he will more often be in a position of having either auditory advantage or visual advantage, but not both.”

Our data generally support the claim by Peissig and Kollmeier (1997) that when the target is at 0°, performance is better if two competing sentences are on the same side, compared with conditions when they are distributed on both sides. Although the analysis of our data is limited by the fact that our results include floor effects for the comparable conditions that were tested, there was a small reduction in error from the condition of two competing sentences that were presented from either the left and right (−90° and +90°) to the condition where both were from the left (−90° and −60°).

In summary, the advantage of binaural listening in a realistic environment was shown to be dependent on specific details of the environment. When the speech was already highly intelligible monaurally, there was no improvement to be had by using binaural listening. When the speech intelligibility was poor monaurally, the improvement in signal-to-

noise ratio by a few decibels was sometimes not enough to substantially improve the intelligibility, although in other environments those few decibels led to large improvements in intelligibility. Therefore, binaural hearing was sometimes able to drastically improve the intelligibility; however, the real benefit comes from having the better monaural ear always available.

### III. LOCALIZATION EXPERIMENT

In the localization experiment, the listeners' task was to perform an absolute identification of source location for a known sentence in the presence of unknown competing sentence(s) presented from unknown locations. The number of competing sentences and the content of the target sentence were fixed throughout a block of trials. After each stimulus presentation, the listener was given approximately 6 s to write down the judged location of the target sentence, specified by the number of the location (an integer in the range 1 to 7).

None of the sentences used in the localization experiment had been previously used as targets in the speech intelligibility experiment (which required 600 target sentences out of 720 sentences in the IEEE corpus). The twelve shortest sentences from this pool of 120 sentences were designated as target sentences for this experiment, therefore guaranteeing that the target sentence was shorter than any of its competitors. The competing sentences were chosen from the remaining sentences in the pool and varied from trial to trial. While sentences spoken by both talkers were used, in any given block the same talker spoke both the target and the competing sentences. Within a trial, all sentences had nearly synchronous onsets, but asynchronous offsets. Each sentence was scaled to the same root-mean-square value, corresponding to approximately 62 dBA, the same as was used in the speech intelligibility experiment. Hence, the effective level of the overall competing sound increased as the number of competing sentences was increased.

#### A. Design

The parameters of interest in this study were the number (one, two or three) and configuration (relative locations) of competing sentences. The target sentence was presented randomly from one of the seven speakers ( $-90^\circ$  to  $+90^\circ$ ) simultaneously with competing sentences (one, two or three) in the same configurations as used in the intelligibility experiment (Table I). Note that the competing sentences were clustered on the left, in front, on the right, or distributed on both sides.

The total number of competing sentence configurations is 15 (7 with one competing sentence and 4 for each with two- and three-competing sentences). All conditions for each number of competing sentences were randomized and then tested in blocks of either 70 or 84 trials each with the content of the target sentence fixed. The number of trials in each block differed to give roughly equal number of blocks for each number of competing sentences. Each configuration was repeated about 15 times<sup>7</sup> for each of the seven target positions. The order of the blocks was randomized and all listeners performed the experiment in the same order. A total

of 1575 trials was tested per listener. In addition, the monaural (virtual-listening) subjects localized the target without competing sources.

Listeners were given minimal training, but no trial-by-trial feedback was given. At the beginning of each testing session, sample sentences were played from each location in order between  $-90^\circ$  to  $+90^\circ$ , twice, and listeners were given several practice trials in which they localized a known sentence. There was no feedback during the training trials, and performance was not scored. The localization experiment required approximately 6 h of testing per listener over multiple sessions.

#### B. Results

To quantify the effect of competing sentence configuration on localization performance, several statistics were calculated for each competing sentence configuration: percent correct, root-mean-square error (rms error) and correlation coefficient  $r$  for the least-squares linear regression. The percent correct is a measure of how many correct identifications are made independent of the size of the errors on incorrect trials; the rms error is a measure of the average size of the errors between observed and perfect performance; and  $r^2$ , the proportion of variance accounted for by a linear regression, is a measure of the size of the deviations between the best fitting line through the data and the actual responses (Good and Gilkey, 1996). Random guessing would result in 14% correct, a rms error of  $85^\circ$ , and a  $r^2$  value of 0. The statistics of rms error and  $r^2$  gave consistent information, therefore only rms error is considered further.

The percent correct and rms error for each listener are shown in Fig. 6. The data for each listener are grouped in clusters, with columns slightly offset according to the number of competing sentences. Results for individual subjects are arranged so that binaural listeners are on the left and monaural listeners are on the right. The symbol shape and shading denotes the number of competing sentences. Average and standard deviation are plotted for all conditions having the same number of competitors. Dark lines are plotted for localization with no competing sentences (only measured for the monaural virtual-listening subjects).

When listening binaurally, overall performance was quite good with 89% correct averaged across all listeners and all competing sentence configurations. Individual binaural listeners accurately localized the target sentence on at least 92% of the sound-field trials and 72% of the virtual-listening trials averaged over all competing sentence configurations. At least 95% of the responses for each binaural listener were within one speaker of the correct location. This corresponds to a rms error typically less than  $30^\circ$ . Two of the three listeners in the virtual listening group (V1 and V2) reported that the sounds appeared to be presented from locations that matched the recorded azimuth, but elevated from the recorded locations. While listeners V1 and V2 have the lowest overall percent correct of the binaural listeners, the rms error observed is similar to the other binaural listeners.

Monaural virtual-listening subjects performed much poorer than the binaural listeners. The monaural listeners had an overall average accuracy of 27%, with individual listen-

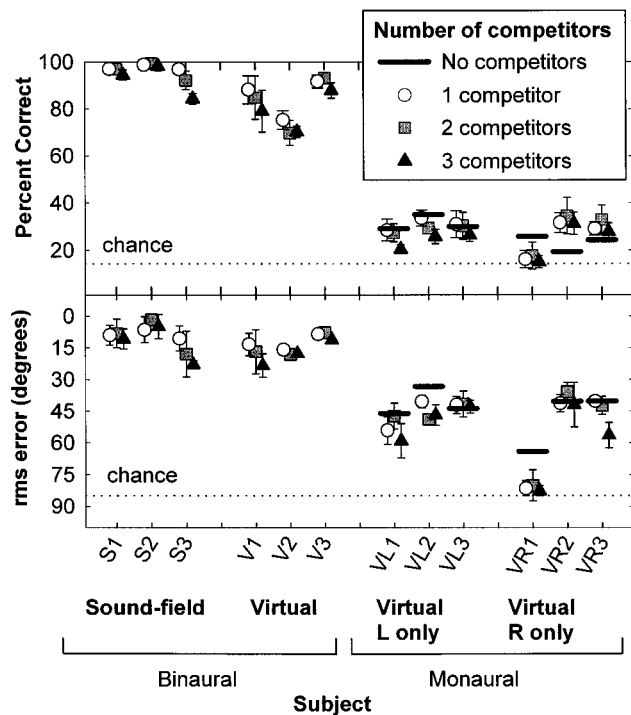


FIG. 6. Results for localization experiments for individual listeners. The percent correct and the root-mean-square (rms) error are plotted for each competing sentence configuration. The symbol and shading denoting the number of competing sentences (white for one, gray for two, black for three). Dark lines denote performance for localization in quiet for monaural listeners.

ers' averages in the range of 17%–31%. (Note that chance performance is 14% since there are only seven possible responses.) The overall rms error was 52°, with individual listeners' averages ranging from 40° to 81°. Localization ability in quiet was not consistently improved over localization ability in the presence of competitors for these listeners.

A two-way mixed design ANOVA [binaural listening condition (sound-field or virtual-listening) × competing sentence configuration (all 15 configurations including one, two or three competitors)] yielded a significant effect of listening condition ( $p < 0.001$ ), but not of competing sentence configuration ( $p > 0.41$ ) or for the interaction between these factors ( $p > 0.98$ ) when evaluated using either percent correct or rms error. The binaural listening condition was a between-subjects factor and the competing sentence configuration was a within-subjects factor. The average percent correct was 96% and 83% for the sound-field and virtual-listening groups, respectively. The average rms error was 10° and 14° for the sound-field and virtual-listening groups, respectively. Although the rms error differs significantly between the groups, the difference in rms error is quite small; average errors are much smaller than the separation between the speakers.

The response distributions of the binaural listening groups for the three-competing-sentence configurations are shown in Figs. 7 and 8 for the sound-field and virtual-listening subjects, respectively. The three-competing-sentence configurations are shown since these configurations typically show the poorest performance observed for the subjects. In these figures, results for a single subject are given in

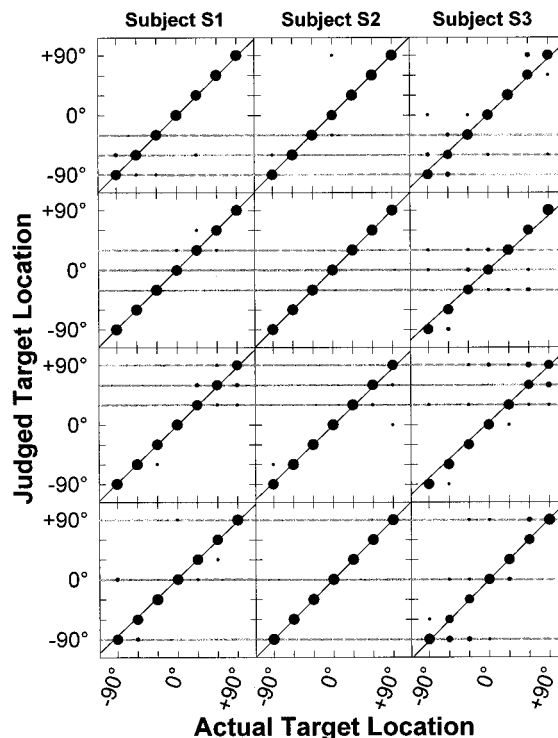


FIG. 7. Confusion matrix of localization responses for the binaural sound-field listening group for three-competing-sentence conditions. Separate listeners are shown in each column. The location of each competing sentence is shown by the gray horizontal lines (first row:  $-90^\circ$ ,  $-60^\circ$  and  $-30^\circ$ ; second row:  $-30^\circ$ ,  $0^\circ$  and  $+30^\circ$ ; third row:  $+30^\circ$ ,  $+60^\circ$  and  $+90^\circ$ ; fourth row:  $-90^\circ$ ,  $0^\circ$  and  $+90^\circ$ ). The size of each circle is proportional to the number of responses judged to that location. The black line is the least square error linear regression for each panel.

each column. Each row of panels shows a different three-competing-sentence configuration with the location of each competing sentence denoted by a horizontal gray line. Each panel shows the distribution of judged target locations for each target location. The size of the circle indicates the relative number of responses for each judged location. Perfect performance corresponds to all judgments along the diagonal. The black line shows the best linear fit for the data in each panel separately.

The pattern of errors for individual subjects highlights the individual differences observed. All three binaural sound-field listeners (Fig. 7) localized the target well, even in the presence of three equal-level competing sentences. The large circles are concentrated along the diagonal and the best linear fit is very close to the diagonal. However, an analysis of the errors observed for the three-competing-sentence configurations shows that while the total error rates are low for all listeners, they occasionally mislocalized by more than 30° from the target location to the location of a competing sentence (S1: 1.3%, S2: 0%, S3: 5.5% of total trials). These infrequent large errors for subject S3 explains the lower percent correct and larger rms error for this subject in these three-competing-sentence configurations as compared with the other sound-field listeners. The virtual-listening subjects (Fig. 8) tend to have more errors that are near the target than was observed for the sound-field listeners. An analysis of the errors observed for the three-competing-sentence configura-

tions shows that, while the total percent correct tends to be lower than was observed for the sound-field listeners, the percentage of errors which were more than 30° and at the location of a competing sentence was not increased (V1: 4.0%, V2: 0.5%, V3: 0.2% of total trials).

Responses for representative monaural (virtual-listening) subjects for the three-competing sentence configurations are shown in Fig. 9 for the poorest performing subject (VR1) and two typical subjects (VL3 and VR2). The three-competing-sentence conditions plotted are the same as shown for the binaural subjects in Figs. 7 and 8. All six monaural-listening subjects showed poorer performance than the binaural subjects did. Responses are no longer tightly clustered near the diagonal, and the best fitting line is often very different from a diagonal line. Subject VR1 shows no relation between his judgments and the actual target location, with correlation coefficients near zero. The other two listeners shown (and the other three not shown) tended to choose the correct side much of the time, but with little differentiation between locations.

### C. Discussion

The major findings of the localization study were: (1) under binaural conditions, localization performance is excellent (average of 90% correct and 12° rms error) for all configurations that were studied; (2) under monaural conditions, localization performance is poor for all configurations that were studied; (3) localization in the sound-field gives significantly smaller rms error than localization under virtual lis-

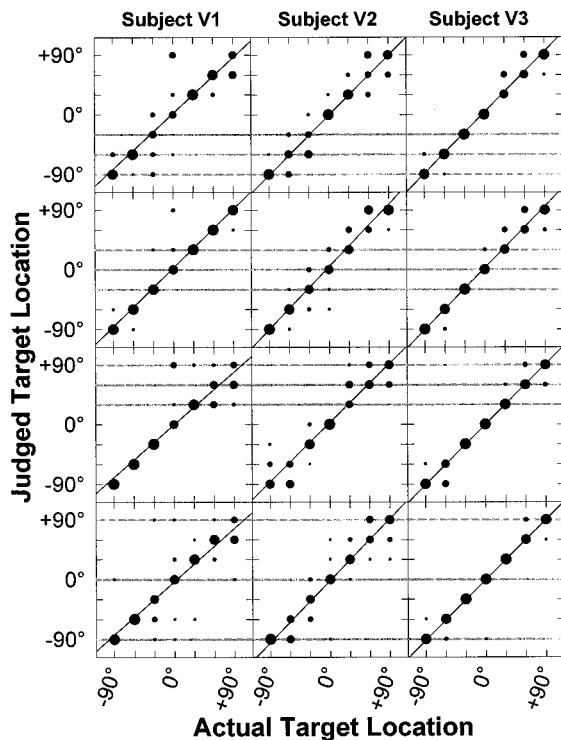


FIG. 8. Confusion matrix of localization responses for the binaural virtual-listening group for the three-competing-sentence conditions. See Fig. 7 for explanation.

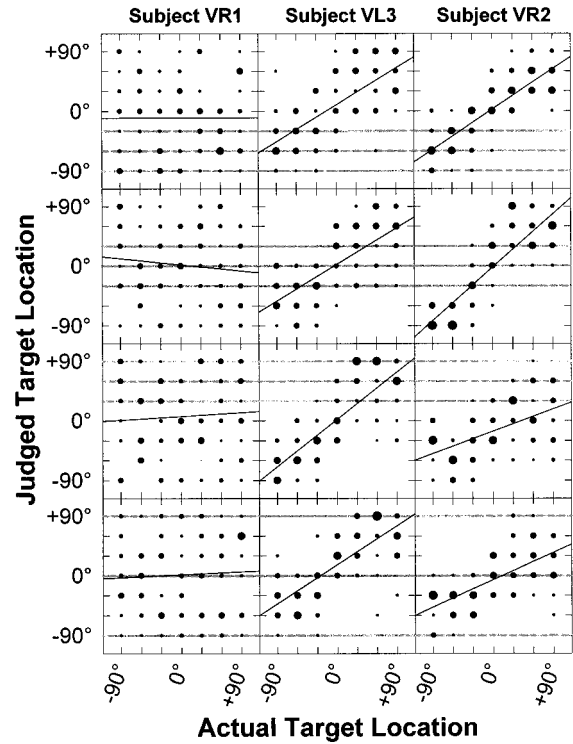


FIG. 9. Confusion matrix of localization responses for representative listeners in the monaural virtual-listening group for the three-competing-sentence conditions. The left column shows listener with accuracy close to chance. The middle and right columns show typical performance for the other five listeners. See Fig. 7 for explanation.

tening conditions; and (4) some subjects showed large infrequent mislocalizations to the location of competing sentences as the environment increased in complexity.

The binaural listeners were able to localize the target sound well even in the presence of competitors for all configurations tested. The stimuli were whole sentences (about 3 s long) presented at a level where the target is detectable even when presented with three competing sentences, although not always intelligible (Hawley and Colburn, 1997). This is consistent with the recent study by Yost *et al.* (1996) in which words that were correctly identified were localized correctly at least 80% of the time.

The observation that localization is better under binaural than monaural conditions is expected (e.g., Middlebrooks and Green, 1991; Slattery and Middlebrooks, 1994; Wightman and Kistler, 1997a), since the localization accuracy in the azimuthal plane is thought to be dominated by the availability of interaural timing information (e.g., Wightman and Kistler, 1992). Monaural listeners do not have interaural timing information and therefore must rely on overall level and spectral differences. Five of six monaural-listening subjects in this study were able to judge the side of the target much of the time presumably based on the level of the target. Since the target was always played at the same level, the level cue would be reliable. The listeners in this study had no experience listening monaurally prior to this study and therefore may not be used to relying on spectral cues. Since the subjects were listening to someone else's (KEMAR) transfer functions, training with feedback may have allowed the subjects to learn these features and improve their performance.

Intersubject differences in listeners' ability to localize sounds under monaural conditions has also been seen by Slattery and Middlebrooks (1994) where some experienced monaural listeners performed better than and others performed comparably to nonexperienced monaural listeners; however, their performance was always poorer than binaural listeners.

Results suggest that listeners' abilities to make absolute identification judgments in the horizontal plane based on virtual sources did not appear to be greatly influenced by the distortions in perceived elevations experienced by subjects V1 and V2. Listeners appear to have categorized their localization judgments in a similar way regardless of the accuracy of the perceived elevation, yielding similar results to those of Besing and Koehnke (1995) who employed comparable virtual-listening methods. The frequent small errors and perceived elevation of sources reported by some of the listeners in the virtual listening group could be due to the use of KEMAR recordings and headphone presentation rather than more precise simulations (e.g., Wightman and Kistler, 1989a, 1989b; Kulkarni, 1997). While features of the head-related transfer functions (HRTFs) differ across listeners, particularly at higher frequencies (e.g., Wightman and Kistler, 1989a; Pralogn and Carlile, 1996), the use of individualized functions have been shown to be less critical for azimuthal localization (e.g., Wenzel *et al.*, 1993).

Our results show a small, yet statistically significant decrease in performance between the sound-field and the virtual-listening subjects. A direct comparison between performance in the sound-field and virtual-listening conditions (e.g., Wightman and Kistler, 1989a; Wenzel *et al.*, 1993) cannot be made since our groups were composed of different listeners. Even for the virtual listener (subject V2) who showed the poorest overall performance (72% correct), the rms error is less than the difference between adjacent speaker locations.

Only two of the binaural listening subjects (listeners S3 and V1) showed increased difficulty localizing sounds as the number of competing sounds increased. For these listeners, the average percent correct decreased by 9.1% and 12.6% and the average rms error increased by 10.0° and 12.4° from the one- to the three-competing sentence environments. For the other four listeners, the average change from the one- to the three-competing-sentence environments was less than 5% correct and less than 3° rms error. Listeners S1 and V1 also showed an increased likelihood of making large mislocalizations to the location of a competing sentence, particularly in the three-competing-sentence environment (4.0% and 5.5% of the total trials). Since all sentences were spoken by the same talker, these listeners (S1 and V1) may have been confused as to which sentence was currently the target sentence and instead localized a competing sentence, although the experimental design intended to minimize this problem by having the target sentence written on the answer sheet in front of the listener and remain consistent for an entire block of trials.

#### IV. INTELLIGIBILITY AND LOCALIZATION IN THE SAME LISTENERS

Intersubject and presentation mode differences noted in the localization experiment did not translate into differences in speech intelligibility performance. Although there were significant differences in localization performance between the binaural groups listening in the sound-field and virtual environments, there were no significant differences between the groups in their intelligibility performance. Results from other experiments (Hawley and Colburn, 1997) show that these sentence stimuli can be localized in quiet at a level 20 dB below the level required for 50% correct key word intelligibility. Since the level of the target stimulus was well above this criterion in all cases, the localization performance for both the sound-field and virtual listening conditions is likely to be close to the asymptote of maximum performance for the subject. It is also likely that the separation of the speakers in the speech intelligibility was too large to show an effect of poorer localization of the stimuli on the intelligibility results. The intersubject differences (not tested statistically) among the monaural listeners is more striking than was seen for the binaural listeners with subject VR1 showing almost no localization ability at all and the other monaural listeners showing better than chance performance. However, the intelligibility performance of all monaural listeners was similar. Therefore, accurate localization of the target source is not necessary for intelligible speech under monaural conditions.

#### V. CONCLUSIONS

- (1) In comparisons between sound-field or virtual listening results, there was no statistically significant difference for the intelligibility experiment and there were small, but significant, differences for the localization experiment.
- (2) The proximity of the competing sentences to the target location was more influential than the number of sources on the intelligibility of sentences under binaural and better-monaural-ear listening conditions. Under poorer-monaural-ear listening conditions, the number of competing sounds was more influential on the observed intelligibility than was proximity of the target and competing sources.
- (3) In environments with equal level target and competitor(s), large differences in intelligibility for binaural over better-monaural-ear listening were observed only for particular multiple-competing-sentence configurations. A model which predicts speech intelligibility in various environments confirmed this finding.
- (4) Localizability of a clearly audible target is quite robust for sentences under binaural conditions even in the presence of three competing sounds regardless of their locations; however, there are occasional mislocalizations to the location of a competing sound.
- (5) Monaural localization is much poorer than binaural localization as expected; however, performance does not typically degrade further even in the presence of three competing sounds. Listeners are generally able to only

judge the correct side that the sound was played from, even in quiet environments.

## ACKNOWLEDGMENTS

This work was supported by NIDCD DC0100. The authors would like to thank Patrick Zurek for supplying the speech recordings used in this study and Abhijit Kulkarni for technical assistance. Leah Dunton and Jennifer Jones collected the data for the binaural virtual-listening conditions for their Senior Project in the Department of Biomedical Engineering. Monika Nalepa provided assistance with the sound-field data collection. Responding to comments by Janet Koehnke, Adelbert Bronkhorst and an additional anonymous reviewer greatly enhanced this paper. Wes Grantham provided thorough and thoughtful suggestions throughout the editorial process.

<sup>1</sup>Kulkarni (1997) showed that the transfer function for the same headphone and listener varies with each placement. This is mentioned only to point out another source of variability in the simulation.

<sup>2</sup>It is unlikely that testing all subjects in the same order corrupted the results since the testing took place over six sessions and measurements for each condition were made during each session.

<sup>3</sup>While the ear with the higher signal-to-noise ratio will also typically give the lower error rate, there can be a discrepancy since the signal-to-noise ratio can improve without improving the intelligibility of the speech. This occurs if the signal-to-noise ratio improvement is due to frequency regions that are not important for speech. For all of the conditions tested in this study, the ear with the lower error rate also has the higher signal-to-noise ratio.

<sup>4</sup>One ear is always more favorable than the other ear in a single competitor environment if the spatial location of the target and competitor are different. However, when there are multiple competitors with symmetrically placed competitors and target from the front, there can be spatial separation without a more favorable ear.

<sup>5</sup>The original model is valid for a single steady noise competitor in anechoic space. Since the competitors are speech instead of noise, to get comparable intelligibility rates between the model and the actual performance the level of the target needed to be decreased by 5 dB. This amount is similar to the 6–8 dB lower thresholds obtained by Festen and Plomp (1990) for a speech competitor than for a noise competitor. To extend the model to multiple competitors, the amount of effective binaural interaction was assumed to be limited by the closest spatial competitor and equal to the binaural interaction as if that were the only competing source.

<sup>6</sup>A proposed rehabilitative strategy for monaural hearing loss is the Contralateral Routing of Signals (CROS) hearing aid configuration (e.g., Harford and Barry, 1965) in which the sound that would have been received at the deaf ear is presented to the good ear through an open ear mold. This strategy was simulated by adding the left and right signals and presenting the sum to the right ear. An additional subject tested in this configuration and performance, relative to having only the right ear stimulus, was comparable to the performance of the better monaural ear only when the target was in front. This resulted in improvements only in the single competitor case (since the ears were not much different in the two- and three-competitor environments). Performance was much worse in conditions in which the right ear was the better monaural ear (target on the same side as the good ear) and not improved when the right ear was the poorer monaural ear. Therefore, limited benefit would be expected from this strategy in multi-source environments.

<sup>7</sup>Due to an error in the making of the tape for this experiment, an unequal number of repetitions were performed for individual target-competitor pairs. The range of repetitions was 8 to 23.

Abouchacra, K. M., Tran, T. V., Besing, J. M., and Koehnke, J. D. (1997). "Performance on a selective listening task as a function of stimuli presentation mode," Association for Research in Otolaryngology Abstracts, 53. Besing, J. M., and Koehnke, J. (1995). "A test of virtual auditory localization," *Ear Hear.* **16**(2), 220–229. Blauert, J. (1997). *Spatial Hearing*, 2nd ed. (MIT, Cambridge, MA).

Bregman, A. S. (1990). *Auditory Scene Analysis: The Perceptual Organization of Sound* (MIT, Cambridge, MA).

Bronkhorst, A. W., and Plomp, R. (1990). "A clinical test for the assessment of binaural speech perception in noise," *Audiology* **29**, 275–285.

Bronkhorst, A. W., and Plomp, R. (1992). "Effect of multiple speech-like maskers on binaural speech recognition in normal and impaired hearing," *J. Acoust. Soc. Am.* **92**, 3132–3139.

Byrne, D., Dillon, H., Tran, K., Arlinger, S., Wilbraham, K., Cox, R., Hagerman, B., Hetu, R., Kei, J., Lui, C., Kiessling, J., Nasser Kotby, M., Nasser, N. H. A., El Kholy, W. A. H., Nakanishi, Y., Oyer, H., Powell, R., Stephens, D., Meredith, R., Sirimanna, T., Tavartkiladze, G., Frolenkov, G. I., Westerman, S., and Ludvigsen, C. (1994). "An international comparison of long-term average speech," *J. Acoust. Soc. Am.* **96**, 2108–2126.

Cherry, E. C. (1953). "Some experiments on the recognition of speech, with one and with two ears," *J. Acoust. Soc. Am.* **25**, 975–979.

Dirks, D. D., and Wilson, R. H. (1969). "The effect of spatially separated sound sources on speech intelligibility," *J. Speech Hear. Res.* **12**, 5–38.

Dunton, L. B., and Jones, J. K. (1996). "Intelligibility and localization of speech in a complex multi-talker environment: Listeners with normal hearing and hearing impairments," senior project, Biomedical Engineering, Boston University, Boston, MA.

Duquesnoy, A. J. (1983). "Effect of a single interfering noise or speech source upon the binaural sentence intelligibility of aged persons," *J. Acoust. Soc. Am.* **74**, 739–743.

Ericson, M. A., and McKinley, R. L. (1997). "The intelligibility of multiple talkers separated spatially in noise," in *Binaural and Spatial Hearing in Real and Virtual Environments*, edited by R. H. Gilkey and T. R. Anderson (Erlbaum, Mahwah, NJ), pp. 701–724.

Festen, J. M., and Plomp, R. (1990). "Effects of fluctuating noise and interfering speech on the speech-reception threshold for impaired and normal hearing," *J. Acoust. Soc. Am.* **88**, 1725–1736.

Giguere, C., and Abel, S. M. (1993). "Sound localization: Effects of reverberation time, speaker array, stimulus frequency, and stimulus rise/decay," *J. Acoust. Soc. Am.* **94**, 769–776.

Gilkey, R. H., and Anderson, T. R. (1995). "The accuracy of absolute localization judgments for speech stimuli," *Journal of Vestibular Research* **5**(6), 487–497.

Good, M., and Gilkey, R. H. (1996). "Sound localization in noise: The effect of signal-to-noise ratio," *J. Acoust. Soc. Am.* **99**, 1108–1117.

Good, M. D. (1994). "The influence of noise on auditory localization in the free field." M.S. thesis, Psychology, Wright State University, Dayton, OH.

Harford, E., and Barry, J. (1965). "A Rehabilitative Approach to the Problem of Unilateral Hearing Impairment: The Contralateral Routing of Signals (CROS)," *Journal of Speech and Hearing Disorders* **30**, 121–138.

Hawley, M. L., and Colburn, H. S. (1997). "Speech intelligibility and localization in a complex environment: relationship to binaural hearing," Second Biennial Hearing Aid Research and Development Conference, Bethesda, MD, p. 60.

Hirsh, I. J. (1950). "The relation between localization and intelligibility," *J. Acoust. Soc. Am.* **22**, 196–200.

IEEE (1969). "IEEE Recommended Practice for Speech Quality Measurements," *IEEE Trans. Audio Electroacoust.* **17**(3), 225–246.

Jerger, J., Carhart, R., and Dirks, D. (1961). "Binaural hearing aids and speech intelligibility," *J. Speech Hear. Res.* **4**(2), 137–148.

Koehnke, J., and Besing, J. M. (1996). "A procedure for testing speech intelligibility in a virtual listening environment," *Ear Hear.* **17**(3), 211–217.

Kulkarni, A. (1997). "Sound localization in real and virtual acoustical environments," Ph.D. dissertation, Biomedical Engineering, Boston University, Boston, MA.

MacKeith, N. W., and Coles, R. R. A. (1971). "Binaural advantages in hearing of speech," *J. Laryngol. Otol.* **85**, 213–232.

Middlebrooks, J. C., and Green, D. M. (1991). "Sound localization by human listeners," *Annu. Rev. Psychol.* **42**, 135–159.

Nabelek, A. K., and Robinson, P. K. (1982). "Monaural and binaural speech perception in reverberation for listeners of various ages," *J. Acoust. Soc. Am.* **71**, 1242–1248.

Nilsson, M. J., and Soli, S. D. (1994). "Norms for a headphone simulation of the hearing in noise test: comparison of physical and simulated spatial separation of sound sources," *J. Acoust. Soc. Am.* **95**, 2994.

Peissig, J., and Kollmeier, B. (1997). "Directivity of binaural noise reduc-

- tion in spatial multiple noise-source arrangements for normal and impaired listeners," *J. Acoust. Soc. Am.* **101**, 1660–1670.
- Plomp, R. (1976). "Binaural and monaural speech intelligibility of connected discourse in reverberation as a function of azimuth of a single competing sound wave (Speech or Noise)," *Acustica* **31**, 200–211.
- Plomp, R., and Mimpen, A. M. (1979). "Improving the reliability of testing the speech reception threshold for sentences," *Audiology* **18**, 43–52.
- Plomp, R., and Mimpen, A. M. (1981). "Effect of the orientation of the speaker's head and the azimuth on 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.
- Pralong, D., and Carlile, S. (1996). "The role of individualized headphone calibration for the generation of high fidelity virtual auditory space," *J. Acoust. Soc. Am.* **100**, 3785–3793.
- Shaw, E. A. G., and Vaillancourt, M. M. (1985). "Transformation of sound pressure level from the free field to the eardrum in numerical form," *J. Acoust. Soc. Am.* **78**, 1120–1123.
- Slattery, 3rd W. H. I., and Middlebrooks, J. C. (1994). "Monaural sound localization: Acute versus chronic unilateral impairment," *Hearing Res.* **75**, 38–46.
- Walker, H. M., and Lev, J. (1953). *Statistical Inference* (Holt, New York).
- Wenzel, E. M., Arruda, M., Kistler, D. J., and Wightman, F. L. (1993). "Localization using nonindividualized head-related transfer functions," *J. Acoust. Soc. Am.* **94**, 111–123.
- Wightman, F. L., and Kistler, D. J. (1989a). "Headphone simulation of free-field listening. I: Stimulus synthesis," *J. Acoust. Soc. Am.* **85**, 858–867.
- Wightman, F. L., and Kistler, D. J. (1989b). "Headphone simulation of free-field listening. II: Psychophysical validation," *J. Acoust. Soc. Am.* **85**, 868–878.
- Wightman, F. L., and Kistler, D. J. (1992). "The dominant role of low-frequency interaural time differences in sound localization," *J. Acoust. Soc. Am.* **91**, 1648–1661.
- Wightman, F. L., and Kistler, D. J. (1997a). "Monaural sound localization revisited," *J. Acoust. Soc. Am.* **101**, 1050–1063.
- Wightman, F. L., and Kistler, D. J. (1997b). "Sound localization in the presence of multiple distracters," *J. Acoust. Soc. Am.* **101**, A3105.
- Yost, W. A. (1992). "Auditory perception and sound source discrimination," *Current Directions in Psychological Science* **1**(6), 179–184.
- Yost, W. A. (1997). "The Cocktail Party Problem: Forty Years Later," in *Binaural and Spatial Hearing in Real and Virtual Environments*, edited by R. H. Gilkey and T. R. Anderson (Erlbaum, Hillsdale, NJ), pp. 329–347.
- Yost, W. A., Dye, R. H. Jr., and Sheft, S. (1996). "A Simulated 'Cocktail Party' with Up to Three Sound Sources," *Percept. Psychophys.* **58**(7), 1026–1036.
- Zurek, P. M. (1993). "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.

# Acoustic pursuit of invisible moving targets by cats

Ralph E. Beitel<sup>a)</sup>

Department of Otolaryngology, 513 Parnassus Avenue, S-762, University of California, San Francisco, California 94143-0732

(Received 4 June 1998; revised 14 December 1998; accepted 9 February 1999)

Head movements evoked by an invisible acoustic target were used as a metric to analyze localization of moving sources of sound in naive cats. The target was presented in the lateral sound field and moved along an arc at constant angular speeds. Head-movement trajectories were characterized by a large-magnitude orienting component that undershot the target, and a tracking component elicited by the target during acoustic pursuit. The tracking component was characterized by a succession of stepwise head movements that maintained a relatively close alignment of the median plane of the head with the moving acoustic target. © 1999 Acoustical Society of America. [S0001-4966(99)03205-1]

PACS numbers: 43.66.Qp, 43.66.Gf [REB]

## INTRODUCTION

Barn owls and some species of echolocating bats and porpoises are able to detect and track the position of moving acoustic targets (Konishi, 1973; Simmons, 1987; Suga, 1984). However, relatively little is known about auditory-tracking behavior in terrestrial mammals. Human subjects may learn to track acoustic targets with a combination of saccadic and smooth-pursuit eye movements (Zambarbieri *et al.*, 1980). Other studies have shown, however, that smooth-tracking eye movements do not occur in humans engaged in an acoustic pursuit task unless visual or proprioceptive cues for target position are available (Buizza *et al.*, 1979; Gauthier and Hofferer, 1976).

In naive cats, an unexpected sound elicits a rapid head-orienting response (OR) that points the eyes and pinnae toward the sound source. The alignment of the median plane of the head relative to the position of an acoustic target has been used as a metric to assess the cat's ability to detect sound direction (Beitel, 1997; Beitel and Kaas, 1993). In the present study, the naive cat's natural tendency to point its head at a sound source was used to measure head tracking of a moving sound. The primary objective of the study was to investigate acoustic pursuit in a species that is used frequently in behavioral and neurophysiological studies of hearing.

## I. METHODS

Seven naive, normal-hearing adult cats were tested in this study. Performance was not rewarded, and testing in six of the animals was limited, therefore, to four trials each to offset response habituation. The seventh cat (C111) was selected specifically for this study, based on informal observations of vigorous orienting behavior to sounds occurring in the animal housing facility. This cat was tested in two sessions of ten trials each.

Prior to testing, a cat was anesthetized with pentobarbital sodium (40 mg/kg, i.p.), the optic nerves were surgi-

cally sectioned, and a percutaneous post was cemented to the midline of the skull. After recovery from surgery ( $\geq 21$  days), a head-position indicator was attached to the post during a testing session to provide a constant midline reference for subsequent analysis of the head trajectories. All procedures followed NIH guidelines for the maintenance and care of laboratory animals.

Stimuli were clicks produced by 1.0-ms rectangular pulses delivered from a Grass stimulator to an amplifier and an attenuator in series with a 3-in.-diameter speaker. For six of the cats, clicks were presented at one pulse rate (5.0, 7.7, or 60 pps). For cat C111, clicks were presented at 7.7 pps in one session and at 60 pps in a second session. Stimulus level was adjusted to 65 dB-A SPL (*re*  $20 \mu\text{N/m}^2$ ) at the region occupied by a cat's head during testing. Stimulus durations were 2.6 or 3.0 s, depending on the angular speed of speaker rotation (16 deg/s or 12 deg/s, respectively).

The cat was harnessed to a platform located in a sound-insulated testing enclosure (Beitel and Kaas, 1993). The speaker was mounted on an L-shaped rod attached to a carousel located below the platform. The position of the speaker was approximately at the level of the interaural horizontal plane at a radial distance of about 43 cm from the cat's head. During a trial, the speaker assembly was rotated at a constant speed by a quiet electric motor that was inaudible inside the testing enclosure.

Because the cat's head was not restrained, the resting position of the head varied from trial to trial. A trial was initiated when (a) the midline indicator was aligned approximately in the straight ahead (0 deg) direction, (b) the azimuthal angle ( $\phi_a$ ) of the speaker relative to the resting position of the cat's head was sufficiently eccentric (mean  $\phi_a = 82 \text{ deg} \pm 14.9 \text{ deg}$ ) to evoke a large-magnitude OR, and (c) the cat's pinnae were aligned in a normal vertical resting position.

Two types of trials were run. On reversal trials, the cat was required to reverse the direction of its head movement to track the acoustic target. For example, if a target moving toward the right appeared on the cat's left, following an OR toward the left, head tracking toward the right was required

<sup>a)</sup>Electronic mail: beitel@phy.ucsf.edu



to pursue the target. On nonreversal trails, if a target moving toward the right appeared on the cat's right, right-ward head movements were required to pursue the target. During a session, each cat was tested at one target speed and one pulse rate. Minimum intertrial interval was 2 min and trials were counterbalanced for target location (right, left) and trial type to offset habituation effects.

During a trial, head and speaker positions were photographed with a constant-speed 16-mm movie camera. The camera was enclosed in a sound-attenuated cabinet mounted above the testing enclosure and was inaudible. The positions of the midline indicator and the speaker were measured with simple drafting tools from projected images of the film. At a click rate of 5 pps, the shutter speed (frames/s) and the click rate were synchronized (200-ms intersample interval); the successive positions of the midline indicator and the center of the speaker were measured frame by frame. At click rates of 7.7 and 60 pps, the shutter speed was 13.3 frames/s (1 frame/75 ms), and measurements were made in odd-numbered frames (150-ms intersample interval). Measurement error was less than 1 deg. On some trials, small pinna movements were observed during head tracking, but features of these movements (e.g., frequency and direction) could not be identified and measured from the film.

## II. RESULTS

Head trajectories had two distinct components: an orienting head saccade (OR) that pointed the midline of the head toward the target, and a tracking component during acoustic pursuit that was frequently preceded by a brief pause. All cats executed an undershooting OR ( $OR < \phi_a$ ) toward the acoustic target on every trial. The results shown in Fig. 1(A) (click rate=7.7 pps) and (B) (click rate=60 pps) for cat C111 include data from reversal and nonreversal trials. The ORs were completed essentially within 500 ms after stimulus onset (0.0 s) and terminated in a position that was maintained for a duration corresponding to at least two frames. The mean relative magnitude ( $OR/\phi_a$ ) of the ORs and the mean residual errors ( $\phi_e = \phi_a - OR$ ) were not significantly different for click rates of 7.7 and 60 pps ( $OR/\phi_a = 0.918$  and  $0.891$ ;  $\phi_e = 6.6$  deg and  $9.7$  deg, respectively;  $t$ -test;  $p > 0.05$ ).

The tracking component of the responses illustrated in Fig. 1 is clustered near the filled symbols that represent target positions. The azimuthal angle ( $\phi_a$ ) of a sound source relative to the head is an important variable contributing to accuracy in orientation (Beitel and Kaas, 1993; May and Huang, 1996), and differences in  $\phi_a$  at the onset of the stimulus probably contributed to the variability associated with head tracking shown in Fig. 1. However, variations in the tracking trajectories are obscured by the scale required on the ordinates in Fig. 1 to display the ORs. Therefore, the tracking components shown in Fig. 1 have been replotted in Fig. 2 to show head position relative to target position at 600 ms after stimulus onset (0.0 s on the abscissas) for reversal [Fig. 2(A)] and nonreversal [Fig. 2(B)] trials. In each plot, the first and last trials are shown with fine and bold dashed curves, respectively. The similar rates at which head position

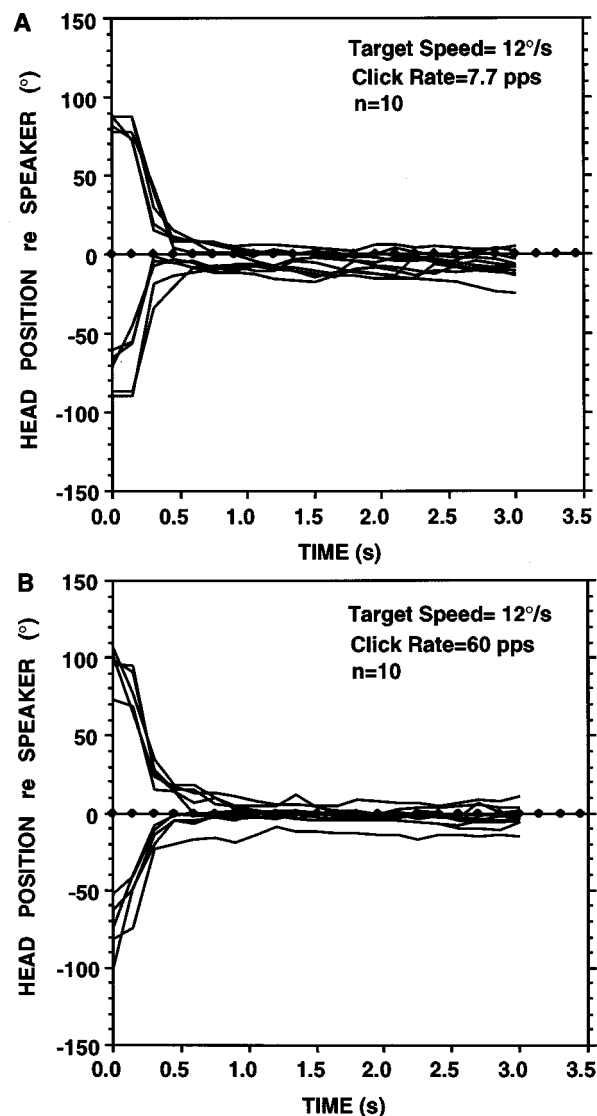


FIG. 1. Head trajectories by cat C111 to moving acoustic targets. Target position is represented by filled symbols. Intersampling interval=150 ms. Ordinates show head midline positions to the right (positive values) and to the left (negative values) of target.

changes over time suggest that response habituation was not an important factor in this study.

The results shown in Fig. 2 illustrate two characteristic features of head tracking of acoustic targets in the naive cat. First, tracking typically occurred as a succession of step-wise head movements. Segments of the curves with negative slopes occur if the speed of the target exceeded the speed of the head movement (e.g., if the cat paused). This pursuit strategy maintained a reasonably good match between the speed of acoustic pursuit and the speed of the target. By averaging the slopes of individual trajectories, the calculated mean speed of acoustic pursuit for cat C111 was  $10.2$  deg/s ( $\pm 1.95$  deg/s: s.d.) during tracking runs of 2.4-s duration. The speed of the target was  $12$  deg/s. Differences in the mean slopes of trajectories at click rates of 7.7 and 60 pps, between reversal and nonreversal trials, and between trajectories to targets on the cat's right or left side were not significant ( $t$ -tests;  $p > 0.05$ ).

Second, pursuit strategy typically aligned the median

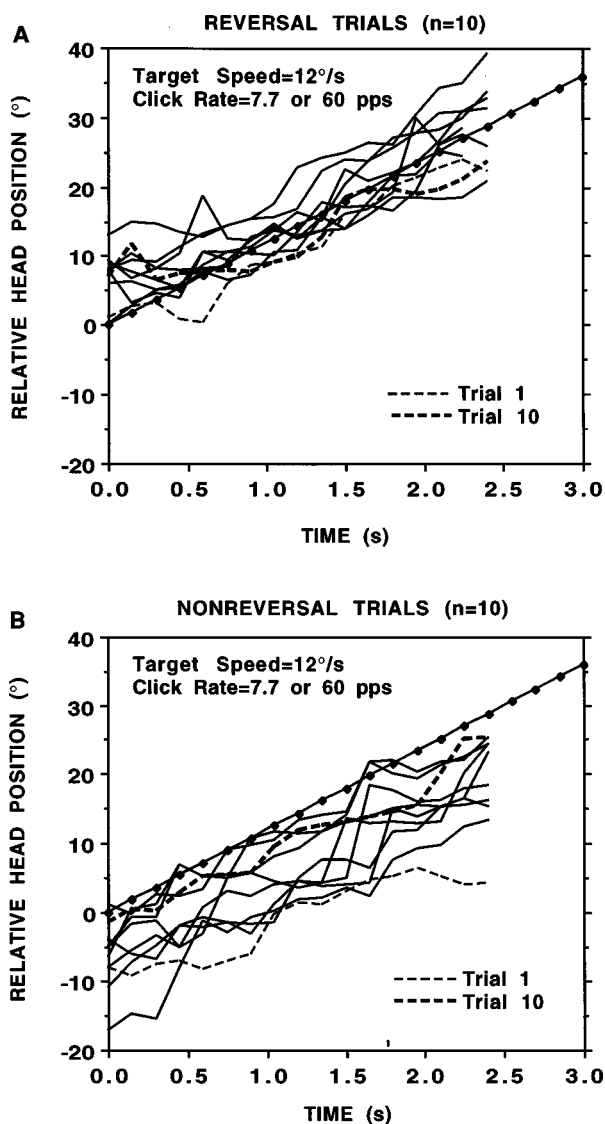


FIG. 2. Head positions (cat C111) relative to target position at 600 ms after sound onset (0.0 on abscissas). Target position: filled symbols. Intersampling interval=150 ms.

plane of the head toward the acoustic target, indicating that the cats localized the position (direction) of the target during its rotation. On reversal trials [Fig. 2(A)], the tracking trajectories for cat C111 are distributed around the symbols representing target trajectory over time; head positions either lead or lag target position (points located above and below the filled symbols, respectively). During tracking runs of 2.4-s duration, the median disparity (error) between head position and target position was 1.0 deg. On nonreversal trials [Fig. 2(B)], head position consistently lags target position; the median disparity between head position and target position was  $-7.2$  deg. This result indicates that when the target moved more laterally, the undershoot or residual error at the completion of the OR was not eliminated during pursuit of the acoustic target.

The difference in disparities between nonreversal and reversal trials is probably due to a procedural difference in the two kinds of trials, rather than to a difference in tracking strategies. Thus, a brief pause in head movement frequently occurred at the completion of an undershooting OR. On a

reversal trial, the residual error at the completion of the OR was reduced during the pause as a consequence of target movement in the direction opposite to the head movement. On a nonreversal trial, the target eccentricity increased during the pause, thereby increasing the disparity between head and target positions. Furthermore, at the more eccentric location of targets on nonreversal trials, poorer resolution of target position would be expected (Beitel and Kaas, 1993; May and Huang, 1996).

In Fig. 3(A)–(H), examples of pursuit trajectories are shown for each cat tested. To facilitate comparisons of pursuit trajectories among cats, head position in Fig. 3 was normalized to coincide with target position at 600 ms after stimulus onset (0.0 s on the abscissas). Examples of reversal and nonreversal trials are illustrated in Fig. 3. The intersampling intervals (200 ms) between data points depicting head positions in Fig. 3(A)–(D) are four times longer than the average latency of ORs in naive cats (50 ms; Beitel and Kaas, 1993). If step-wise movements occurred within intersampling intervals, they would not be reflected in the plots of the pursuit trajectories. Nonetheless, all of the trajectories illustrated in Fig. 3 are serrated, reflecting the fact that tracking typically occurred as a succession of stepwise head movements. For three cats tested at click rates of 7.7 or 60 pps and a target speed of 12 deg/s ( $n=24$  trials), and for four cats tested at a click rate of 5 pps and a target speed of 16 deg/s ( $n=11$  trials), the overall mean speeds of acoustic pursuit were 10.6 deg/s ( $\pm 2.17$  deg/s: s.d.) and 13.8 deg/s ( $\pm 3.6$  deg/s: s.d.). Compared to the speed of the acoustic target, the approximately 2 deg/s-slower average speed of acoustic pursuit reflects a tendency for the disparity between head and target positions to increase during the course of a pursuit trial. Disparity, *per se*, is presumably the adequate condition to induce head tracking. However the tendency for disparity to increase during pursuit may indicate that the target speeds used in this study exceeded the cat's ability to head track an acoustic target.

### III. DISCUSSION

In most studies on directional hearing, a subject's head and the sound source are stationary during each experimental trial (Middlebrooks and Green, 1991; Heffner and Heffner, 1990). In natural listening conditions, however, the head and the acoustic target may be in motion simultaneously, confronting an animal with more complex acoustic situations than those typically investigated experimentally. The unusual feature of the present study is that a relatively natural listening situation was modeled behaviorally by measuring head tracking of a moving acoustic target, i.e., both the cat's head and the acoustic target were in motion during every trial.

The strategy for acoustic pursuit by the cat had two distinct components: a saccadic head-orienting response, and after a brief pause, pursuit head tracking. The OR pointed the pinnae and eyes in the direction of the target. All of the ORs were hypometric, i.e., every OR terminated in an undershoot. In the cat, undershooting is a characteristic feature of head-orienting responses and saccadic eye movements to eccentrically located acoustic targets (Beitel and Kaas, 1993; Beitel, 1997; May and Huang, 1996; Populin and Yin, 1998). Hy-

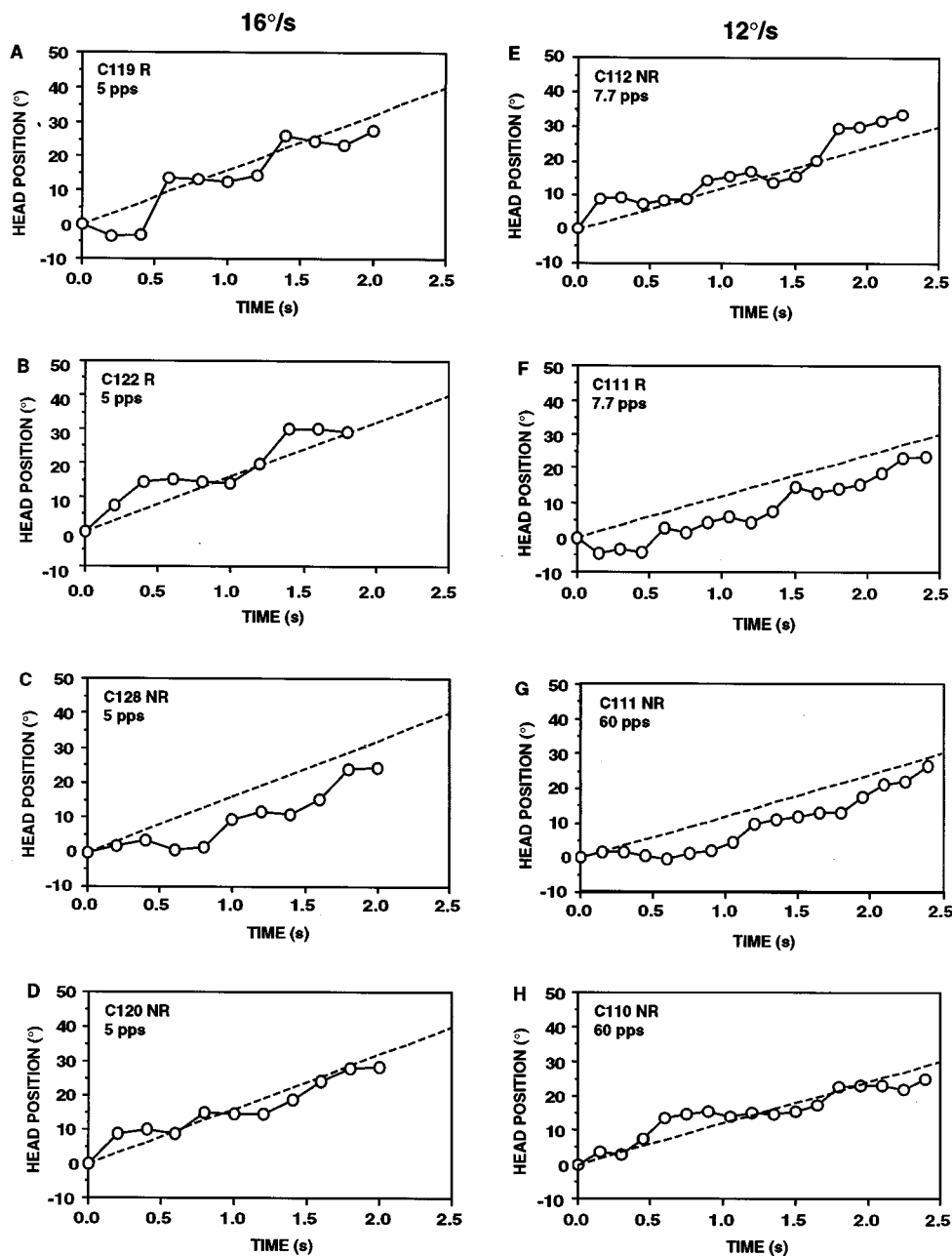


FIG. 3. Examples of acoustic pursuit trajectories by seven cats. Head position is normalized to coincide with target position at 600 ms after sound onset (0.0 on abscissas). (A)–(D): Intersampling interval=200 ms. (E)–(H): Intersampling interval=150 ms. R=reversal trial; NR=nonreversal trial. Target trajectories: dashed curves.

pometria, in fact, is a typical feature of acoustical orientation to eccentrically located sounds, and undershooting head or eye movements are common in owls (Beitel, 1991; Knudsen *et al.*, 1979), monkeys (Jay and Sparks, 1990), and humans (Perrott *et al.*, 1987; Zambambieri *et al.*, 1982).

Head tracking was accomplished with a succession of small-magnitude, stepwise head movements that usually maintained a relatively close alignment of the median plane of the head with the moving acoustic target. Superficially at least, head tracking in cats resembles acoustic pursuit by eye tracking in humans. Sustained episodes of smooth-pursuit tracking were not observed in the naive cat, and in humans, smooth-pursuit eye movements either do not occur (Buizza *et al.*, 1979; Gauthier and Hofferer, 1976) or they may emerge with practice as unstable components during periods of stepwise eye tracking of acoustic targets (Zambambieri *et al.*, 1980).

In previously reported experiments (Beitel, 1991; Beitel

and Kaas, 1993), it has been shown that ORs evoked by eccentrically located broadband signals frequently realigned the sound source to within the animal's frontal sound field. If the sound continued after completion of an undershooting OR, animals were able to reduce residual error on the basis of a secondary, corrective response. May and Huang (1996) obtained similar results in cats using a pair of broadband noise bursts: Undershooting ORs to the first noise burst were followed by corrective head movements to the second burst of noise.

Measurements of free field to eardrum-pressure transformations in the cat indicate that broadband sounds located in the frontal sound field have well-defined, pinna-dependent spectral profiles that may provide information on source location (Musicant *et al.*, 1990; Rice *et al.*, 1992). For eccentrically located sound sources, these putative spectral cues are poorly defined and provide less reliable information on source location. It has been suggested that the degraded spec-

tral profiles associated with eccentrically located sound sources may contribute to the undershooting phenomenon (May and Huang, 1996).

The considerations above suggest a rudimentary hypothesis for acoustic pursuit in the cat. At the onset of a train of clicks, interaural time and level differences and directional amplification effects of the pinnae provided the cat with information on the lateral direction of the eccentrically located sound source. An OR was initiated on the basis of one or more of these putative cues and realigned the moving sound source within the frontal sound field. Because head tracking occurred in a stepwise manner, a reasonable assumption is that the spectral profile of a moving click was sufficiently complex to provide information on its location relative to preceding clicks. The cat might, therefore, move its head stepwise, approximately in phase with a detectable change in the spectral profile brought about by disparity between the position of the sound source and the position of its pinnae.

### ACKNOWLEDGMENTS

I thank Dr. Gregg H. Recanzone for valuable comments on an earlier version of the manuscript. Funding was provided by NINDS Grants NS-5326 and NS-6225 to the Department of Neurophysiology, University of Wisconsin, Madison, WI, and by the Coleman Fund at the University of California, San Francisco, CA.

- Beitel, R. E. (1991). "Localization of azimuthal sound direction by the great horned owl," *J. Acoust. Soc. Am.* **90**, 2843–2846.
- Beitel, R. E. (1997). "Contribution of auditory cortex to acoustical orientation in cats under conditions of discordant auditory reafference," *J. Neurophysiol.* **78**, 3468–3474.
- Beitel, R. E., and Kaas, J. H. (1993). "Effects of bilateral and unilateral ablation of auditory cortex in cats on the unconditioned head orienting response to acoustic stimuli," *J. Neurophysiol.* **70**, 351–369.
- Buizza, A., Legar, A., Berthoz, A., and Schmid, R. (1979). "Otolithic-acoustic interaction in the control of eye movement," *Exp. Brain Res.* **36**, 509–522.

- Gauthier, G. M., and Hofferer, J. M. (1976). "Eye movements to real and apparent motions of acoustic targets," *Percept. Mot. Skills* **42**, 963–971.
- Heffner, R. S., and Heffner, H. E. (1990). "Evolution of sound localization in mammals," in *The Evolutionary Biology of Hearing*, edited by D. B. Webster, R. R. Fay, and A. N. Popper (Springer, New York), pp. 691–715.
- Jay, M. F., and Sparks, D. L. (1990). "Localization of auditory and visual targets for the initiation of saccadic eye movements," in *Comparative Perception, Vol. 1, Basic Mechanisms*, edited by M. A. Berkley and W. C. Stebbins (Wiley, New York), pp. 351–374.
- Knudsen, E. I., Blasdel, G. G., and Konishi, M. (1979). "Sound localization by the barn owl (*Tyto alba*) measured with the search coil technique," *J. Comp. Physiol.* **133(A)**, 1–11.
- Konishi, M. (1973). "How the owl tracks its prey," *Am. Sci.* **61**, 414–424.
- May, B. J., and Huang, A. Y. (1996). "Sound orientation behavior in cats. I. Localization of broadband noise," *J. Acoust. Soc. Am.* **100**, 1059–1069.
- Middlebrooks, J. C., and Green, D. M. (1991). "Sound localization by human listeners," *Annu. Rev. Psychol.* **42**, 135–159.
- Musicant, A. D., Chan, J. C. K., and Hind, J. E. (1990). "Direction-dependent spectral properties of cat external ear: New data and cross-species comparisons," *J. Acoust. Soc. Am.* **87**, 757–781.
- Perrott, D. R., Ambarsoom, H., and Tucker, J. (1987). "Changes in head position as a measure of auditory localization performance: Auditory psychomotor coordination under monaural and binaural listening conditions," *J. Acoust. Soc. Am.* **82**, 1637–1645.
- Populin, L. C., and Yin, T. C. T. (1998). "Behavioral studies of sound localization in the cat," *J. Neurosci.* **18**, 2147–2160.
- Rice, J. J., May, B. J., Spirou, G. A., and Young, E. D. (1992). "Pinna-based spectral cues for sound localization in cat," *Hearing Res.* **58**, 132–152.
- Simmons, J. A. (1987). "Directional hearing and sound localization in echolocating animals," in *Directional Hearing*, edited by W. A. Yost and G. Gourevitch (Springer, New York), pp. 214–225.
- Suga, N. (1984). "The extent to which biosonar information is represented in the bat auditory cortex," in *Dynamic Aspects of Neocortical Function*, edited by G. M. Edelman, W. E. Gall, and W. M. Cowan (Wiley, New York), pp. 315–373.
- Zambarbieri, D., Schmid, R., Magenes, G., and Prablanc, C. (1982). "Saccadic responses evoked by presentation of visual and auditory targets," *Exp. Brain Res.* **47**, 417–427.
- Zambarbieri, D., Schmid, R., Prablanc, C., and Magenes, G. (1980). "Characteristics of eye movements evoked by the presentation of acoustic targets," in *Progress in Oculomotor Research*, edited by A. F. Fuchs and W. Becker (Elsevier, New York), pp. 559–566.

# Sound localization in noise in hearing-impaired listeners<sup>a)</sup>

Christian Lorenzi, Stuart Gatehouse,<sup>b)</sup> and Catherine Lever

MRC Institute of Hearing Research (Scottish Section), Royal Glasgow Infirmary, Queen Elizabeth Building, 16 Alexandra Parade, Glasgow G31 2ER, Scotland, United Kingdom

(Received 25 September 1997; revised 19 October 1998; accepted 10 March 1999)

The present study assesses the ability of four listeners with high-frequency, bilateral symmetrical sensorineural hearing loss to localize and detect a broadband click train in the frontal-horizontal plane, in quiet and in the presence of a white noise. The speaker array and stimuli are identical to those described by Lorenzi *et al.* (in press). The results show that: (1) localization performance is only slightly poorer in hearing-impaired listeners than in normal-hearing listeners when noise is at 0 deg azimuth, (2) localization performance begins to decrease at higher signal-to-noise ratios for hearing-impaired listeners than for normal-hearing listeners when noise is at  $\pm 90$  deg azimuth, and (3) the performance of hearing-impaired listeners is less consistent when noise is at  $\pm 90$  deg azimuth than at 0 deg azimuth. The effects of a high-frequency hearing loss were also studied by measuring the ability of normal-hearing listeners to localize the low-pass filtered version of the clicks. The data reproduce the effects of noise on three out of the four hearing-impaired listeners when noise is at 0 deg azimuth. They reproduce the effects of noise on only two out of the four hearing-impaired listeners when noise is at  $\pm 90$  deg azimuth. The additional effects of a low-frequency hearing loss were investigated by attenuating the low-pass filtered clicks and the noise by 20 dB. The results show that attenuation does not strongly affect localization accuracy for normal-hearing listeners. Measurements of the clicks' detectability indicate that the hearing-impaired listeners who show the poorest localization accuracy also show the poorest ability to detect the clicks. The inaudibility of high frequencies, "distortions," and reduced detectability of the signal are assumed to have caused the poorer-than-normal localization accuracy for hearing-impaired listeners. © 1999 Acoustical Society of America. [S0001-4966(99)04106-5]

PACS numbers: 43.66.Qp, 43.66.Sr [DWG]

## INTRODUCTION

For listeners with sensorineural hearing loss, there are now considerable data describing the ability to localize or lateralize a sound in quiet environments (for reviews, see Durlach *et al.*, 1981, and Noble *et al.*, 1994), but very few studies investigating this ability in the presence of an interfering noise (e.g., Viehweg and Campbell, 1960; Abel and Hay, 1996). Therefore, it remains unclear whether people with cochlear damage experience greater difficulties than normal-hearing listeners when they have to locate a sound source in a noisy background (a fairly typical situation).

The ability to identify the azimuth of a sound is primarily dependent on three acoustic cues: interaural time differences (ITD), interaural level differences (ILD), and spectral shape cues. According to the "duplex" theory, localization judgments in the frontal horizontal plane are primarily based on analysis of ITD cues at low frequencies and ILD and spectral shape cues at high frequencies. (Lord Rayleigh, 1907; Mills, 1972; Kistler and Wightman, 1992). As high frequencies are often inaudible to listeners with a sensorineural hearing loss, these listeners would have reduced access to ILD and spectral cues and should therefore rely mostly on ITD cues. A number of studies show that such an impoverished situation does not substantially impair localization accuracy in the frontal-horizontal plane *in quiet* as long as the

stimuli are broadband and the hearing loss is bilateral and symmetric (e.g., Nordlund, 1964; Röser, 1966; Noble and Byrne, 1990; Noble *et al.*, 1994). When an interfering noise is present, however, it is unclear whether this reduced-cue situation would lead to judgments in the frontal-horizontal plane as accurate as in a multiple-cue situation.

Viehweg and Campbell (1960) studied the ability of listeners with normal hearing and listeners with unilateral hearing losses (at least 30 dB asymmetry between the ears, mean loss in the better ear 1.2 dB over 500–2000 Hz, mean loss in the poorer ear 63 dB) to localize a speech signal, in quiet and in a background noise, emanating from two speakers placed on the right and the left of the listener. Of the 51 impaired subjects, 33 had completely normal better ears, and 18 had some degree of high-frequency loss, the mean of which was 57.2 dB at 4000–8000 cps. All listeners were instructed to hold their heads steady during a stimulus presentation. The results showed that: (1) both the number of identification errors and magnitude of these errors were substantially larger for the hearing-impaired listeners than for the normals in both quiet and noise, and (2) performance amongst the unilateral subjects did not depend on the amount of high-frequency loss in the better ear. Durlach *et al.* (1981) pointed out that the second result was surprising inasmuch as the hearing-impaired listeners must have been operating essentially monaurally (so that localization information must have come primarily from high-frequency spectral cues or head movements). They suggested that listeners may not have held their heads fixed as instructed.

<sup>a)</sup>Parts of this work were presented at the 133rd meeting of the Acoustical Society of America (Lorenzi *et al.*, 1997).

<sup>b)</sup>Electronic mail: stuart@ihr.gla.ac.uk

A recent study performance by Abel and Hay (1996) is more pertinent to the present investigation because the impaired listeners who participated in their experiment showed a bilateral, symmetrical, high-frequency, sensorineural hearing loss. The hearing threshold levels (HTLs) of these listeners were below 25 dB SPL at 0.5 kHz and between 40 and 60 dB SPL at 4 kHz. Abel and Hay measured their ability to localize low- and high-frequency narrow-band signals in quiet and in the presence of a continuous white noise. Noise was presented by means of three speakers arranged to create a homogeneous sound field (though by inference, the noises presented to the three loudspeakers were not independent). All listeners were instructed to hold their heads steady during a stimulus presentation. The results showed that localization performance was poorer for the hearing-impaired listeners than for the normal-hearing listeners, in both quiet and noise. For both groups of listeners, the introduction of the noise decreased only the localization performance for the low-frequency signals, the decrease in performance being similar for both groups of listeners. These results suggested that: (1) noise disrupts mainly ITD cues, and (2) this effect is similar in both listeners with normal hearing and listeners with cochlear damage. However, Abel and Hay measured localization performance at only one signal-to-noise ratio (+15 dB S/N) and localization performance was assessed with a relatively small set of response alternatives: six speakers surrounded the listener at azimuth angles 60 deg apart. It could therefore be argued that their study did not provide the level of precision required to differentiate completely between the localization performance in noise of normal-hearing and hearing-impaired listeners.

In a companion paper, Lorenzi *et al.* (1999) obtained data with normal-hearing listeners exposed to low-pass filtered clicks that may be taken as a prediction (or a simulation) of the effects of a high-frequency hearing loss on local-

ization performance in noise. As in the preceding studies, all listeners were instructed to hold their heads steady during a stimulus presentation. The results showed that the ability to localize a click-train signal low-pass filtered at 1.6 kHz was not greatly affected by a white-noise masker when the latter was presented straight ahead, but began to decrease at a signal-to-noise ratio of 0 to 6 dB when noise was presented at the side of the listener (see Figs. 6 and 7 in Lorenzi *et al.*, 1999). For broadband and high-pass filtered clicks, localization judgments remained accurate when the noise masker was presented straight ahead. When noise was presented at the sides of the listener, localization accuracy decreased in both the broadband and high-pass conditions, but the decrease began at a lower signal-to-noise ratio (S/N) (-6 to 0 dB) and the error score at the lowest S/N (-9 dB) was smaller compared to the low-pass condition.

If the inaudibility of high-frequency cues is the main determinant of the localization performance in noise (in the frontal-horizontal plane) of listeners with sensorineural hearing loss, then performance should be similar for listeners with normal-hearing listening to low-pass filtered clicks and for listeners with a steep, high-frequency sensorineural hearing loss listening to unfiltered clicks. Differences between the observed and predicted performance would indicate the presence of factors other than attenuation (i.e., the presence of "distortions," according to the terminology of Plomp, 1978). The present study aimed to verify this prediction made with normal-hearing listeners and low-pass filtered stimuli. A first set of experiments investigated the ability of listeners with a bilateral, symmetrical, high-frequency, sensorineural hearing loss to localize a broadband click train in the frontal-horizontal plane, in quiet and in noise. As in Lorenzi *et al.* (1999), the effects of signal-to-noise ratio and masker location on localization performance were tested. A second set of experiments investigated the relation be-

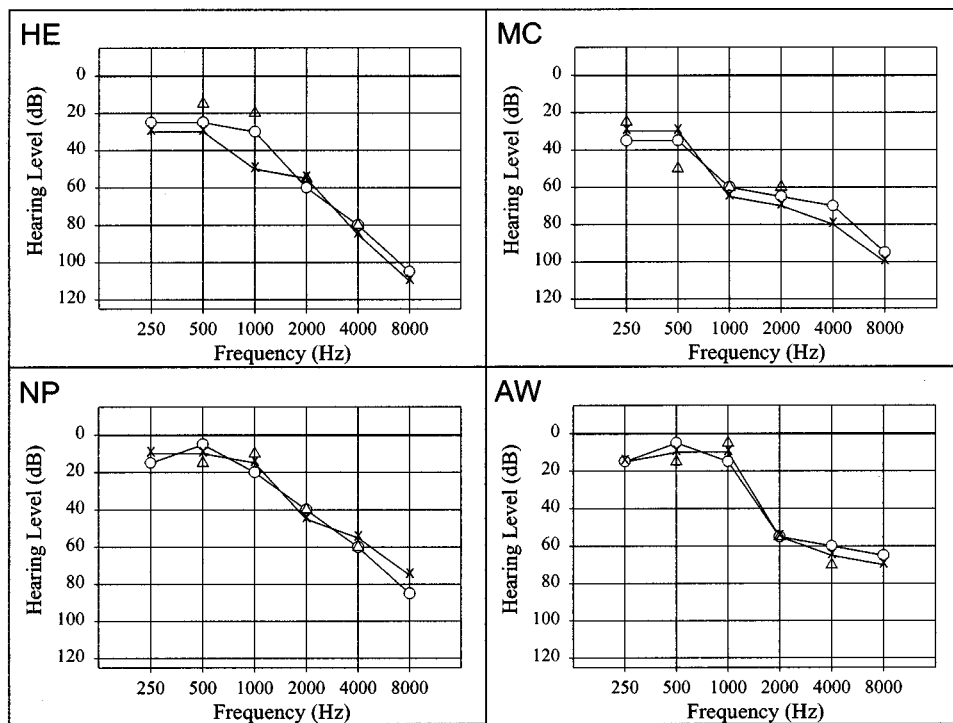


FIG. 1. Audiograms of the four hearing-impaired listeners with sensorineural hearing loss. Continuous lines with circles and "x" show the air-conduction thresholds for the right and left ear, respectively. Triangles show bone-conduction thresholds.

TABLE I. Clinical data for the four hearing-impaired listeners of the study. The table shows: the name and the age of each listener, the duration of hearing loss, and the low-frequency pure-tone average (LFPTA) (mean of HTLs at 0.25, 0.5, and 1 kHz for both ears).

Name	Age	Reported duration of hearing impairment	LFPTA [0.25–1 kHz]
AW	74	30	11.6
HE	52	30	31.6
MC	69	10	42.5
NP	64	39	12.5

tween localization performance and the detectability of the broadband signal as a function of signal and masker location.

## I. METHOD

### A. Listeners

Four hearing-impaired listeners (AW, HE, MC, and NP) participated in the experiments. All hearing-impaired listeners were male and ranged in age between 52 and 74 years. They were selected with the constraints that their hearing loss was sensorineural, bilateral, and symmetrical, and that they showed a steep high-frequency hearing loss. The interaural difference level was less than 9 dB averaged across the frequencies 0.5, 1, 2, and 4 kHz. All four listeners had normal hearing below 1 kHz. Their audiograms are presented in Fig. 1, and their low-frequency pure-tone binaural average (LFPTA) (mean of HTLs at 0.25, 0.5, and 1 kHz for both ears), age, and the duration of hearing loss are presented in Table I. Listeners AW, HE, and NP were diagnosed with unknown sensorineural-hearing losses, and listener MC was diagnosed with noise-related sensorineural-hearing loss.

Four listeners with normal hearing and no reported history of hearing impairment also took part in the experiment: three male (CH, PA, and TO) and one female (CA). Some of their results are reported in Lorenzi *et al.* (in press). These listeners ranged in age from 23 to 33 years. Hearing-threshold levels did not exceed 15 dB at any octave frequency from 0.25 to 8 kHz.

### B. Apparatus and stimuli

The experiment was identical to that described in Lorenzi *et al.* (1999). Sound localization ability was assessed by means of a 1.25-m radius hemisphere placed in a sound-treated chamber (3-m long, 3-m wide, 2-m high). Eleven speakers were mounted on the inner surface of the hemisphere, at 0 deg elevation along the horizontal plane. Speakers were 18 deg apart in each plane at azimuths ranging from  $-90$  deg to  $+90$  deg. The surface of the hemisphere was covered with foam so as to attenuate reflections. The speakers' frequency responses were within  $\pm 1.5$  dB SPL over the range 150 Hz to 20 kHz. All SPLs were measured at the listener's head position with the listener absent from the sound field. The gain of each speaker was balanced so that levels emanating from each speaker were within 0.5 dB SPL. Listeners sat in the (theoretical) center of the full sphere, 1.25 m from each speaker, facing the speaker at 0

deg azimuth and 0 deg elevation. The speakers were visible and numbered. (A personal computer and software were used to control the signal presentation.) Listeners responded using a response box with a matrix of buttons configured like the speaker matrix. Listeners were informed that the experiment investigated only localization ability in the frontal-horizontal plane. Because of this instruction, listeners never volunteered that the signal came from behind.

A comparison paper (Lorenzi *et al.*, 1999) describes in some detail the reverberation characteristics of the experimental system (combination of room and hemisphere). Briefly, the ITD and ILD cues for signals presented from the source loudspeakers, as measured in a Kemar mannequin, are sufficiently preserved for localization and the extraneous cues due to reverberation sufficiently reduced at the signal-to-noise ratios where errors become material.

The signal and masker were identical to those used by Lorenzi *et al.* (1999). They were generated digitally at a sampling rate of 44.1 kHz. The signal was a train of 23- $\mu$ s pulses, which were repeated at a rate of 100 Hz. The pulse train had a duration of 300 ms and was shaped with 25-ms cosine ramps. The pulse train was low-pass filtered digitally at 11 kHz in the "broadband" condition, and low-pass filtered at 1.6 kHz in the "low-pass" condition. In each case, rejection slopes were 96 dB/oct. The resulting pulse train was attenuated digitally and presented through one channel of a 16-bit digital-to-analog converter at a sampling rate of 44.1 kHz. The overall level of the signal was fixed to 70 dB SPL (rms), so as to mimic a conversational speech level.

The masker was a 900-ms white noise shaped with 25-ms cosine ramps. It was low-pass filtered digitally at 14 kHz (96 dB/oct rejection slope). Twenty different samples of filtered noise were generated. For each trial, a sample was chosen at random from this set of 20 noise maskers. The selected noise masker was attenuated digitally and presented through a separate channel of the 16-bit digital-to-analog converter used to play the signal. Six noise levels were used: 52, 58, 64, 70, 76, and 79 dB SPL, so that the signal-to-noise ratio ranged from  $-9$  to  $+18$  dB. Masker was presented at either  $-90$  deg, 0 deg, or  $+90$  deg azimuth. The signal and masker were switched to the appropriate speakers. The switching hardware also added the signal and masker when they were spatially coincident. The signal was temporally centered within the masker.

## C. Procedure

### 1. Localization

Listeners were tested using a forced-choice speaker-identification paradigm. Accordingly, they made a forced-choice decision as to which of the 11 numbered speakers was the most likely source of the target sound presented on each experimental trial. Before each experimental session, listeners were instructed to look straight ahead and to hold their heads and eyes steady during a stimulus presentation. More specifically, they had to look at the speaker set at 0 deg azimuth and 0 deg elevation. Between trials, they were allowed to move their heads in order to look more precisely at the speakers and give their answer on the matrix of buttons

put on their laps. The experimenter could observe the listeners, and check that listeners kept their heads steady.

An experimental session consisted of three separate listening sessions, one for each of the three masker locations ( $-90$  deg,  $0$  deg, or  $+90$  deg azimuth). A listening session comprised 40 experimental runs. Within a run, there were 77 trials, one presentation of the 11 speakers at each of the seven noise levels (including the “in quiet” condition) in random order. Within a run, successive target sounds were varied at random through  $\pm 1$  dB from the nominal level to minimize any (residual) absolute level cues associated with a particular speaker which might be used for source identification. The intertrial interval was fixed at 1.5 s. No feedback was given during the run. To eliminate the effects of training, the order of experimental sessions was randomized among listeners. In addition, ten training runs were provided at the beginning of each experimental session.

## 2. Detection

In a second set of experiments, the audibility of the broadband clicks was assessed by measuring the sensitivity ( $d'$ ) of each impaired listener using a blocked procedure for four combinations of the signal (S) and the masking noise (N) locations, referred to as  $N_0S_0$ ,  $N_0S_{90}$ ,  $N_{90}S_{90}$ , and  $N_{90}S_{-90}$  (see Fig. 1 in Lorenzi *et al.*, 1999). The broadband click-train signal and the white-noise masker were identical to those described in Sec. I B. The noise masker was either presented straight ahead or at the side, and signal and noise were either spatially coincident or maximally separated. In each condition and for each listener, we measured 5-point psychometric functions for the detection of the signal, using a single interval, yes/no task based on 50 trials per point. The gated noise masker was present in each trial. The signal was present in 50% of the trials. The silent interval between trials was 1.5 s. The five most adverse signal-to-noise ratios used in the localization experiments were involved in the detection experiment. Listeners received no training or feedback during this experiment.

## D. Data analysis

For each experimental condition, overall localization performance was estimated by computing:

- (1) the root-mean-square (rms) error,  $D$ , which corresponds to the rms average of the difference between the azimuth of the source and the listener's response. The run rms error ( $\bar{D}$ ) corresponds to the rms average across speakers of the rms error for each speaker.  $D$  and  $\bar{D}$  are used to quantify a listener's accuracy in localizing sound sources. In the present experiment,  $\bar{D}$  should vary between  $0$  deg (perfect localization) and  $80.4$  deg (random guessing).
- (2) The proportion of variance accounted for by the best-fitting linear relation between the perceived angles and the actual source angles,  $r^2$ . This statistic indicates the variability of listeners' choices.  $r^2$  can be considered primarily as a measure of localization consistency, though with a potential contribution from localization accuracy.

It was introduced by Good and Gilkey (1996) because it is less sensitive to systematic biases in the pattern of responses than the root-mean-square error.

In the data analysis, we decided to use the  $\bar{D}$  statistic as the primary measure of overall localization accuracy, and the  $r^2$  statistic for confirmatory purposes. It should be kept in mind that the  $\bar{D}$  statistic includes the contribution of systematic response bias, while the  $r^2$  statistic does so to a lesser extent (Good and Gilkey, 1996). Any effects which might be evident in *both* the  $\bar{D}$  and  $r^2$  statistics are therefore unlikely to be simple consequences of the response biases that individual listeners might have adopted. This is particularly important given the restricted range of responses available to the listeners in the experiment.

## II. RESULTS AND DISCUSSION

### A. Localization data

#### 1. Hearing-impaired listeners

Figure 2 shows a selection of results obtained in the broadband condition for one hearing-impaired listener (MC). In each panel, the response angle is plotted as a function of the signal angle, for one of three signal-to-noise ratios (in quiet,  $0$  dB S/N, and  $-9$  dB S/N), and for one of the three masker locations ( $-90$  deg,  $0$  deg, and  $+90$  deg azimuth). The three top panels therefore represent three replications of the same condition (obtained in different runs). The area of each symbol on the scatter plot represents the number of responses made at that angle for any given signal location. Each vertical column then sums to 40 repeated measures. Ideal performance would be represented by all points lying on the major diagonal. In quiet, most responses fall close to the major diagonal of each panel, except those obtained for signal angles near  $-90$  deg and  $+90$  deg azimuth. This indicates that localization performance is good near  $0$  deg azimuth, but decreases in the spatial periphery. Figure 2 also shows that responses diverge from the major diagonal at  $0$  dB S/N, and are only weakly related to the actual signal angle at  $-9$  dB S/N. For these two signal-to-noise ratios, localization performance appears poorer when noise is at the sides than when noise is straight ahead.

Figure 3 shows similar representations for the responses of the four hearing-impaired listeners, for the three masker locations, at a single (adverse) signal-to-noise ratio of  $-6$  dB. We used the  $-6$  dB S/N ratio to represent the data, as it was the most appropriate experimental condition to show the patterns of biases of the four hearing-impaired listeners. At  $-9$  dB S/N (the most adverse signal-to-noise ratio used here), localization performance of hearing-impaired listeners is close to chance levels and biases are less apparent. These scatter plots reveal in more detail the presence of response biases in two out of the four hearing-impaired listeners (HE and MC), although the patterns of biases appear highly idiosyncratic. Responses at azimuths close to that of the masker dominate in the distributions of responses of listener HE, which could be interpreted as a “pulling effect” similar to that reported by Butler and Naunton (1964). A similar type of bias was observed in normal-hearing listeners by Good



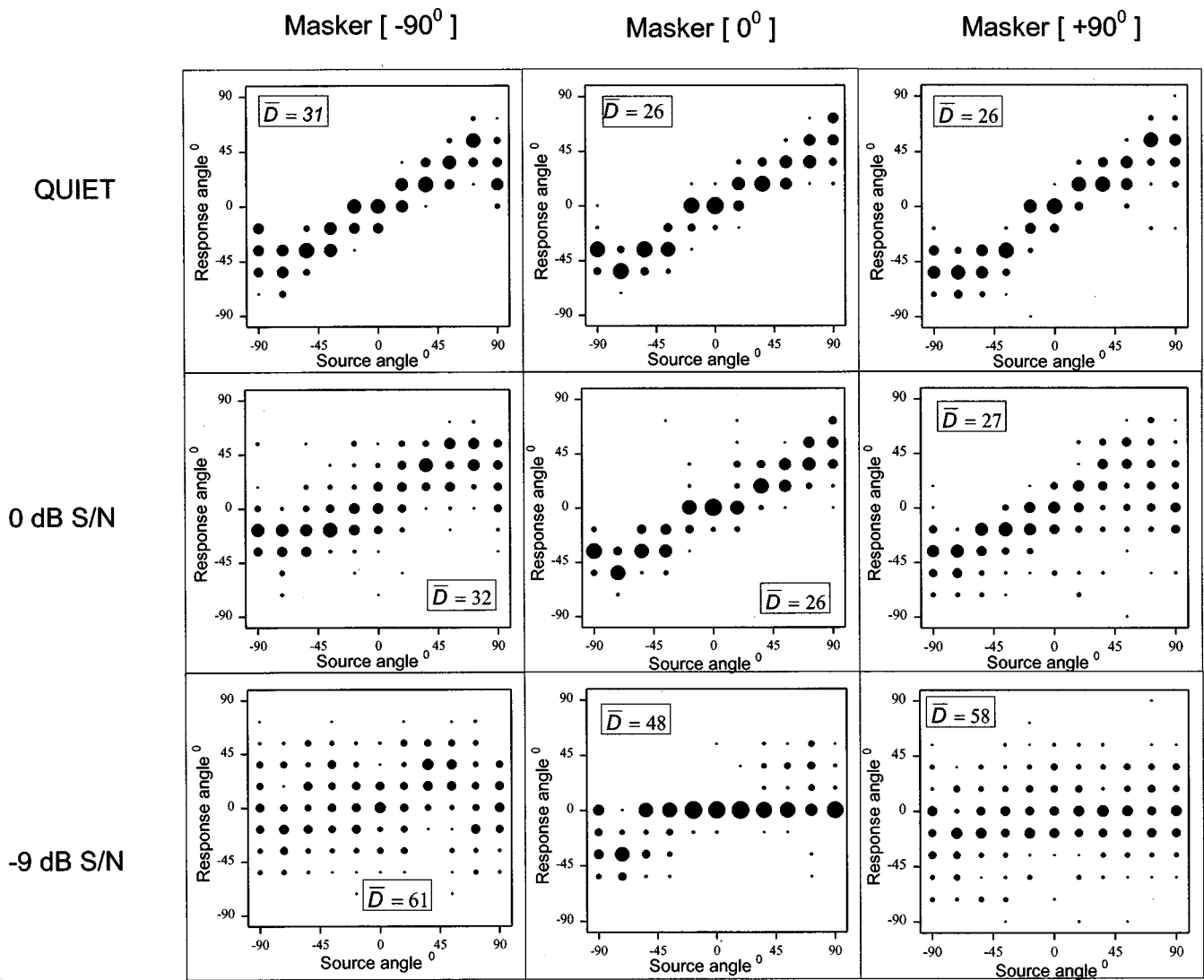


FIG. 2. Localization data for one hearing-impaired listener (MC) at three signal-to-noise ratios [in quiet (top panels), at 0 dB S/N (middle panels), and at -9 dB S/N (bottom panels)], and three masker locations: -90 deg azimuth (left panels), 0 deg azimuth (middle panels), and +90 deg azimuth (right panels), using broadband clicks. In each panel, the response angle is plotted as a function of the actual source angle. The area of each symbol on the scatter plot represents the number of responses made at that angle for any given source location. Values of  $\bar{D}$  (rms error) are also shown in each panel.

and Gilkey (1996) and Lorenzi *et al.* (1999). A different pattern occurs in hearing-impaired listener MC. In this listener, responses are clustered around 0 deg azimuth without regard to masker location, which suggests that MC adopted the rule of answering systematically “straight ahead” when uncertainty was present. Taken together, the present data and those reported in a comparison paper (Lorenzi *et al.*, in press) reveal the presence of a variety of responses biases for both normal-hearing and hearing-impaired listeners, such as systematic shifts in the perceived angle towards or away from the masker, or systematic shifts towards 0 deg azimuth. The idiosyncrasy of the biases therefore suggests (but does not demonstrate) that these systematic biases may have a cognitive rather than a purely sensory origin.

Localization accuracy ( $\bar{D}$ ) and (primarily) consistency ( $r^2$ ) were measured with the broadband click train for each hearing-impaired listener. The data are presented in Figs. 4 and 5, which show  $\bar{D}$  and  $r^2$  as a function of signal-to-noise ratio for each masker location, respectively. In each panel, each curve corresponds to the data of a given listener and the

unconnected symbol corresponds to the “in quiet” condition. For comparison, the data for hearing-impaired listeners are plotted along with the data obtained by Lorenzi *et al.* (1999) in identical conditions with normal-hearing listeners. Filled symbols represent the data obtained with the four hearing-impaired listeners of the present experiment, and unfilled symbols represent the data for the four normal-hearing listeners of Lorenzi *et al.*'s study. Overall, the ability to localize the click train is poorer in hearing-impaired listeners than in normal-hearing listeners. Figures 4 and 5 also show that the masker location has a greater effect on localization accuracy and consistency in hearing-impaired listeners than in normal-hearing listeners. When noise is straight ahead, the data of the four hearing-impaired listeners are relatively consistent: when the signal-to-noise ratio is above 0 to 6 dB, localization accuracy and consistency are not strongly affected by noise. Below 0 to 6 dB S/N, however, localization accuracy and consistency in hearing-impaired listeners decrease and reach a relatively low level (45 deg <  $\bar{D}$  < 50 deg; 0.25 <  $r^2$  < 0.5); in contrast, localization accuracy and consis-

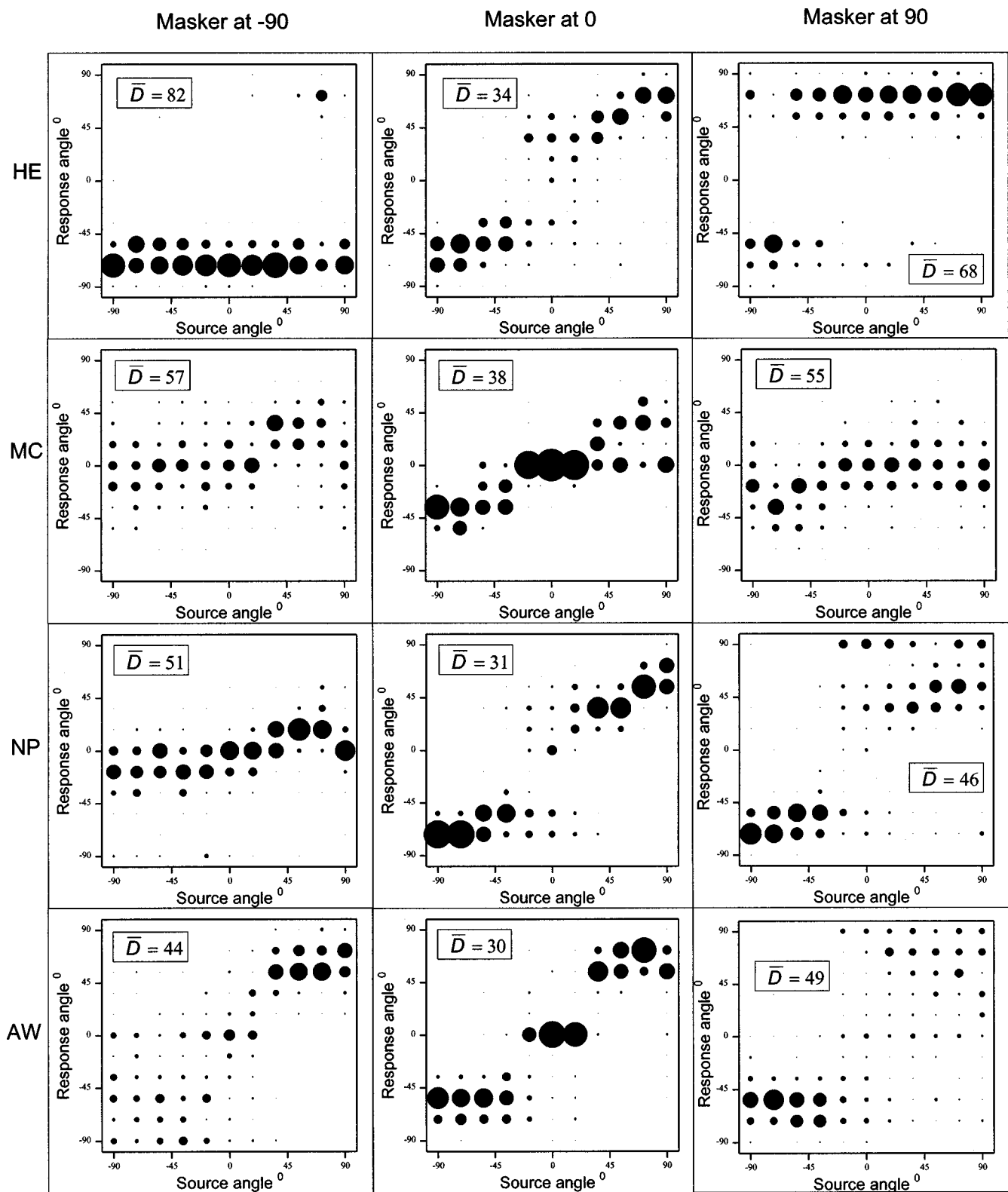


FIG. 3. Localization data for each hearing-impaired listener and for the three masker locations:  $-90$  deg azimuth (left panels),  $0$  deg azimuth (middle panels), and  $+90$  deg azimuth (right panels). The data were obtained with broadband clicks at a signal-to-noise ratio of  $-6$  dB. Values of  $\bar{D}$  (rms error) are also shown in each panel. Other details are as in Fig. 2.

tency in normal-hearing listeners are only slightly affected by noise ( $\bar{D} < 35$  deg;  $r^2 > 0.75$ ). When noise is presented at the sides of the listener, the localization data are less consistent across hearing-impaired listeners. In both quiet and noise, hearing-impaired listeners HE and MC (filled squares

and circles, respectively) show poorer-than-normal localization accuracy and consistency, while hearing-impaired listeners AW and NP (filled diamonds and triangles, respectively) show localization judgments relatively similar to those of normal-hearing listener PA (unfilled triangles). Localization

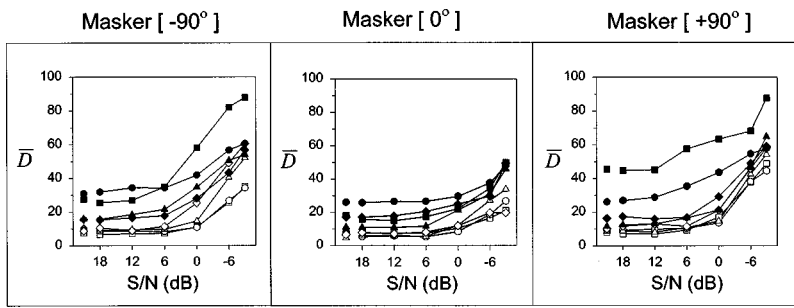


FIG. 4. Values of  $\bar{D}$  (rms error) are plotted as a function of signal-to-noise ratio for each normal-hearing and hearing-impaired listener and for the three masker locations:  $-90$  deg azimuth (left panels),  $0$  deg azimuth (middle panels), and  $+90$  deg azimuth (right panels). The data were obtained with broadband clicks. The unconnected symbols correspond to the "in quiet" condition. Normal-hearing listeners: Unfilled circles (CH), unfilled squares (CA), unfilled triangles (PA), unfilled diamonds (TO). Hearing-impaired listeners: Filled circles (MC), filled squares (HE), filled triangles (NP), filled diamonds (AW).

judgments remain more accurate and more consistent in hearing-impaired listener HE when noise is at  $-90$  deg azimuth than when at  $+90$  deg azimuth, while localization performance decreases relatively similarly in the other three hearing-impaired listeners when noise is at  $-90$  deg and  $+90$  deg azimuth. Such an asymmetry was also observed in two out of the four normal-hearing listeners (Lorenzi *et al.*, in press). Figures 4 and 5 also show that, when noise is presented at the sides of the listener, localization accuracy and consistency begin to decrease at a higher signal-to-noise ratio in the four hearing-impaired listeners compared to the normal-hearing listeners; the decrease begins at 6 to 18 dB S/N in hearing-impaired listeners, compared to 0 to 6 dB S/N in normal-hearing listeners.

Figures 4 and 5 show that, for listeners HE and MC, the accuracy and consistency of localization judgments in quiet are affected by the context in which they are measured. For listener HE, for example,  $\bar{D}$  measured in quiet is lower when measured within an experimental session for which noise is at  $0$  deg azimuth ( $\bar{D}=18.3$  deg) than when measured within sessions for which noise is at  $-90$  deg and  $+90$  deg azimuth ( $\bar{D}=27.5$  deg and  $45.5$  deg, respectively);  $r^2$  measured in quiet is higher when measured within an experimental session for which noise is at  $0$  deg azimuth ( $r^2=0.9$ ) than when measured within sessions for which noise is at  $-90$  deg and  $+90$  deg azimuth ( $r^2=0.62$  and  $0.78$ , respectively). This so-called "context effect" is unlikely to be due to a training effect as the order of experimental sessions was randomized between listeners: noise was presented at  $0$  deg, then  $+90$  deg, and then  $-90$  deg azimuth for listener HE, and at  $+90$  deg, then  $-90$  deg, and then  $0$  deg for listener MC.

## 2. Hearing-impaired listeners compared to normal-hearing listeners

Both low-frequency (primarily ITD) cues and high-frequency (primarily ILD and spectral shape) cues were available to normal-hearing listeners because of the broad amplitude spectrum of the click-train signal. However, low-

frequency cues were likely to be the only cues available to impaired listeners because of their steep, high-frequency hearing loss. This suggests that the poorer-than-normal localization performance observed in hearing-impaired listeners was caused by the inaccessibility of high-frequency cues. We tested this assumption by comparing the present localization data for hearing-impaired listeners to those measured by Lorenzi *et al.* (1999) in normal-hearing listeners with low-pass filtered clicks. Figures 6 and 7 present these data when noise comes from straight ahead (left panels) or from the right of the listener (right panels). First, Figs. 6 and 7 indicate that, when noise is at  $0$  deg azimuth and the signal-to-noise ratio is greater than  $0$  dB, localization accuracy and consistency are relatively similar for normal-hearing listeners and hearing-impaired listeners AW, HE, and NP. However, the hearing-impaired listener MC still shows worse localization accuracy than all other listeners. Second, Figs. 6 and 7 indicate that, when noise is at  $+90$  deg azimuth, localization accuracy and consistency are similar for normal-hearing listeners and hearing-impaired listeners AW and NP, but listeners HE and MC still show localization accuracy and consistency poorer than all other listeners. These data indicate that the poorer localization performance observed in hearing-impaired listeners when noise is at the side is, at least partly, caused by the inaudibility of high-frequency cues. However, the selective filtering applied to the clicks does not account for the poorer scores of two out of the four hearing-impaired listeners.

An inspection of the clinical data presented in Table I reveals that the hearing-impaired listeners showing the poorer localization accuracy (HE and MC) also show the higher low-frequency HTLs: the LFPTA ranges between  $31$ – $42$  dB HL for listeners HE and MC, and between  $11$ – $12$  dB HL for listeners AW and NP. This suggests a relationship between localization performance and the degree of hearing loss, an issue which has already been addressed by several studies. Pröschel and Döring (1990) compared localization

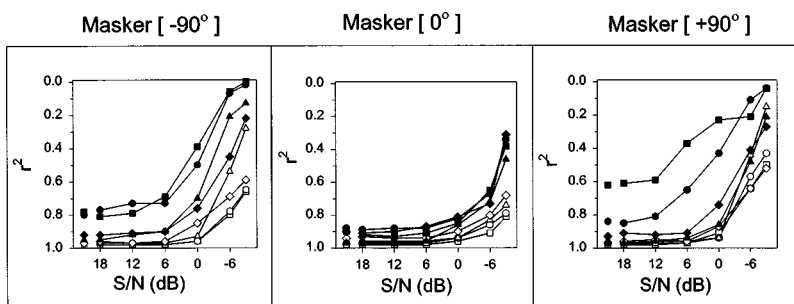


FIG. 5. Values of  $r^2$  are plotted as a function of signal-to-noise ratio for each normal-hearing and hearing-impaired listener and for the three masker locations:  $-90$  deg azimuth (left panels),  $0$  deg azimuth (middle panels), and  $+90$  deg azimuth (right panels). The data were obtained with broadband clicks. Other details are as in Fig. 4.

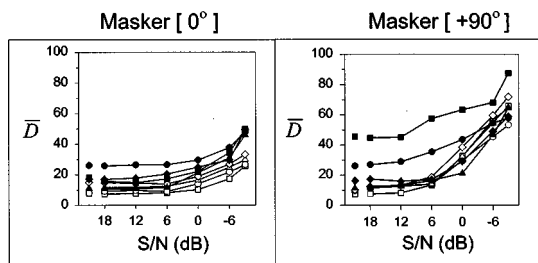


FIG. 6. Values of  $\bar{D}$  are plotted as a function of signal-to-noise ratio for each normal-hearing and hearing-impaired listener and for two masker locations: 0 deg azimuth (left panels) and +90 deg azimuth (right panels). The data were obtained with low-pass filtered clicks in normal-hearing listeners and with broadband clicks for hearing-impaired listeners. Other details are as in Fig. 4.

ability in the horizontal plane in quiet in three groups of listeners with increasing average severity of bilateral, sensorineural hearing-impairment. They found that the group of normal-hearing listeners (HTL < 20 dB through 0.25–6 kHz) had significantly better localization performance than the group with HTLs between 20 and 40 dB; the latter group, in turn, showed significantly better localization performance than the group with HTLs > 40 dB. Noble *et al.* (1994) studied localization ability in different groups of impaired listeners. Compared to the strong relationship reported by Pröschel and Döring, Noble *et al.* found only a modest correlation of 0.4 between frontal-horizontal plane accuracy and low-frequency HTLs in listeners with bilateral, symmetrical, sensorineural impairment. This modest correlation indicated that factors other than attenuation may affect localization ability.

### 3. Supplementary experiment with attenuated signal and noise

In Lorenzi *et al.* (1999), localization ability was measured in four normal-hearing listeners with a low-pass filtered signal. In the rationale of the present paper, we regarded these results as a prediction (or a simulation) of the effects of a high-frequency hearing loss on localization ability. By using this method, we also assumed that hearing-impaired listeners have low-frequency HTLs within the normal range. As pointed out above, this is the case for listeners AW and NP, but not for listeners HE and MC. In the following experiment, we studied the effect of a low-frequency

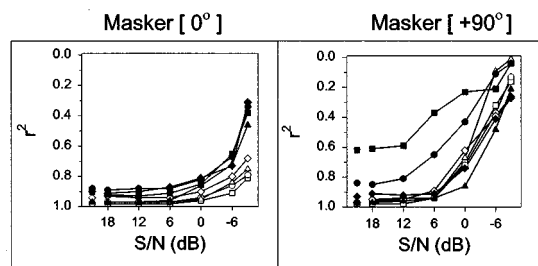


FIG. 7. Values of  $r^2$  are plotted as a function of signal-to-noise ratio for each normal-hearing and hearing-impaired listener and for two masker locations: 0 deg azimuth (left panels) and +90 deg azimuth (right panels). The data were obtained with low-pass filtered clicks in normal-hearing listeners and with broadband clicks for hearing-impaired listeners. Other details are as in Fig. 4.

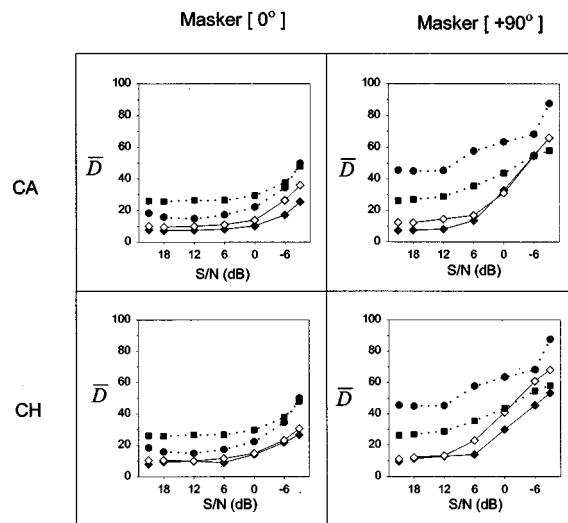


FIG. 8. Values of  $\bar{D}$  are plotted as a function of signal-to-noise ratio for two normal-hearing listeners, CA (top panels) and CH (bottom panels), and for two masker locations: 0 deg azimuth (left panels) and +90 deg azimuth (right panels). The signal was a click train low-pass filtered at 1.6 kHz. In each panel, the solid lines with filled diamonds show the data obtained with a signal presented at 70 dB SPL and signal-to-noise ratios ranging from -9 to 18 dB, the solid lines with unfilled diamonds show the data obtained with the same stimuli attenuated by 20 dB. The data of MC (dotted lines with filled circles) and HE (dotted lines with filled squares), are replotted from Fig. 4 for comparison.

hearing loss by measuring localization accuracy in two out of the four normal-hearing listeners tested by Lorenzi *et al.* (CA and CH) with the low-pass filtered signal and the noise subjected to an attenuation of 20 dB. The results are presented in Fig. 8. In each panel, the localization data obtained by Lorenzi *et al.* (in press) for a given normal-hearing listener are plotted, along with the localization data obtained with the same stimuli attenuated by 20 dB. For comparison, the data are also plotted along with the localization data of the two hearing-impaired listeners HE and MC. Figure 8 shows that the 20-dB attenuation produces only small effects on localization accuracy. Moreover, the localization judgments of CA and CH remain more accurate than the judgments of HE and MC, at least for signal-to-noise ratios above 0 dB S/N. Although this result does not demonstrate that the poorer localization accuracy in listeners HE and MC was independent of audibility in the low-frequency region, it shows that the hypothesis of another cause is at least tenable. Characteristics of hearing impairment in addition to the audibility of low- and high-frequency cues (i.e., “distortion” characteristics) may therefore have caused the poorer localization accuracy of these two hearing-impaired listeners.

The hearing-impaired listeners who participated in this study were not matched for age. Although there is some evidence in the literature that aging can disrupt the encoding of ITD cues (Hermann *et al.*, 1977; Abel and Hay, 1996), it is important to point out that the *youngest* hearing-impaired listener of this study (HE) showed the *poorest* localization performance. Therefore, if age does exert a material influ-

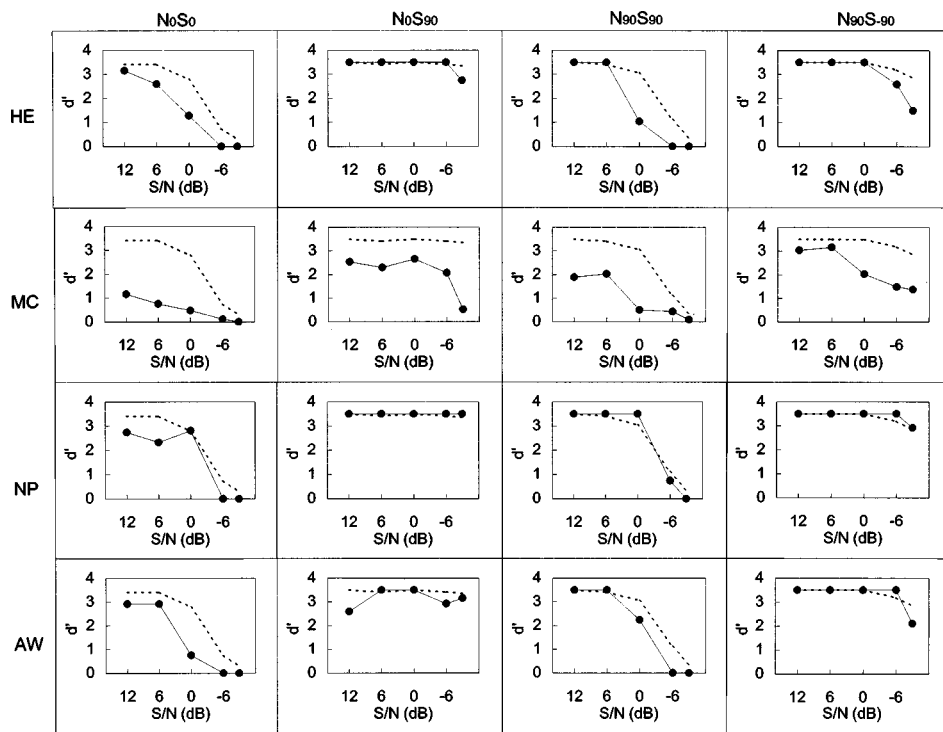


FIG. 9. Psychometric functions showing binaural-detection performance of the broadband click train as a function of signal-to-noise ratio. Each row presents the data for a given hearing-impaired listener. Each column presents the data obtained in one of the four experimental conditions defined by Lorenzi *et al.* (in press) (referred to as  $N_0S_0$ ,  $N_0S_{90}$ ,  $N_{90}S_0$ , and  $N_{90}S_{-90}$ ). In each panel, the data from a given hearing-impaired listener (continuous line with filled circles) are plotted along with the average data from normal-hearing listeners (dotted line without symbol).

ence, the effects that we attribute to hearing impairment would be an underestimate of its actual influences.

## B. Detection data

Because of binaural masking-level differences (BMLD, i.e., the changes in signal detectability as a function of signal and masker interaural time configurations) and head-shadow effects (i.e., the contralateral attenuation of high-frequency sounds performed by the head), the detectability of the clicks is not constant across azimuths but increases with the spatial separation between clicks and noise. Previous studies reported smaller BMLD in hearing-impaired listeners than in normal-hearing listeners (Quaranta and Cervellera, 1974; Hall *et al.*, 1984; Staffel *et al.*, 1990). This suggests that the poorer localization accuracy shown by hearing-impaired listeners may have been caused by poorer detectability of the clicks rather than by a deficit in the auditory localization function. To address this issue, the sensitivity ( $d'$ ) to the presence of the broadband clicks was measured as a function of signal-to-noise ratio in the four experimental conditions specified by Lorenzi *et al.* (1999; see Fig. 1), using a blocked procedure. A limited number of both signal and nonsignal trials ( $n=50$ ) was used. To compute a finite value for  $d'$  in the case of perfect performance (50/50), hit rate was artificially adjusted to 49/50 and false-alarm rate to 1/50. This process results in a ceiling value of  $d'=3.5$ . The psychometric functions of the four hearing-impaired listeners are presented in Fig. 9. In each panel, the dotted line shows the average psychometric function measured in four normal-hearing listeners in identical conditions (replotted from Lorenzi *et al.*, 1999). Figure 9 shows that the detection performance of listeners NP and AW is relatively close to that measured in normal-hearing listeners for the four experi-

mental conditions. Listener HE shows poorer detection performance than listeners NP and AW only when noise is at the side. Listener MC shows poorer detection performance than the other three listeners for the four experimental conditions. As pointed out above, listeners HE and MC show the poorer localization performance and have the higher low-frequency HTLs. The detection performance of hearing-impaired listeners is therefore consistent with both their degree of hearing loss and their localization ability. Distortion characteristics were suggested to explain the localization data of hearing-impaired listeners HE and MC. The present detection data suggest two ways in which such distortions could affect localization: (1) distortions (produced by the damaged cochlea) degrade the detectability of the signal which, in turn, degrades the accuracy of localization judgments (i.e., poor localization accuracy reflects only the small number of speakers for which the clicks are above detection threshold), and (2) distortions degrade ITD cues and the detectability of the signal, and both degrade localization accuracy.

## III. CONCLUSIONS

This paper describes the ability of hearing-impaired listeners to locate and detect a broadband signal in the frontal-horizontal plane, in quiet and in the presence of a noise masker. The hearing-impaired listeners showed a bilateral, symmetrical, sensorineural hearing loss in the high frequencies. The experimental tested the effects of signal-to-noise ratio and masker location on localization accuracy and consistency.

The results show that, as for normal-hearing listeners, hearing-impaired listeners adopt different response patterns (“biases”) at the more adverse signal-to-noise ratios. More-

over, localization accuracy and consistency of hearing-impaired listeners are affected by the masker location in the same way as for normal-hearing listeners: the disruptive effect of the interfering noise is greater when it is located at the side of the listener than straight ahead. However, when noise is at the side, localization accuracy and consistency begin to degrade at higher signal-to-noise ratios in hearing-impaired listeners compared to normal-hearing listeners, and the overall localization accuracy and consistency are poorer in two out of the four hearing-impaired listeners of the study. The rapid decrease in localization performance can be reproduced with normal-hearing listeners by low-pass filtering the click-train signal. However, the poor overall localization performance observed in the two hearing-impaired listeners cannot be fully replicated by altering the audibility of either high- or low-frequency cues. Without a direct demonstration, this result suggests that characteristics of hearing impairment other than audibility (i.e., distortion characteristics) may also affect localization ability in both quiet and noise. The results of detection experiments performed with the broadband signal and masker indicate that the ability to hear the signal is reduced in the hearing-impaired listeners showing the poorer localization performance. Distortion characteristics and reduced detectability in noise are therefore likely to have caused the poorer-than-normal localization ability of impaired listeners.

In conclusion, the general result of the localization experiments indicates that, for broadband stimuli, the accuracy and consistency of localization judgments are poorer for hearing-impaired listeners than for normal-hearing listeners. This supports the idea that, in daily living, patients with cochlear damage experience greater difficulties than normal-hearing listeners when they have to locate a sound source in a noisy background or in a competing situations.

## ACKNOWLEDGMENTS

The authors are particularly indebted to M. A. Akeroyd for his advice during the course of this study. The authors also wish to thank M. P. Haggard and Q. Summerfield for comments on a previous version of this paper, and W. M. Hartmann and B. C. J. Moore for helpful discussions. We would like to acknowledge the comments of the Associate Editor, Dr. W. Grantham, Dr. Fred Wightman, and two anonymous reviewers on an earlier version of this manuscript.

- Abel, S. M., and Hay, V. H. (1996). "Sound localization. The interaction of aging, hearing loss and hearing protection," *Scand. Audiol.* **25**, 3–12.
- Butler, R. A., and Naunton, R. F. (1964). "Role of stimulus frequency and duration in the phenomenon of localization shift," *J. Acoust. Soc. Am.* **36**, 917–922.
- Durlach, N. I., Thompson, C. L., and Colburn, H. S. (1981). "Binaural interaction in impaired listeners: A review of past research," *Audiol.* **20**, 181–211.
- Good, M., and Gilkey, R. H. (1996). "Sound localization in noise: The effect of signal-to-noise ratio," *J. Acoust. Soc. Am.* **99**, 1108–1117.
- Hall, J. W., Tyler, R. S., and Fernandes, M. A. (1984). "Factors influencing the masking level difference in cochlear hearing-impaired and normal-hearing listeners," *J. Speech Hear. Res.* **27**, 145–154.
- Hermann, G. E., Warren, L. R., and Wagener, J. W. (1977). "Auditory lateralization: age differences in sensitivity to dichotic time and amplitude cues," *J. Gerontol.* **32**, 187–191.
- Kistler, D. J., and Wightman, F. L. (1992). "A model of head-related transfer functions based on principal components analysis and minimum-phase reconstruction," *J. Acoust. Soc. Am.* **91**, 1637–1647.
- Lorenzi, C., Gatehouse, S., and Lever, C. (1999). "Localization in noise in normal-hearing listeners," *J. Acoust. Soc. Am.* **105**, 1810–1820.
- Lorenzi, C., Gatehouse, S., and Lever, C. (1997). "Sound localization in noise in hearing-impaired listeners," *J. Acoust. Soc. Am.* **101**, 3104.
- Mills, A. W. (1972). "Auditory localization," in *Foundations of Modern Auditory Theory*, Vol. 2, edited by J. V. Tobias (Academic, New York), pp. 303–348.
- Noble, W., and Byrne, D. (1990). "A comparison of different binaural aid systems for sound localization in the horizontal and vertical planes," *Br. J. Audiol.* **24**, 335–346.
- Noble, W., Byrne, D., and Lepage, B. (1994). "Effects on sound localization of configuration and type of hearing impairment," *J. Acoust. Soc. Am.* **95**, 992–1005.
- Nordlund, B. (1964). "Directional audiometry," *Acta Oto-Laryngol.* **57**, 1–18.
- Plomp, R. (1978). "Auditory handicap of hearing impairment and the limited benefit of hearing aids," *J. Acoust. Soc. Am.* **63**, 533–549.
- Pröschel, U. L. J., and Döring, W. H. (1990). "Richtungshören in der horizontalebene bei störungen der auditiven selektionsfähigkeit und bei seitengleicher innenhorschwerhörigkeit, Teil I, Teil II," *Audiol. Akustik.* **3**, 98–107, 170–177.
- Quaranta, A., and Cervellera, G. (1974). "Masking level differences in normal and pathological ears," *Audiol.* **13**, 428–431.
- Rayleigh, Lord (1907). "On our perception of sound direction," *Philos. Mag.* **13**, 214–232.
- Röser, D. (1966). "Directional hearing in persons with hearing disorders," cited in Durlach, N. I., Thompson, C. L., and Colburn, H. S. (1981). "Binaural interaction in impaired listeners," *Audiol.* **20**, 181–211.
- Staffel, J. G., Hall, J. W., Grose, J. H., and Pillsbury, H. C. (1990). "NoSo and NoS $\pi$  detection as a function of masker bandwidth in normal-hearing and cochlear-impaired listeners," *J. Acoust. Soc. Am.* **87**, 1720–1727.
- Viehweg, R., and Campbell, R. (1960). "Localization difficulty in monaurally impaired listeners," *Ann. Otol. Rhinol. Laryngol.* **69**, 622–634.
- Wightman, F. L., and Kistler, D. (1992). "The dominant role of low-frequency interaural time differences in sound localization," *J. Acoust. Soc. Am.* **91**, 1648–1661.

# Temporal integration of loudness in listeners with hearing losses of primarily cochlear origin<sup>a)</sup>

Søren Buus<sup>b)</sup>

Communication and Digital Signal Processing Center, Department of Electrical and Computer Engineering (409DA), Northeastern University, Boston, Massachusetts 02115

Mary Florentine

Department of Speech-Language Pathology and Audiology (133FR) and Institute for Hearing, Speech, and Language, Northeastern University, Boston, Massachusetts 02115

Torben Poulsen

Department of Acoustic Technology (Building 352), Technical University of Denmark, DK-2800 Lyngby, Denmark

(Received 5 November 1998; revised 22 January 1999; accepted 11 March 1999)

To investigate how hearing loss of primarily cochlear origin affects the loudness of brief tones, loudness matches between 5- and 200-ms tones were obtained as a function of level for 15 listeners with cochlear impairments and for seven age-matched controls. Three frequencies, usually 0.5, 1, and 4 kHz, were tested in each listener using a two-interval, two-alternative forced-choice (2I, 2AFC) paradigm with a roving-level, up-down adaptive procedure. Results for the normal listeners generally were consistent with published data [e.g., Florentine *et al.*, *J. Acoust. Soc. Am.* **99**, 1633–1644 (1996)]. The amount of temporal integration—defined as the level difference between equally loud short and long tones—varied nonmonotonically with level and was largest at moderate levels. No consistent effect of frequency was apparent. The impaired listeners varied widely, but most showed a clear effect of level on the amount of temporal integration. Overall, their results appear consistent with expectations based on knowledge of the general properties of their loudness-growth functions and the equal-loudness-ratio hypothesis, which states that the loudness ratio between equal-SPL long and brief tones is the same at all SPLs. The impaired listeners' amounts of temporal integration at high SPLs often were larger than normal, although it was reduced near threshold. When evaluated at equal SPLs, the amount of temporal integration well above threshold usually was in the low end of the normal range. Two listeners with abrupt high-frequency hearing losses (slopes >50 dB/octave) showed larger-than-normal maximal amounts of temporal integration (40 to 50 dB). This finding is consistent with the shallow loudness functions predicted by our excitation-pattern model for impaired listeners [Florentine *et al.*, in *Modeling Sensorineural Hearing Loss*, edited by W. Jesteadt (Erlbaum, Mahwah, NJ, 1997), pp. 187–198]. Loudness functions derived from impaired listeners' temporal-integration functions indicate that restoration of loudness in listeners with cochlear hearing loss usually will require the same gain whether the sound is short or long. © 1999 Acoustical Society of America. [S0001-4966(99)04206-X]

PACS numbers: 43.66.Sr, 43.66.Cb, 43.66.Mk, 43.66.Ba [JWH]

## INTRODUCTION

The purpose of the present study is to investigate temporal integration of loudness in listeners with hearing impairments of primarily cochlear origin. Whereas considerable knowledge exists about the loudness of long-duration sounds in impaired listeners (for review, see Hellman and Meiselman, 1993; Moore, 1995; Moore and Glasberg, 1997), little is known about the loudness of brief sounds. Because most natural sounds are not steady state, but have amplitude peaks

that typically are much shorter than the 50–150-ms integration time generally assumed for loudness (for review, see Scharf, 1978), knowledge of the loudness functions for brief sounds may be important for hearing-aid fitting and for general understanding of impaired listeners' auditory perception. Moreover, our recent studies (Florentine *et al.*, 1996; Buus *et al.*, 1997; Florentine *et al.*, 1998) indicate that temporal integration of loudness may provide rather direct information about the form of the loudness functions for both brief and long sounds.

To our knowledge, Pedersen and Poulsen's (1973) study on temporal integration of loudness in impaired hearing is the only one published. They tested 24 listeners with cochlear impairments due to presbycusis. The results indicated that the impaired listeners' amount of temporal integration—defined as the level difference between equally

<sup>a)</sup>Parts of this paper were presented at the Hearing Aid Research and Development Conference, September 1997, NIDCD/VA Bethesda, MD [Buus *et al.*, *Proceedings of the 2nd Biennial Hearing Aid Research and Development Conference* (NIDCD/VA, Bethesda, MD, Sept. 1997), 42] and at the ASHA meeting, November 1997, Boston, MA [Buus *et al.*, *ASHA Leader* **2**(15), 143].

<sup>b)</sup>Electronic mail: buus@neu.edu

loud short and long tones—was normal at 75 dB SPL and larger than normal at 95 dB SPL. The normal and larger-than-normal temporal integration for loudness found by Pedersen and Poulsen (1973) contrasts with the reduced temporal integration that is generally found in measurements of detection thresholds for brief signals in impaired listeners (e.g., Florentine *et al.*, 1988; Carlyon *et al.*, 1990; for review, see Moore, 1995) and certainly warrants further study.

To expand on Pedersen and Poulsen's (1973) study, the present study aims to investigate how the amount of temporal integration for loudness varies with level in impaired listeners. The effect of level is important because the amount of temporal integration varies markedly with level in normal listeners (e.g., Stephens, 1973; Poulsen, 1981; Florentine *et al.*, 1996, 1998; Buus *et al.*, 1997) and the outcome of comparisons between normal and impaired hearing appears to depend on the SPL (Pedersen and Poulsen, 1973). The outcome may also depend on whether normal and impaired listeners are compared at equal SL or equal SPL, as indicated by the contrast between temporal integration for detection (an equal-SL comparison) and for loudness (compared at equal SPLs). To provide insight into the effects of audiometric configurations, measurements were obtained using listeners with cochlear impairments in a variety of audiometric configurations. To allow comparisons at equal SLs and equal SPLs, temporal integration for detection and temporal integration of loudness was measured as a function of level.

## I. METHOD

### A. Stimuli

The stimuli were tone bursts with equivalent rectangular durations of 5 and 200 ms. The 5-ms tones consisted of a 6.67-ms raised-cosine rise followed immediately by a 6.67-ms raised-cosine fall. The 200-ms tones had a 195-ms steady-state segment between the 6.67-ms rise and fall. These envelope shapes ensured that most of the tone bursts' energy was contained within the critical bandwidth for all test frequencies. Even for the 5-ms tone burst, the energy within the 100-Hz-wide critical band centered at 500 Hz (Zwicker, 1961; Scharf, 1970) was only 1.3 dB less than the overall energy. [The auditory-filter bandwidth at 500 Hz is only about 75 Hz, but loudness of a constant-SPL stimulus remains constant until its bandwidth exceeds the larger critical bandwidth (see Moore and Glasberg, 1986). The present stimuli were chosen with the aim of keeping loudness the same as that for a narrow-band sound. Thus, the critical bandwidth was used for evaluation of the 5-ms tones' spectral splatter.]

The fixed-level stimuli ranged from 5 dB SL to about 120 dB SPL for the 5-ms tone or the highest level that the listener could comfortably tolerate, whichever was lower. Levels were chosen in 5-dB steps up to 30 dB SL and in 10-dB steps at higher SLs. The test frequencies usually were 0.5, 1, and 4 kHz, but one normal listener was also tested at 8 kHz and one impaired listener was tested at 8 kHz instead of 0.5 kHz to obtain measurements at two frequencies with hearing loss.

## B. Procedure

### 1. Absolute thresholds

To obtain a reference for setting the sensation levels, absolute thresholds were measured for each of the stimuli used in the loudness-balance experiment. The thresholds were obtained with an adaptive procedure in a two-interval, two-alternative forced-choice (2I, 2AFC) paradigm. Each trial contained two observation intervals, which were marked by lights and separated by 500 ms. The signal was presented in either the first or the second observation interval with equal *a priori* probability. The listener's task was to indicate which interval contained the signal by pressing a key on a small computer terminal. Two hundred milliseconds after the listener responded, the correct answer was indicated by a 200-ms light. Following the feedback, the next trial began after a 500-ms delay.

A single threshold measurement was based on three interleaved adaptive tracks. For each track, the level of the signal initially was set approximately 15 dB above the listener's expected threshold. It decreased following three consecutive correct responses and increased following one incorrect response. The step size was 5 dB until the second reversal, after which it was reduced to 2 dB. Reversals occurred when successive signal levels changed direction from increasing to decreasing, or vice versa. On each trial, the track was selected at random among the tracks that had not yet ended, which they did after five reversals. The threshold for one track was calculated as the average of the signal levels at the fourth and fifth reversals. One threshold measurement was taken as the average threshold across the three tracks. This procedure converged on the signal level yielding 79.4% correct responses (Levitt, 1971). Three such threshold measurements (for a total of nine tracks) were obtained for each listener and stimulus. The average across all measurements was used as the reference to set the sensation level in the loudness-balance experiments.

### 2. Loudness matches

Loudness matches between 5- and 200-ms tones were obtained with a roving-level adaptive procedure in a 2I, 2AFC paradigm, similar to that used in a recent study (Buus *et al.*, 1998). On each trial, the listener heard two tones separated by 500 ms. The fixed-level tone followed the variable tone or the reverse with equal *a priori* probability. The listener's task was to indicate which sound was louder by pressing a key on a small computer terminal. The response initiated the next trial after a 700-ms delay. No feedback was given.

To reduce biases that may occur when only a single fixed sound is presented in a series of trials (e.g., Florentine *et al.*, 1996), six to 12 interleaved adaptive tracks were used to obtain concurrent loudness matches at three to six levels with both the short and the long tone varied. The track for each trial was selected at random from all tracks that had not yet ended, which they did after nine reversals. This procedure produced a random variation in overall loudness, which forced the listeners to base their responses only on loudness judgments of the two sounds presented in the trial. [For fur-



ther discussion of roving-level loudness-matching procedures, see Buus *et al.* (1997).]

Because no more than six levels could be tested concurrently, the complete range of fixed levels usually was divided into two or three ranges as necessary to encompass each listener's dynamic range. For normal listeners, a low range included fixed levels between 5 and 20 dB SL, a middle range included levels between 25 and 50 dB SL, and a high range included levels between 60 and 90 dB SL. For impaired listeners, the ranges varied among listeners and frequencies to ensure that each block of trials encompassed reasonable ranges of fixed levels for both short and long tones. If two or fewer levels within a range used for normal listeners could be tested for a given impaired listener and frequency, the number of ranges was reduced and the complete range of levels was redistributed to yield blocks with approximately equal numbers of tracks. If fewer than seven levels could be tested for a given listener and frequency, all levels were usually tested in a single block.

Each track began with the variable stimulus set approximately 15 dB below the expected equal-loudness level. (If that level was below threshold, the variable stimulus was set to threshold.) This choice of starting levels ensured that the listener heard some trials in which the short tone was clearly louder and some in which the long tone was clearly louder. For each track, the level of the variable tone was adjusted according to a simple up-down procedure. If the listener indicated that the variable tone was louder than the fixed tone, its level was reduced; otherwise, it was increased. The step size was 5 dB until the fourth reversal, after which it was reduced to 2 dB. This procedure made the variable tone converge towards the level at which it was judged louder than the fixed tone in 50% of the trials (Levitt, 1971). The average level of the last four reversals was used as an estimate of the level at which the variable tone had the same loudness as the fixed-level tone (cf. Jesteadt, 1980). Three such matches were obtained for each listener and condition.

### C. Apparatus

An APR 486/33 PC-compatible computer with a signal processor (TDT AP2) generated the stimuli, sampled the listeners' responses, and executed the adaptive procedures. The tone bursts were generated digitally with a 50-kHz sample rate and reproduced by a 16-bit digital-to-analog (D/A) converter (TDT DD1). The output from the D/A was attenuated (TDT PA4), low-pass filtered (TDT FT5,  $f_c = 20$  kHz, 190 dB/octave), attenuated again (TDT PA4), and led to a headphone amplifier (TDT HB6), which fed one earpiece of a Sony MDR-V6 headset. This setup ensured that the stimulus level could be controlled with a dynamic range of at least 130 dB. The listeners were seated in a sound-attenuating booth.

For routine calibration, the output of the headphone amplifier was led to a 16-bit analog-to-digital (A/D) converter (TDT DD1), such that the computer could sample the waveform, calculate its spectrum and rms voltage, and display the results before each set of matches. [The SPLs reported below presume a frequency-independent output of 116 dB SPL for an input of 1 V rms. The Sony MDR-V6 headset has a rea-

sonably flat ( $\pm 5$  dB) free-field response across the range of test frequencies. The SPL of 116 dB is close to that measured at 1 kHz in a 6-cc coupler (B&K 4152).]

### D. Listeners

Sixteen ears of 15 listeners with hearing impairments of primarily cochlear origin were tested. The origin of the hearing loss was diagnosed by the listeners' ENT physicians on the basis of a standardized audiological test battery, medical tests, and the listeners' history. Table I shows the diagnostic information for the listeners. According to excitation-pattern models of loudness in impaired listeners (Florentine and Zwicker, 1979; Florentine *et al.*, 1997; Moore and Glasberg, 1997), the form of the loudness function may depend on the impaired listener's audiometric configuration. Certainly, it has been shown that level discrimination in impaired listeners depends on the audiometric configuration in a way that is consistent with predictions of excitation-pattern models (Florentine *et al.*, 1993). Therefore, the impaired listeners were divided into five groups based on the overall configurations of their hearing losses as indicated by the listener IDs shown in the first column. (A sixth group of normal controls is shown at the bottom.) If the hearing loss increased more than 50 dB over any octave, it was characterized as an abrupt loss (listeners A1 and A2); such a steep loss is likely to indicate that inner-hair cells are missing or nonfunctional in some frequency region starting at the beginning of the slope (Hellman and Meiselman, 1993; Florentine *et al.*, 1997). If the hearing loss was not abrupt and decreased 15 dB or more over any two-octave interval, it was characterized as a rising loss (R1 to R3). If the loss was neither abrupt nor rising, it was characterized as flattish if it increased 20 dB or less over all two-octave intervals (F1 to F3), and as falling otherwise. Because eight ears of seven listeners fell into the falling-loss category, they were further subdivided into mild-to-moderate losses (MF1 to MF4) and moderately-severe-to-severe losses (SF1 to SF3).

The second column shows the listeners' genders and the third, their ages. The impaired listeners' ages ranged from 32 to 72 years, with an average of 46 years and a standard deviation of 12.5 years. The test ear is indicated in the fourth column and the listeners' etiologies are shown in the fifth column. Columns six through 15 show the audiometric thresholds in dB HL at the standard audiometric frequencies. The test frequencies used for each listener are indicated by the bold entries, which show that the measurements encompassed conditions with hearing losses ranging from no significant loss to 90 dB HL.

Seven listeners with normal hearing served as age-matched controls. Their ages ranged from 25 to 58 years. The average age was 46 years and the standard deviation 11.7 years. They had no history of hearing difficulties and their audiometric thresholds were 15 dB HL or less at and below 4 kHz; most of the middle-aged listeners had a small threshold elevation at 6 kHz and N7 had a threshold of 25 dB HL at 8 kHz. Listener N1 is the first author and N2 is the third author. They and listener N3 were highly trained listeners; the remaining normal listeners had no prior experience in psychoacoustic experiments. Due to timing constraints,

TABLE I. Diagnostic information for the listeners. (See the text for further information.)

ID	Gender	Age	Ear	Etiology	125	250	500	1 k	1.5 k	2 k	3 k	4 k	6 k	8 kHz
A1	F	51	L	Hereditary, noncongenital	-10	-10	-5	<b>60</b>	60	55	45	<b>45</b>		65
A2	F	61	L	Hereditary	5	10	<b>5</b>	<b>5</b>	35	<b>60</b>	85	90	90	95
R1	F	50	R	Hereditary, noncongenital	-5	5	<b>20</b>	<b>45</b>	50	45	35	<b>30</b>		60
R2	M	37	L	Hereditary Konigsmark	55	50	<b>50</b>	<b>40</b>		35		<b>30</b>	55	55
R3	M	53	L	Unknown	60	75	<b>70</b>	<b>65</b>	65	55	50	<b>45</b>		35
F1	M	36	L	Hereditary		35	<b>40</b>	<b>45</b>		55		<b>55</b>		45
F2	F	33	L	Hereditary	45	50	<b>50</b>	<b>65</b>		65		<b>70</b>		70
F3	F	35	L	Congenital, jaundice		70	<b>70</b>	<b>80</b>		80		<b>90</b>		85
MF1	M	45	R	Unknown	5	10	15	<b>5</b>		0	20	<b>35</b>	50	<b>55</b>
MF2	M	65	L	Hereditary, noncongenital	0	0	<b>15</b>	<b>30</b>	40	40	50	<b>65</b>		75
MF3	F	49	L	Hereditary, noncongenital	0	10	<b>20</b>	<b>35</b>	45	45	45	<b>50</b>		80
MF4	F	40	R	Hereditary, congenital	15	20	<b>45</b>	<b>50</b>		45		<b>50</b>		45
SF1	M	72	L	Unknown, noncongenital	20	25	40	<b>50</b>	55	55	60	<b>60</b>	75	80
SF2	M	32	R	Hereditary, noncongenital	10	20	<b>35</b>	70	75	65	65	75		85
SF3	F	35	R	Sudden, unknown	10	40	<b>60</b>	<b>80</b>		75		<b>70</b>		70
N1	M	45	L	Normal	5	-5	<b>0</b>	<b>0</b>		5	0	<b>0</b>	20	<b>10</b>
N2	M	53	R	Normal	5	0	<b>0</b>	<b>0</b>		0	0	<b>0</b>	30	15
N3	M	25	L	Normal	5	0	<b>5</b>	<b>0</b>		0		<b>-5</b>		5
N4	M	35	R	Normal	5	0	<b>0</b>	<b>5</b>		5	15	<b>15</b>	15	5
N5	F	50	R	Normal	5	5	<b>5</b>	<b>5</b>		0	0	<b>10</b>	0	10
N6	F	53	L	Normal	15	5	<b>0</b>	<b>0</b>		0	10	15	20	15
N7	M	58	R	Normal	5	0	<b>0</b>	<b>5</b>		10		10	20	25

naive normal and impaired listeners practiced at most for 15 minutes, but careful instructions usually rendered practice unnecessary.

### E. Data analysis

For each listener, frequency, and SL, two separate points of subjective equality were calculated for each stimulus pair: the average of the tracks with the short tone fixed and the average of the tracks with the long tone fixed. For each listener and frequency, polynomial fits were then made to the combined data obtained with the long and the short tone varied. Such fits were used because they provided a good description of the entire data set, while avoiding problems of how to average measurements taken at different loudness levels. Visual inspection was used to determine a range of polynomial orders that provided a good fit to the data. The lowest-order polynomial within the range was used unless a higher-order polynomial appeared to follow the data markedly better over some range. The order of the polynomials ranged from two to eight, but third- or fourth-order polynomials were used in most cases. Generally, the fits were quite good as indicated by an average  $r^2$  of 0.968. The resulting polynomials were then used to summarize the amount of temporal integration as a function of level for each listener and frequency by calculating the level difference between the

5- and 200-ms tones as a function of level of the 200-ms tone.

To examine the statistical significance of the effects of stimulus variables and differences among normal listeners, a four-way analysis of variance (ANOVA) (SL of fixed stimulus×frequency×long or short variable×listener) for repeated measures was performed (DATA DESK 6.0.2, Data Description, Inc., Ithaca, NY, 1997). The dependent variable for this analysis was the level difference ( $L_{5\text{ ms}} - L_{200\text{ ms}}$ ) between two equally loud 5- and 200-ms tones. Scheffé *post hoc* tests for contrast (DATA DESK 6.0.2, 1997) were performed when appropriate to explore sources of significant effects and interactions. For all statistical tests, differences were considered significant when  $p \leq 0.05$ .

To examine the effects of stimulus variables and differences among the different audiometric configurations, a four-way ANOVA [SL of fixed stimulus×frequency×long or short variable×group (five levels: abrupt, rising, flat, falling, or normal)] was performed (DATA DESK 6.0.2, 1997) with the level difference ( $L_{5\text{ ms}} - L_{200\text{ ms}}$ ) as the dependent variable. To limit the number of missing data, the analysis included only levels between 5 and 40 dB SL and only frequencies of 0.5, 1, and 4 kHz. Again, Scheffé *post hoc* tests for contrast (DATA DESK 6.0.2, 1997) were used to explore sources of significant effects and interactions.

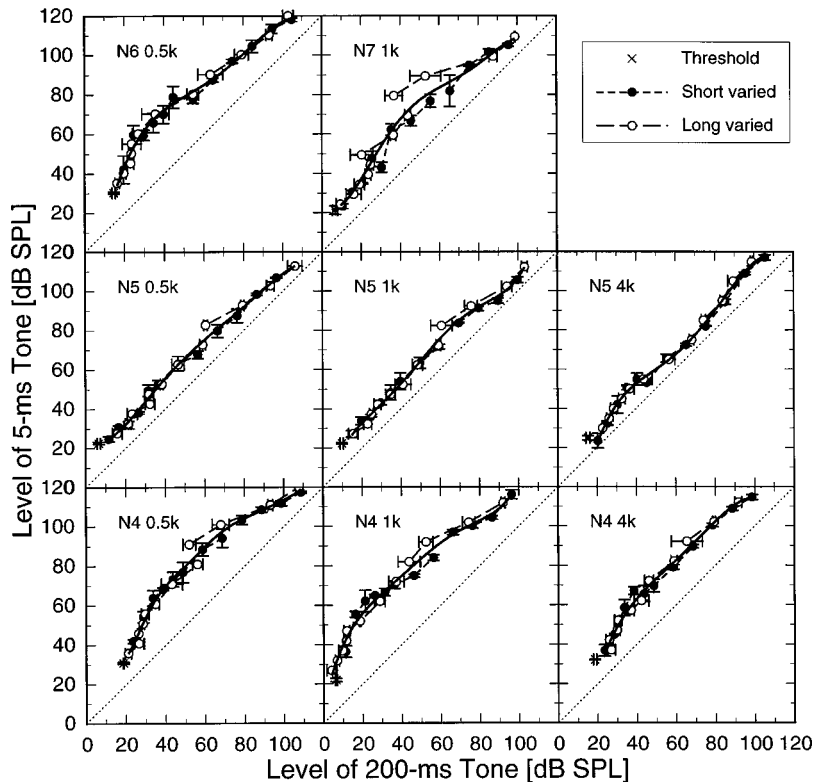


FIG. 1. Loudness-balance judgments by four untrained normal listeners. The level of the 5-ms tone is plotted as a function of the level of the 200-ms tone. The unfilled symbols show data obtained when the 200-ms tones were varied; the filled symbols show data obtained when the 5-ms tones were varied. Each column shows the data for a different frequency. The top row shows the data for listeners N6 and N7, who were tested at only one frequency. The two lower rows of panels show the data for listeners N4 and N5, who were tested at all three frequencies. The error bars show plus and minus one standard error calculated across the three repetitions for each condition. The solid lines show polynomial fits to the data and the dotted lines indicate equal SPLs for the 5- and 200-ms tones. The distance between the dotted and solid lines indicates the amount of temporal integration.

## II. RESULTS

### A. Normal listeners

Figure 1 shows the data obtained for the untrained normal listeners. The listeners' judgments were generally quite consistent, as indicated by the small error bars. The average standard error was 2.3 dB. (It was even smaller for the trained normal listeners, 1.8 dB; across all the normal listeners it was 2.0 dB.) The roving-level procedure appeared successful in reducing biases. The data obtained with the 5-ms tone varied and with the 200-ms tone varied generally lie on a single smooth function, but a few notable exceptions are apparent (e.g., listener N4 at 0.5 and 1 kHz and N7 at 1 kHz). For these data sets, the judgments obtained with the 5-ms

tone fixed at 60 and 70 dB SL (i.e., when the long tone was varied and the loudness level was in the low end of the high range) deviate considerably from the smooth function and show much larger amounts of temporal integration than the other data. This phenomenon also was apparent in several data sets for the trained normal listeners and in some data sets for impaired listeners tested at frequencies with near-normal thresholds. Despite these occasional deviations, a polynomial fitted to the combined data for a single listener and frequency generally summarized the data quite accurately, as indicated by the solid lines. Across all the normal listeners (a total of 18 functions),  $r^2$  ranges from 0.957 to 0.997 with an average of 0.982. In the following, these poly-

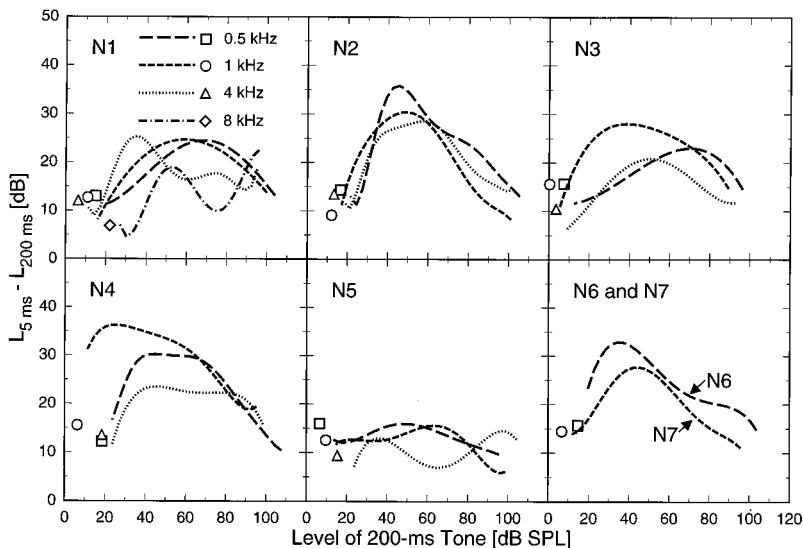


FIG. 2. The amount of temporal integration, defined as the level difference between equally loud 5- and 200-ms tones, for normal listeners is plotted as a function of the SPL of the 200-ms tones. Each panel shows data for a different listener, except that the lower-right panel combines the data for N6, who was tested only at 0.5 kHz, and N7, who was tested only at 1 kHz. The symbols show the data for detection thresholds and the lines show the data for loudness matches, which are derived from the fitted polynomials (see Fig. 1). As indicated by the legend in the top-left panel, different symbols and line types indicate frequency.

nomials are used to characterize how the amount of temporal integration varies with level for each listener and frequency.

Figure 2 shows the amount of temporal integration as a function of level for all the normal listeners. The lines show the level difference between equally loud 5- and 200-ms tones plotted as a function of the level of the 200-ms tones. The threshold data are shown by the symbols. Each panel shows data for a different listener, except that the data for N6 and N7 are combined in the lower-right panel. The amounts of temporal integration for detection range from 7 to 16 dB with no obvious differences among listeners and frequencies, except that the 7-dB value is obtained at 8 kHz (the second-lowest value is 9 dB at 1 kHz for N2). The average amounts of temporal integration for detection are 14.5 dB at 0.5 kHz, 13.3 dB at 1 kHz, and 11.8 dB at 4 kHz.

The amounts of temporal integration for loudness vary considerably more across listeners, but, again, there is no obvious trend across frequency. The temporal-integration functions are quite similar across frequency for some listeners (e.g., N2 and N5) and when clear differences are apparent, they are not consistent across listeners (e.g., N1 and N4). For all listeners (except perhaps N5) and all frequencies, the amount of temporal integration clearly varies nonmonotonically with level and is largest at moderate levels. With a few exceptions, the amount of temporal integration for loudness at low levels approaches that obtained for detection. As the level increases, the amount of temporal integration increases, often dramatically. The maximal amounts of temporal integration vary from 13 dB (N5 at 4 kHz; 14.5 dB if the second maximum at 97 dB SPL is used) to 36 dB (N2 at 0.5 kHz and N4 at 1 kHz). The averages of the maxima are 27 dB ( $\pm 7.3$  dB standard deviation) at 0.5 kHz, 27 dB at 1 kHz ( $\pm 7.7$  dB), and 22 dB ( $\pm 5.9$  dB) at 4 kHz. No obvious relations are apparent among the maximal amount of temporal integration for loudness, the level (SPL or SL) at which the maximum occurs, the amount of temporal integration for detection, and the detection threshold. At high levels, the amounts of temporal integration generally decrease to reach roughly the same values as those obtained for detection.

To provide a comparison for the data from the impaired listeners, the combined data for the normal listeners are shown as functions of SPL in Fig. 3 and SL in Fig. 4. Because no consistent effect of frequency was apparent across listeners, the data are shown together for frequencies of 0.5 (long-dashed lines), 1 (short-dashed lines), and 4 kHz (dotted lines). The two graphs are quite similar because the normal listeners' thresholds encompass only a relatively narrow range of SPLs. Both show that the amounts of temporal integration encompass quite large ranges, especially at moderate levels where the range exceeds 20 dB. The gray areas show the ranges that will be considered as normal in evaluating the data for impaired listeners. Because the set of normal data is relatively large, comprising a total of 17 functions, the two extreme values at any level have been excluded from these "standard" normal ranges. Even with this trimming the standard normal ranges are considerable; at moderate levels they range from about 12 to well over 30 dB of temporal integration. At low and high SPLs, the standard normal range is much narrower, ranging from 7 or 8 dB to

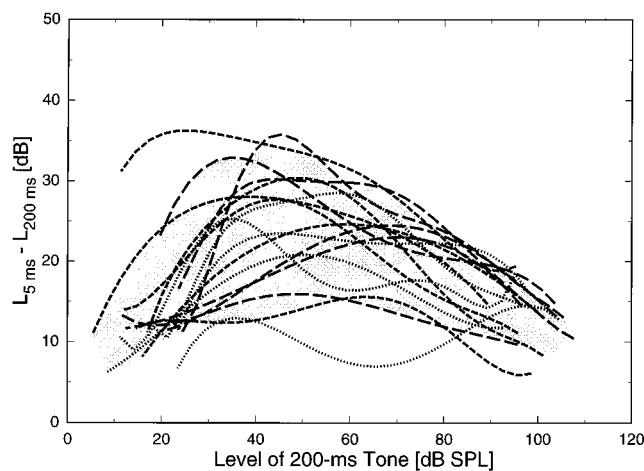


FIG. 3. The amount of temporal integration for loudness is plotted as a function of the SPL of the 200-ms tone. Each line shows the data obtained for a normal listener tested at 0.5 (long-dashed lines), 1 (short-dashed lines), or 4 kHz (dotted lines). The gray area indicates a normal range that will be used as a "standard" against which impaired listeners' data can be evaluated. It encompasses the range from the second lowest to the second highest amount of temporal integration.

about 13 dB (see Fig. 3); at high SPLs it is also relatively narrow, ranging from about 8 to 17 dB, but at low SPLs it ranges from about 7 to 23 dB (see Fig. 4). Clearly, the amount of temporal integration varies widely among normal listeners.

The statistical analyses support these observations. A three-way (frequency  $\times$  duration  $\times$  listener) ANOVA for repeated measures of the thresholds for the listeners who were tested at all three frequencies showed significant effects of duration [ $F(1,4) = 1580, p \leq 0.0001$ ] and listener [ $F(4,60) = 270.5, p \leq 0.0001$ ], as well as significant interactions of listener and frequency [ $F(8,60) = 78.01, p \leq 0.0001$ ] and of listener, frequency, and duration [ $F(8,60) = 7.59, p \leq 0.0001$ ]. No effect of frequency and no interaction between frequency and duration were present. These findings indicate that thresholds differ among listeners and are higher for the 5-ms tones than for the 200-ms tones. On the average, the thresholds do not differ among the three frequencies, but they differ among listeners and among frequencies within

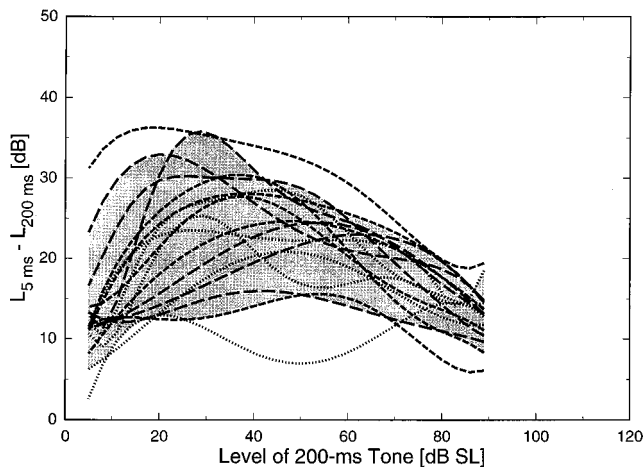


FIG. 4. Same as Fig. 3, except that the amount of temporal integration is plotted as a function of the 200-ms tone's sensation level, SL.

TABLE II. Four-way analysis of variance for repeated measures of loudness matches by the five normal listeners who were tested at all three frequencies. The dependent variable is the level difference between the short and the long tone when they were judged to be equally loud. The stimulus variables frequency (Frq; 3 levels: 0.5, 1, and 4 kHz), sensation level (SL; 12 levels: 5,10,15,20,25,30,40,...,90 dB SL), and variable (Var; 2 levels: short varied or long varied) are fixed factors. Listener (Lsr; 5 levels: N1 through N5) is a random factor.

Source	<i>df</i>	Error <i>df</i>	Sums of squares	Mean square	<i>F</i> -ratio	Prob
Const	1	4	371 191	371 191	98.99	0.0006
Frq	2	8	2 309	1 155	2.630	0.13
SL	11	44	16 879	1 534	7.320	≤0.0001
Frq×SL	22	87	1 306	59.37	1.629	0.058
Var	1	4	160.4	160.4	10.12	0.034
Frq×Var	2	8	22.08	11.04	0.4471	0.65
SL×Var	11	44	3 754	341.3	9.133	≤0.0001
Frq×SL×Var	22	87	492.3	22.38	1.339	0.17
Lsr	4	716	14 999	3 749	226.1	≤0.0001
Frq×Lsr	8	716	3 512	439.1	26.47	≤0.0001
SL×Lsr	44	716	9 223	209.6	12.64	≤0.0001
Frq×SL×Lsr	87	716	3 171	36.45	2.198	≤0.0001
Var×Lsr	4	716	63.42	15.85	0.9560	0.43
Frq×Var×Lsr	8	716	197.6	24.70	1.489	0.16
SL×Var×Lsr	44	716	1 644	37.37	2.253	≤0.0001
Frq×SL×Var×Lsr	87	716	1 454	16.71	1.007	0.46
Error	716		11 874	16.59		
Total	1073		79 641			

listeners. The average amounts of temporal integration differ neither across frequency nor among listeners, but highly reliable differences in temporal integration across frequency are present within listeners.

Table II shows the results of the ANOVA for the amounts of temporal integration calculated from the loudness-balance data. The top half shows the significance of stimulus effects and the bottom half shows the significance

of differences among listeners. Frequency and all interactions with it are not significant. This finding indicates that no consistent effect of frequency is present in the data. In contrast, the effect of SL is highly significant, which reflects the nonmonotonic effect of level on the amount of temporal integration for loudness. The effect of variable (5- or 200-ms tone varied) and the interaction between SL and variable are also significant. These findings indicate that equal SL does

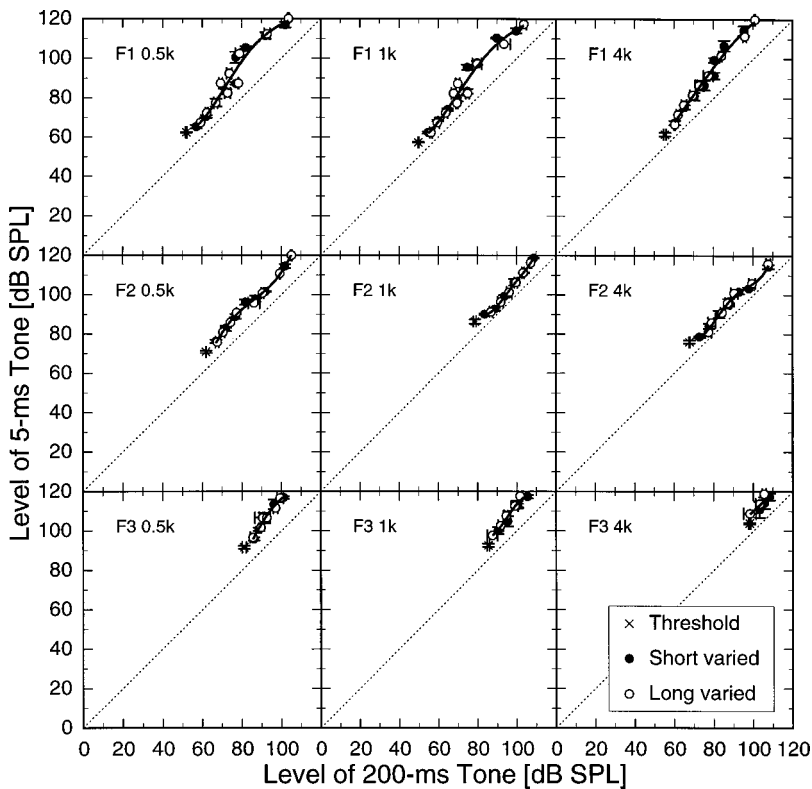


FIG. 5. Same as Fig. 1, but for three listeners with flattish hearing losses.

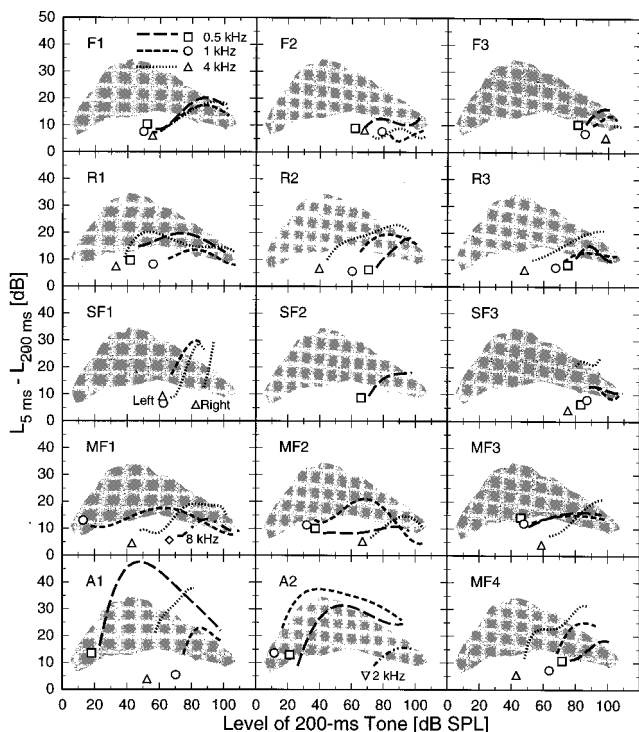


FIG. 6. The amount of temporal integration for loudness obtained for 15 listeners with impairments of predominantly cochlear origin plotted in the same manner as Fig. 2. The gray area in each panel indicates the standard normal range for the amount of temporal integration as a function of SPL (see Fig. 3). Except for the bottom row, each row shows data for listeners with different overall audiometric configurations (see Sec. 1D for the classification criteria). Data for flattish audiograms are shown in the top row, rising audiograms in the second row, moderately severe and severe losses with falling audiograms in the third row, and mild-to-moderate losses with falling audiograms in the fourth row. The bottom row shows data for two listeners with high-frequency losses and abruptly falling audiograms and for the fourth listener with a mild-to-moderate loss and a falling audiogram. Labels by selected functions in the panels for SF1, MF1, and A2 designate test conditions that cannot be identified from the legend in the top-left panel.

not yield equal loudness for short and long tones. Loudness balances with the short tone fixed at some SL are performed at a different loudness level than those with the long tone fixed at the same SL, and the results obtained in the two conditions differ because the amounts of temporal integration depend on loudness level. The bottom half of the table shows that the effects of frequency, SL, and the interaction between frequency and SL vary among listeners. These findings indicate that individual listeners show highly reliable differences in the form of their temporal-integration functions at different frequencies, but as noted above the differences are not consistent across listeners.

Additional analyses of covariance (ANCOVAs) showed that the maximal amount of temporal integration for loudness was uncorrelated with the level at which it occurred, the listeners' threshold at the test frequency, and their temporal integration for detection. Likewise, the level at which the maximal amount of temporal integration occurred was uncorrelated with both the threshold for the 200-ms tone and with the amount of temporal integration for detection. However, a two-way ANCOVA (age $\times$ frequency) revealed that the maximal amount of temporal integration for loudness increased significantly with age [ $F(1,11) = 7.06, p = 0.022$ ],

whereas there was no significant effect of age on the level at which the maximum occurred. The effect of age on the maximal amount of temporal integration for loudness is surprising and may be a spurious result caused by the relatively small amounts of temporal integration obtained for listener N5, who was the second youngest of the normal listeners.

## B. Listeners with cochlear impairments

To provide an example of the data obtained for impaired listeners, Fig. 5 shows the loudness matches obtained for the three listeners with flattish hearing losses. The impaired listeners' judgments were usually very consistent, as indicated by the small error bars. The average standard error for the data in Fig. 5 was 1.3 dB. (The average standard error for all the impaired listeners' data was slightly larger, 1.6 dB, but was still smaller than that obtained for the trained normal listeners.) As for the normal listeners, the data obtained with the 5-ms tone varied and with the 200-ms tone varied generally lie on a single smooth function, except that the data for listener F1 show some separation between the low- and high-level segments of the functions, qualitatively similar to the separation observed in some normal listeners. For this listener, 25 dB SL was included in both sets. The results were the same in both sets when the 4-kHz long tone was varied, but the remaining data at 25 dB SL show 10 dB more temporal integration in the high than in the low set. Listener F2 had 20 dB SL included in both the low and high sets at 4 kHz and listener F3 had 15 dB SL included in both sets at 0.5 and 1 kHz. No discrepancy between the low and high sets is apparent in these conditions.

Despite the occasional deviations between the low and high sets, the fitted polynomials generally summarize the data quite accurately, as shown by the solid lines. Across all the impaired listeners (a total of 43 functions), the average  $r^2$  is 0.962. The range of  $r^2$  is 0.910 to 0.998, except that two functions yielded  $r^2$  of 0.720 (F3 at 4 kHz) and 0.876 (MF2 at 0.5 kHz). The low correlation for F3 at 4 kHz is a result of her six data points encompassing only about a 10-dB range of levels. The low correlation for MF2 at 0.5 kHz results because he shows a relatively large range effect and his data encompassed only about a 55-dB range of levels.

Figure 6 shows the amount of temporal integration as a function of SPL for all the impaired listeners plotted in the same manner as Fig. 2. For comparison, the "standard" normal range from Fig. 3 is shown together with the data for each listener. The amounts of temporal integration for detection by the impaired listeners range from 4 to 14 dB and are mostly lower than those for the normal listeners. Generally, the amount of temporal integration for detection appears less than normal whenever the threshold is elevated. Listener MF3 at 0.5 kHz is the only exception. The impaired listeners' average amounts of temporal integration for detection are 10.0 dB at 0.5 kHz, 8.6 dB at 1 kHz, and 6.0 dB at 4 kHz; these values are about 5 dB lower than those for normal listeners.

The amounts of temporal integration for loudness vary even more for the impaired than for the normal listeners. As for normal listeners, no obvious trend across frequency is apparent and the amount of temporal integration for loudness

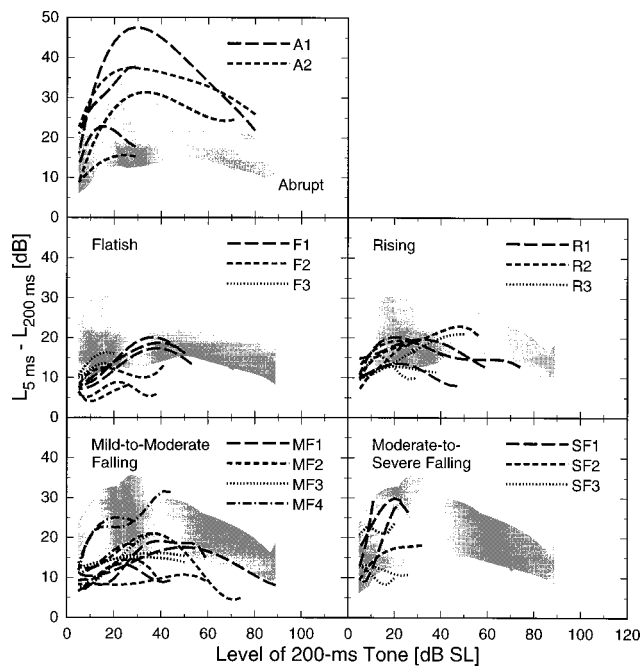


FIG. 7. The amount of temporal integration for loudness as functions of SL in the 15 impaired listeners is plotted in separate panels for each audiometric group. The gray area shows the standard normal range for the amount of temporal integration as a function of SL (see Fig. 4).

at low sensation levels generally approaches that obtained for detection, but a few exceptions may be present (e.g., A1 and SF3 at 4 kHz). For the majority of listeners and frequencies, the amount of temporal integration varies nonmonotonically with level and is largest at some moderately low sensation level. As the level increases above 5 dB SL, the amount of temporal integration usually increases, often dramatically. The maximal amounts of temporal integration vary from 9 (F2 at 4 kHz; 8 dB if the maximum of F2's 1-kHz function is used, but this function does not appear to reach its maximum within the range of levels tested) to 48 dB (A1 at 0.5 kHz). The averages of the maxima are 20 dB ( $\pm 9.7$  dB standard deviation) at 0.5 kHz, 20 dB at 1 kHz ( $\pm 7.4$  dB), and 21 dB ( $\pm 8.4$  dB) at 4 kHz. These averages are 7 dB lower than normal at 0.5 and 1 kHz, and 1 dB lower than normal at 4 kHz. The maximal amount of temporal integration for loudness in impaired listeners does not appear related to the level (SPL or SL) at which it occurs or to the detection threshold, if the data for the listeners with abrupt losses are omitted. It also appears unrelated to the amount of temporal integration for detection. On the other hand, the SPL at which the maximum occurs generally increases with the threshold, whereas the SL of the maximum decreases as the threshold increases.

Except near threshold, the impaired listeners' amounts of temporal integration usually are within the normal range, and at high SPLs they exceed the top of the normal range in about half the conditions tested. A few impaired listeners are clearly different from normal. Listener F2 shows less-than-normal amounts of temporal integration for loudness at almost all SPLs, and the two listeners with abrupt losses, A1 and A2, have obviously abnormal temporal-integration functions. As shown in the bottom row of Fig. 6, these listeners

yield larger-than-normal amounts of temporal integration at the frequencies just below their abrupt losses (0.5 kHz for A1 and 1 kHz for A2). The maximal amount of temporal integration for A1 is nearly 48 dB, which is more than 11 dB above the largest amount of temporal integration obtained for any normal listener at any frequency. A large maximum is also obtained for listener A1 at 4 kHz and for listener A2 at 0.5 kHz. In summary, once the level is somewhat above threshold, the impaired listeners usually show amounts of temporal integration for loudness that are as large as, or sometimes even larger than, that obtained for normal listeners tested at the same SPL. The normal or larger-than-normal temporal integration for loudness contrasts with impaired listeners' reduced temporal integration for detection. As discussed below, this contrast may indicate that the comparison between normal and impaired listeners depends on whether it is made at equal SPLs or equal SLs, because the data for detection necessarily represent an equal-SL comparison.

Whereas the amounts of temporal integration at high SPLs are larger than normal in about half the conditions tested when normal and impaired listeners are compared at equal SPLs, the tendency towards abnormally large temporal integration disappears almost completely if the comparison between normal and impaired listeners is made at equal SLs. This is illustrated in Fig. 7, which shows the amounts of temporal integration obtained for the impaired listeners plotted as a function of SL. Only five functions have segments above the standard normal area. Four of those come from the listeners with abrupt losses. The fifth function is for MF4 at 4 kHz. It exceeds the normal range for a few dB near the highest SLs. On the other hand, 16 of the 43 functions have segments that fall below the normal range, and the central tendency is clearly towards the low end of the normal range for the groups with flattish, rising, and mild-to-moderate falling audiometric configurations. Thus, it appears that when normal and impaired listeners are compared at equal SLs, the amounts of temporal integration are generally normal or, for some groups, slightly less than normal. The clearest exception is listeners with abrupt hearing losses, who show unusually large amounts of temporal integration for loudness whether the comparison to normal listeners is at equal SPL or equal SL.

The preceding figures showed the results with listeners divided according to the overall configuration of their audiograms. Apart from the two listeners with abrupt losses, no clear differences among the groups were apparent. One might hypothesize that a clearer separation among the temporal-integration functions could be obtained by considering the course of the hearing loss above the test frequency. This hypothesis follows from findings that the amount of temporal integration appears to be related to the slope of the loudness function (Florentine *et al.*, 1996, 1998; Buus *et al.*, 1997) and that the growth of loudness depends to some extent on upward spread of excitation (e.g., Hellman, 1974; Florentine *et al.*, 1997). However, grouping the temporal-integration functions according to the local shape of the audiogram as defined by the change in hearing loss over one (for the 4-kHz functions) or two octaves (for lower frequencies) above the test frequency did not improve the order of

TABLE III. Four-way analysis of variance for loudness matches by all the listeners grouped according to the overall configuration of their audiograms. The dependent variable is the level difference between the short and the long tone when they were judged to be equally loud. The independent variables sensation level (SL; 7 levels: 5, 10, 15, 20, 25, 30, and 40 dB SL), frequency (Frq; 3 levels: 0.5, 1, and 4 kHz), variable (Var; 2 levels: short varied or long varied), and audiometric group (AGp; 5 levels: normal, abrupt, falling, flatish, and rising) are fixed factors.

Source	<i>df</i>	Sums of squares	Mean square	<i>F</i> -ratio	Prob
Const	1	596 837	596 837	16 803	≤0.0001
SL	6	18 369	3 061	86.19	≤0.0001
Frq	2	190.2	95.12	2.678	0.069
SL×Frq	12	237.3	19.78	0.5568	0.88
Var	1	1 386	1 386	39.01	≤0.0001
SL×Var	6	731.1	121.9	3.431	0.0023
Frq×Var	2	65.66	32.83	0.9243	0.40
SL×Frq×Var	12	304.5	25.38	0.7145	0.74
AGp	4	20 682	5 170	145.6	≤0.0001
SL×AGp	24	3 941	164.2	4.623	≤0.0001
Frq×AGp	8	3 863	482.8	13.59	≤0.0001
SL×Frq×AGp	48	1 978	41.21	1.160	0.21
Var×AGp	4	986.5	246.6	6.943	≤0.0001
SL×Var×AGp	24	749.4	31.23	0.8791	0.63
Frq×Var×AGp	8	222.0	27.75	0.7811	0.62
SL×Frq×Var×AGp	47	794.4	16.90	0.4758	1.00
Error	2072	73 598	35.52		
Total	2280	148 392			

the data appreciably. Whether the temporal-integration functions were compared at equal SPLs or at equal SLs, the consistency within classes was, at best, only slightly better than that obtained when listeners were grouped according to their overall audiometric configurations. Although this finding may seem counterintuitive, it agrees with the excitation-pattern model for impaired listeners. Upward spread of excitation combined with abnormally fast growth of specific loudness can restore normal growth of loudness, even if the hearing loss limits the spread of excitation to only about one octave (Florentine *et al.*, 1997).

The statistical analyses are consistent with the observations on the SL data. As shown in Table III, the ANOVA showed significant effects of SL and variable (short or long), and a significant interaction of variable and SL. The first effect indicates that the amount of temporal integration varies with sensation level for both normal and impaired listeners and generally increases with SL over the 5-to-40-dB range included in the analysis. The latter effects most likely reflect the findings that short and long stimuli at equal SL generally are not equally loud and that different amounts of temporal integration are obtained at different loudness levels. In addition to these stimulus effects, the effect of overall audiometric configuration and all two-way interactions with the stimulus variables are significant. Some of these effects undoubtedly reflect effects of hearing loss, but because the hearing-impaired groups consist of two to four listeners, the effects may also reflect interlistener differences, which are large.

The Scheffé *post hoc* tests for contrast showed highly significant differences among the groups. Averaged over all stimulus conditions included in the analysis, the listeners with abrupt losses showed significantly larger amounts of temporal integration than the normal listeners, whose tempo-

ral integration was larger than those for the rising and falling losses. The latter two groups did not differ significantly, but their amounts of temporal integration were significantly larger than those for the listeners with flattish losses. The interactions with SL, frequency, and variable show that the differences among listeners depended somewhat on these stimulus variables, but significant differences similar to those between the grand means for each group were present for a number of SLs and at most frequencies.

Additional ANCOVAs on the effects of frequency, threshold for the 200-ms tone, and age on the impaired listeners' amounts of temporal integration for detection showed a significant effect of frequency [ $F(2,29) = 3.568, p = 0.041$ ] and a significant interaction between the threshold for the 200-ms tone and frequency [ $F(2,29) = 3.505, p = 0.043$ ]. The effects of age and interactions with it were not significant. To explore the interaction between the threshold for the 200-ms tone and frequency, separate ANCOVAs were performed on the relation between the threshold and the amount of temporal integration. These analyses showed that the amount of temporal integration decreased significantly as the threshold increased at 0.5 [ $F(1,11) = 28.58, p = 0.0002$ ] and 1 kHz [ $F(1,12) = 10.88, p = 0.0064$ ], whereas the relation did not reach significance at 4 kHz.

An ANCOVA of the impaired listeners' maximal amount of temporal integration of loudness showed no significant dependence on frequency, the threshold for the 200-ms tone, or the SPL at which the maximum occurred when the listeners with abrupt losses were excluded from the analysis. (The low thresholds and large amounts of temporal integration obtained for the abrupt losses at low frequencies produced significant effects of the threshold and SPL at which the maximum occurred.) On the other hand, the SPL at which the maximal amount of temporal integration oc-



curred increased significantly with the threshold for the 200-ms tone at the test frequency, whether the abrupt-loss listeners were included [ $F(1,35)=74.86, p\leq 0.0001$ ] or not [ $F(1,30)=34.68, p\leq 0.0001$ ], and the SL at which the maximum occurred decreased significantly as the threshold for the 200-ms tone increased both with [ $F(1,35)=43.39, p\leq 0.0001$ ] and without [ $F(1,30)=63.74, p\leq 0.0001$ ] the abrupt-loss listeners.

### III. DISCUSSION

#### A. Comparison with other data

The present data for normal hearing are in good agreement with data from previous studies (Florentine *et al.*, 1996, 1998) that used 1-kHz tones with the same temporal parameters as those in the present study. The average amount of temporal integration at 5 dB SL was about 13 dB in the present study, similar to the values of about 12 dB obtained by Florentine *et al.* (1996) and 15 dB obtained by Florentine *et al.* (1998). Likewise, the amount of temporal integration obtained at 90 dB SPL in the present study was 13 dB, which is close to the 11 dB reported by Florentine *et al.* (1996) and 14 dB reported by Florentine *et al.* (1998). The maximal amount of temporal integration of the average data was 26 dB in the present study, which is close to the maximum of 27 dB obtained by Florentine *et al.* (1998) but somewhat larger than the 18 dB obtained by Florentine *et al.* (1996). Given that the individual listeners in the three studies show maxima varying from about 12 to about 36 dB, it is not surprising that the maximal amounts of temporal integration vary somewhat across studies. [It should be noted that N1 and N3 also were tested by Florentine *et al.* (1996). The maximal amounts of temporal integration for these listeners at 1 kHz are about 5 dB larger in the present than in the previous experiment. This discrepancy may reflect differences between the roving-level procedure used in the present experiment and the fixed-level procedure used by Florentine *et al.* (1996).]

The only previous data for impaired listeners are difficult to compare directly with the present data because comparisons between tones of widely different durations were only made indirectly. Pedersen and Poulsen (1973) varied only the shorter tone in a pair and the durations of the tones always differed by a factor of 2. The total amount of temporal integration between 5- and 320-ms tones was estimated by adding the level differences measured for all the pairs between these durations. Whereas this procedure makes the listeners' judgments relatively easy, the estimates of the total amount of temporal integration are susceptible to accumulation of systematic biases. At high levels, the variable tone typically is set to a level slightly below that required for equal loudness (e.g., Stevens, 1955; Scharf, 1961; Port, 1963; Florentine *et al.*, 1996), which would cause the amount of temporal integration to be underestimated when only the short tone is varied. Thus, it is not surprising that the amounts of temporal integration reported by Pedersen and Poulsen (1973) are smaller than those obtained in the present study. Nevertheless, the present study is qualitatively consistent with the previous study in that both found a ten-

dency towards larger amounts of temporal integration in impaired listeners than in normal listeners when the groups are compared at equal, high SPLs. However, when normal and impaired listeners are compared at equal SLs, the amount of temporal integration for impaired listeners tends to be in the low end of the normal range, except when the hearing loss is abrupt (see Fig. 7). Indeed, the amount of temporal integration for detection (i.e., at 0 dB SL) generally is reduced when the threshold is elevated by cochlear hearing losses, as has been shown by many previous studies (for review, see Florentine *et al.*, 1988), as well as in the present study.

It also may be instructive to compare the effects of hearing impairment on temporal integration of loudness to those of partial masking, because it often has been suggested that partial masking increases the rate of loudness growth near masked threshold in a manner similar to the abnormally fast growth of loudness that is typical of listeners with cochlear hearing losses (e.g., Steinberg and Gardner, 1937; Richards, 1973; see also Florentine and Buus, 1986; Buus and Florentine, 1989). Two studies have measured temporal integration of loudness under partial masking (Richards, 1977; Florentine *et al.*, 1998). Although neither study aimed to simulate any particular impaired listener, qualitative comparisons may be made with the present study. Both studies of partial masking found that temporal integration at high levels was unaffected by the presence of a masker, which is similar to the finding that the amount of temporal integration for most impaired listeners tested at high SPLs is equal to or slightly larger than that for normal listeners tested at the same SPLs. In addition, Florentine *et al.* (1998) found that the maximal amount of integration decreased and occurred at increasingly higher-tone SPLs as the masker level increased. The SL at which the maximum occurred decreased as the masker level increased. These findings are qualitatively similar to those for impaired listeners. The present study shows that the maximal amount of temporal integration tends to be smaller in impaired listeners than in normal listeners, except in cases of abrupt loss. The SPL at which the maximum occurs increases as the hearing loss increases, but the SL decreases.

It should be noted that the present study did not find an orderly relation between the listeners' thresholds and their maximal amounts of temporal integration, contrary to the orderly decrease of the maximal amount of temporal integration with increasing masker level reported by Florentine *et al.* (1998). This difference between hearing loss and masking may be understood by considering the mechanisms of threshold elevation in the two cases. In Florentine *et al.*'s (1998) study, the threshold elevation was caused by on-frequency masking, which most likely is excitatory (Delgutte, 1990) and increases threshold by increasing the activity in the auditory channel tuned to the signal, thus requiring the signal to produce more activity before it can be heard. In any event, elevating the threshold by masking evokes a uniform mechanism of threshold elevation across listeners, and it is not surprising that an orderly relation is obtained between the amount of threshold elevation and the maximal amount of temporal integration, which reflects the shallowest slope of the loudness function, as explained later.

Cochlear hearing loss appears to involve two indepen-

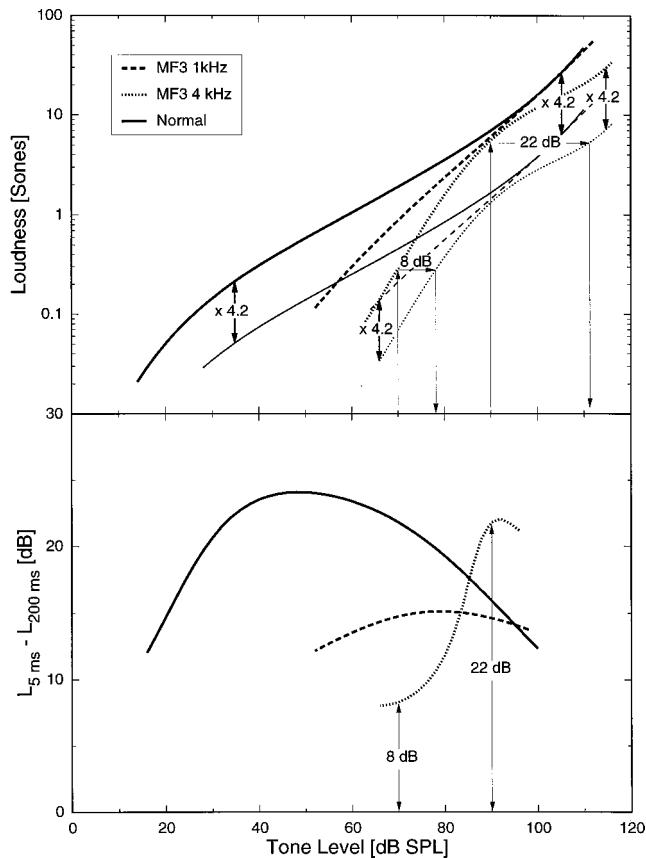


FIG. 8. Relation between the amount of temporal integration for loudness and the loudness function for normal and impaired listeners. The bottom panel shows the temporal-integration functions for normal listeners (averaged across listeners and frequencies; solid line) and for impaired listener MF3 at 1 (dashed line) and 4 kHz (dotted line). The top panel shows loudness functions for 200- (thick lines) and 5-ms tones (thin lines) derived from these temporal-integration functions under the assumption that the ratio of their loudnesses is 4.2 at all SPLs both for normal listeners and for the impaired listener. The arrows to the dotted lines illustrate the application of the equal-loudness-ratio hypothesis to MF3's data. When the 200-ms tone is at 70 dB SPL, the amount of temporal integration is just over 8 dB and its loudness is just below 0.3 sones. The loudness of the 5-ms tone is 4.2 times less, or about 0.07 sones. To match the temporal-integration data, the loudness of the 5-ms tone must grow from 0.07 sones at 70 dB SPL to 0.3 sones at just over 78 dB SPL. This indicates that the loudness function must be quite steep; the exponent is about  $10 \text{ dB} \times \log(4.2)/8 \text{ dB} = 0.79$ . When the 200-ms tone is at 90 dB SPL, the amount of temporal integration is near 22 dB and the loudness is about 5.5 sones. The loudness of the 5-ms tone is again 4.2 times less, or about 1.3 sones. To match the temporal-integration data, the loudness of the 5-ms tone must grow from 1.3 sones at 90 dB to 5.5 sones at 112 dB SPL. This indicates that the loudness function must be rather shallow; the exponent is about  $10 \text{ dB} \times \log(4.2)/22 \text{ dB} = 0.28$ .

dent mechanisms of threshold elevation. One is a loss of transduction efficiency by the inner hair cells, which produces a loss that may be considered similar to an attenuation of the neural output (e.g., Florentine and Zwicker, 1979) or an attenuation of the acoustic signal (Patuzzi, 1993; Moore and Glasberg, 1997). In either case, the hearing loss does not change the form of the loudness function, but shifts it down along the loudness axis or to the right along the signal-level axis. The other mechanism is a loss of mechanical amplification of the basilar-membrane vibration by the outer-hair cells, which causes a loss of compression and alters the form of the loudness function (Yates, 1990b; Ruggero and Rich, 1991; Patuzzi, 1993; Moore and Glasberg, 1997). Because a

given amount of hearing loss can result from different amounts of inner- and outer-hair-cell damage, the amount of alteration to the loudness function caused by a given amount of hearing loss can vary among listeners. Accordingly, it is not surprising that no orderly relation between the amount of hearing loss and the maximal amount of temporal integration is obtained in the present study.

In conclusion, it appears that the effects of partial masking on temporal integration of loudness are very similar to those of a typical cochlear impairment, but are much more orderly. The difference in orderliness probably occurs because masking involves a uniform mechanism of threshold elevation, whereas hearing loss involves two mechanisms that allow the effects of a given hearing loss to vary among listeners.

## B. Relation between temporal integration and growth of loudness in impaired listeners

Previous papers suggested that a close relationship might exist between the temporal-integration function and the growth of loudness (Florentine *et al.*, 1996, 1998; Buus *et al.*, 1997). The relationship can be quantified by assuming that the ratio between the loudnesses of a long and a short tone of equal SPL is independent of the SPL. Although it has yet to be investigated whether this simple relation is strictly true across the entire dynamic range, the assumption of a constant-loudness ratio yields a relation between temporal integration and growth of loudness that appears very useful. If one assumes that the loudness of a 200-ms tone is, for example, four times the loudness of an equal-SPL 5-ms tone, the loudness functions for both tones can be derived from the temporal-integration function, except for a multiplicative constant. In other words, the equal-loudness-ratio hypothesis does not provide an estimate of the absolute loudness. Any pair of loudness functions for 5- and 200-ms stimuli plotted on a logarithmic scale of loudness may be moved vertically without losing consistency with the temporal-integration functions. This vertical movement is similar to the normalization generally applied to direct measurements of the loudness function by the method of magnitude estimation.

As shown in Fig. 8, the equal-loudness-ratio assumption means that the loudness functions for the 5- and 200-ms tones are parallel when they are plotted in the conventional manner with loudness on a logarithmic scale as a function of the SPL of the tone. That is, the vertical distance between the loudness functions is constant, owing to the assumption of a level-independent loudness ratio between them. The amount of temporal integration is the horizontal distance between the functions at some fixed loudness. Accordingly, the loudness functions must be steep when the amount of temporal integration is small, and shallow when it is large.

As illustrated by the arrows in the figure and explained in the caption, the equal-loudness-ratio hypothesis predicts that the slope of the loudness function is roughly inversely proportional to the amount of temporal integration. The proportionality constant depends on the vertical distance between the functions; that is, the value of the fixed-loudness ratio. The ratio is unknown because it depends on the durations of the short and long tones and on how loudness grows

with duration. However, the fixed ratio can be estimated by requiring the loudness function obtained from the temporal-integration function of normal listeners to approximate the "standard" loudness function. For the loudness functions in Fig. 8 the ratio is 4.2, which yields a normal loudness function close to Zwislocki's (1965) modified power function, except at high levels. As noted above, the vertical position of the loudness functions cannot be determined from the temporal-integration functions. The normal loudness functions in Fig. 8 are positioned to yield a monaural loudness of about 0.5 sones for a 200-ms tone at 36 dB SL (corresponding to a binaural loudness of 1 sone at 40 dB SPL in the free field).

The solid lines in Fig. 8 shows the average normal temporal-integration function (lower panel) and the loudness function (upper panel) derived from it when the loudness ratio between 200- and 5-ms tones at equal SPLs is 4.2 at all SPLs. To maintain the constant vertical distance between the 5- and 200-ms loudness functions, they are steep when the amount of temporal integration is small, and shallow when it is large. The primary difference between the normal loudness function shown in Fig. 8 and the standard modified power function (Zwislocki, 1965) is that the former shows a steepening at high levels, because the amount of temporal integration decreases at high levels. It is noteworthy that the normal loudness function in Fig. 8 is at least qualitatively similar to the compressive transfer function used in the models of Moore, Oxenham, and colleagues (Moore *et al.*, 1996; Oxenham and Moore, 1997; Oxenham *et al.*, 1997; Moore and Oxenham, 1998). The loudness function derived from our temporal-integration measurements also shows features similar to the mechanical input/output function measured at the basilar membrane. The amplitude of basilar-membrane vibration grows more slowly with acoustic amplitude at moderate levels than at low and high levels (e.g., Robles *et al.*, 1986; Yates, 1990a).

To explore the implications of the equal-loudness-ratio hypothesis for the present data, one may assume that the loudness ratio between equal-SPL 200- and 5-ms tones is the same in normal and in typical impaired listeners, as illustrated by the vertical arrows. If the equal-loudness-ratio hypothesis holds and if a typical hearing loss does not alter the loudness ratio between short and long tones, the similarities and differences between the normal and impaired temporal-integration functions should reflect similarities and differences between normal and impaired loudness-growth functions. This is illustrated by the broken lines in Fig. 8, which show temporal-integration functions (lower panel) and the corresponding loudness functions (upper panel) for listener MF3 at 1 (dashed lines) and 4 kHz (dotted lines). Again, the vertical position of the loudness functions cannot be determined from the temporal-integration functions. The loudness functions for MF3 are positioned to yield close-to-normal loudness at high levels, consistent with complete recruitment. This position also is consistent with clinical estimates of her loudness function at 4 kHz. The horizontal distance between the normal loudness functions and MF3's loudness functions at 4 kHz (that is, the gain required to produce normal loudness at 4 kHz) is within 2 dB of that estimated

from category-scaling measurements of loudness functions for normal listeners and for MF3 in another laboratory.

As shown by the dotted line in the bottom panel, MF3's temporal-integration function at 4 kHz showed a relatively small amount of temporal integration, just over 8 dB, at 70 dB SPL (about 10 dB SL for the 200-ms tone). Because the vertical distance between the 5- and 200-ms loudness functions is fixed, her loudness functions must be quite steep to yield a horizontal distance of only 8 dB. This is consistent with the recruitment that usually results from cochlear impairment. For high-SPL tones at 4 kHz, MF3 shows more temporal integration than normal listeners, which indicates that her loudness function may be shallower. Overall, MF3's data at 4 kHz appear consistent with a loss of compression at low SLs whereas considerable compression may be present at levels above 90 dB SPL. In terms of excitation-pattern models of loudness in impaired listeners (Florentine and Zwicker, 1979; Moore and Glasberg, 1997), calculations indicate that all of her hearing loss is due to outer-hair-cell damage (Moore, personal communication).

The temporal-integration data for MF3 at 1 kHz are quite different from those at 4 kHz. Near threshold and at high levels her amount of temporal integration is about 12 or 13 dB, which is comparable to that obtained for normal listeners near threshold and around 95 dB SPL. However, MF3's data show only a very small effect of level on the amount of temporal integration, which causes the predicted loudness functions to be nearly perfect power functions. The loudness functions predicted for MF3 at 1 kHz have about the same slope as the normal listeners' loudness functions near threshold, but are steeper than those for normal listeners at the same SPLs. This indicates that some recruitment is present. The slope at high levels is similar to normal when evaluated at equal SPLs, consistent with complete recruitment. Again, the loudness functions derived from MF3's temporal-integration data appear to indicate that her hearing loss is due to outer-hair-cell damage.

According to the excitation-pattern models of loudness in impaired listeners (Florentine and Zwicker, 1979; Moore and Glasberg, 1997), the form of impaired listeners' loudness functions can vary widely depending on the extent to which the hearing loss is due to damage of inner- or outer-hair cells. According to the equal-loudness-ratio hypothesis, the form of the temporal-integration function changes with the form of the loudness function. Therefore, the wide variation of the temporal-integration functions obtained for the impaired listeners is likely to reflect varying proportions of inner- and outer-hair-cell damage. The dominant trends of the present data certainly seem to agree with expectations derived from the general properties of loudness growth in impaired listeners. The normal or slightly larger-than-normal temporal integration obtained for impaired listeners at high SPLs indicates that impaired listeners' loudness functions should have normal or slightly flatter-than-normal slopes at high levels. This seems consistent with measurements and theories of loudness in impaired listeners (e.g., Miskolczy-Fodor, 1960; Florentine and Zwicker, 1979; Florentine *et al.*, 1997; Hellman and Meiselman, 1991; Moore and Glasberg, 1997), which typically show nearly normal growth of loud-

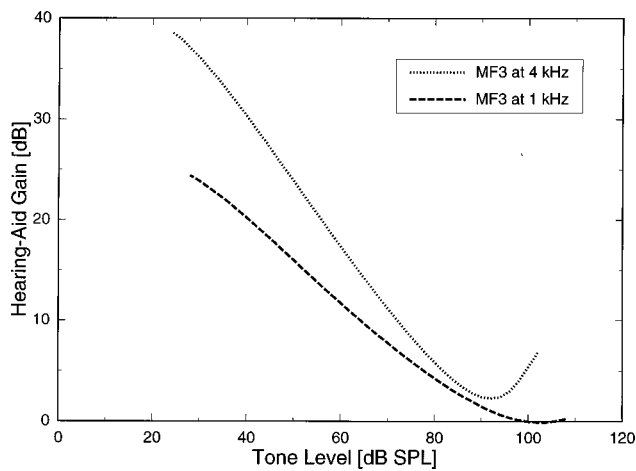


FIG. 9. Predicted hearing-aid gain necessary to restore normal loudness for listener MF3 is plotted as a function of input level for tones at 1 (dashed line) and 4 kHz (dotted line). The predictions are identical for 5- and 200-ms tones.

ness in cochlear-impaired listeners tested at high levels, well above their elevated thresholds. Indeed, the excitation-pattern model of loudness for impaired listeners can predict even shallower-than-normal loudness functions for impaired listeners at high levels (Moore, personal communication). Likewise, the tendency of most impaired listeners to show less-than-normal amounts of temporal integration at and near threshold is consistent with these listeners' steeper-than-normal growth of loudness for tones near threshold (e.g., Miskolczy-Fodor, 1960; Hellman and Meiselman, 1991; Moore and Glasberg, 1997).

The finding that maximal amounts of temporal integration tend to be smaller than normal in most impaired listeners indicates that the most compressive part of the loudness function is steeper in impaired than in normal listeners. This notion is consistent with the idea that cochlear impairments lead to a reduction or even complete loss of compression (Moore and Glasberg, 1997). Listeners with abrupt losses are an exception because their loudness functions are flatter than normal for frequencies close to the highest frequency of near-normal hearing (Hellman, 1994; Florentine *et al.*, 1997; McDermott *et al.*, 1998). Accordingly, one would expect listeners with abrupt losses to show larger-than-normal amounts of temporal integration at frequencies close to the onset of an abrupt loss. The present data for two listeners with abrupt losses fulfill this expectation.

### C. Implications for signal processing in hearing aids

The loudness functions shown in Fig. 8 also have interesting implications for the design of hearing aids intended to restore impaired listeners' perception of loudness to normal. The gain required for loudness restoration may be estimated as the horizontal distance between the loudness function for normal listeners and that for the impaired listener. Figure 9 shows the hearing-aid gain estimated in this manner for MF3 at 1 and 4 kHz. The gain for 1-kHz stimuli decreases steadily with increasing sound level to reach 0 dB at high levels. This estimate is consistent with the notion that wide-dynamic-range compression is best suited to restore impaired listen-

ers' loudness growth (Barfod, 1978; Killion, 1993). The gain function for 4-kHz stimuli is within 2 dB of that derived from clinical category-scaling measurements obtained for MF3, as noted in the discussion of the loudness functions in Fig. 8. This gain function also decreases for input levels up to just over 90 dB SPL, but increases somewhat at higher levels due to the apparent compression in MF3's 4-kHz loudness function at high levels. Despite the departure from constant gain at high levels, the 4-kHz gain function is reasonably close to that provided by a hearing aid with wide-dynamic-range compression. The findings of gain functions consistent with those derived from other measurements of loudness in impaired listeners indicate that the temporal-integration data for impaired listeners are likely to reflect the properties of their loudness functions and a level-independent loudness ratio between long and short tones.

Clinical loudness measurements used to estimate suitable gain characteristics for hearing aids usually employ tones of relatively long durations, such that the gain functions may be most appropriate for long-duration stimuli. Although many natural sounds are characterized by brief periods of relatively high intensity surrounded by periods of lower intensity, little consideration has been given to producing the appropriate gain for brief stimuli. The present data for temporal integration together with the equal-loudness-ratio hypothesis provide estimates of loudness functions for both long and short stimuli, and allow gain functions for both to be derived. If the loudness ratio between long and short stimuli is the same in normal and impaired listeners, the normal and impaired loudness functions for short stimuli are displaced vertically from those for long stimuli by the same amount. Accordingly, the horizontal distances are maintained, which results in the prediction that the gain required to restore loudness of a sound of a given SPL is the same for long and short stimuli.

The requirement of identical gain for short and long stimuli with the same SPL can be approximated if a fast attack time is used for the compression hearing aid. To produce normal temporal variation in the loudness, a fast release time may also be desirable. However, compression schemes in which both the attack and the release are fast reduce the modulation depth of time-varying stimuli. Such a reduction in modulation depth may be undesirable because impaired listeners' sensitivity to amplitude modulation generally is similar to or worse than normal (e.g., Bacon and Viemeister, 1985; Hall and Grose, 1989; Bacon and Gleitman, 1992; Moore *et al.*, 1992). Moreover, the overall loudness of a time-varying sound often is close to the loudness of its peaks (Fastl, 1975; Zwicker, 1977, 1984; Zwicker and Fastl, 1990; Zhang and Zeng, 1997). [The overall loudness may be closer to the average loudness when the modulation depth is moderate (Moore *et al.*, 1998), but in this case the difference between the average loudness and the peak loudness is relatively small.] Thus, compression with a fast attack time may be sufficient to produce near-normal loudness in impaired listeners, and it is unclear whether a fast release time is necessary or even desirable.

The prediction of identical gains for long and short stimuli resulted from the assumption that hearing impairment

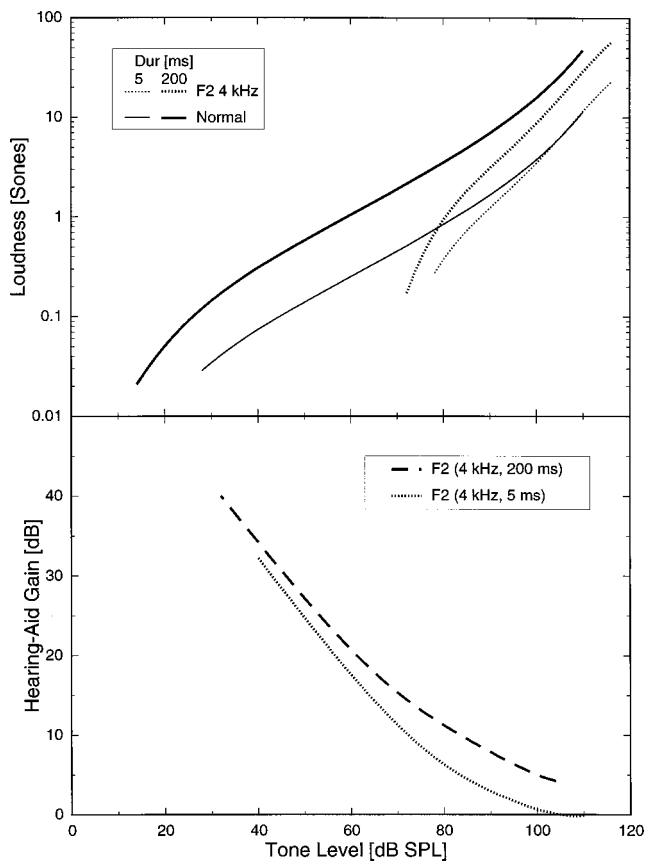


FIG. 10. Loudness functions and predicted hearing-aid gain necessary to restore normal loudness for listener F2, who showed very small amounts of temporal integration. The upper panel shows loudness functions for 5- (thin lines) and 200-ms tones (thick lines) for normal listeners (solid lines) and for listener F2 at 4 kHz (dotted lines). The lower panel shows the hearing-aid gain predicted as the horizontal distance between the normal and impaired loudness functions for 5- (dotted line) and 200-ms tones (dashed line). Because F2's abnormally small amounts of temporal integration were assumed to reflect abnormal adaptation of auditory activity, the equal-loudness-ratio hypothesis predicts that a higher gain is needed to restore loudness for 200- than for 5-ms tones.

did not alter the loudness ratio between long and short tones. Whereas this assumption seems reasonable for most listeners with cochlear impairments, it has yet to be verified. In addition, it may be incorrect for some listeners such as F2, who shows markedly lower-than-normal amounts of temporal integration. Unusually small amounts of temporal integration may reflect an abnormally rapid adaptation of neural activity in the auditory system, which is likely to reduce the rate at which loudness increases with duration. This would result in a smaller-than-normal loudness ratio between long and short tones. The amount of temporal integration for F2 at 4 kHz was only about 4 dB near her threshold of 68 dB SPL, increased to a maximum of about 9 dB around 90 dB SPL, and decreased to about 6 dB at high SPLs. These small amounts of temporal integration would imply very steep loudness functions, if the loudness ratio between long and short tones were normal. If the loudness ratio between 200- and 5-ms tones is assumed to be reduced to about 2.3 by abnormal adaptation, however, the loudness functions predicted for F2 have the same slope as the normal loudness functions at high SPLs as shown in Fig. 10.

Because the abnormal adaptation ought to reduce the loudness of long-duration stimuli, the vertical position of the predicted loudness functions in the top panel of Fig. 10 was chosen to show normal loudness for 5-ms tones at high levels, but reduced loudness (incomplete recruitment) for long tones. The gain functions predicted for loudness restoration are shown in the lower panel. As for the listener with a normal loudness ratio between long and short tones (and, presumably, normal adaptation), the gain functions require wide-dynamic-range compression. However, due to F2's altered loudness ratio between long and short tones, the gain required for 5-ms tones is no longer the same as that for 200-ms tones. The gain for a 5-ms tone (dotted line) decreases to 0 dB at high levels as follows from the assumption of complete recruitment; the gain for a 200-ms tone (dashed line) also decreases with increasing level, but is always *higher* than that for the 5-ms tone. Whereas the gain for any given input SPL and the detailed form of the gain function depend critically on the vertical position chosen for F2's loudness functions, the prediction of higher gain for long than for short stimuli does not. It results from the assumption of a reduced loudness ratio.

The goal of having higher gain for long than for short sounds may be met by properly designed digital signal processing. However, such processing is likely to introduce an undesirable delay and/or distortion to the signal envelope, which could reduce speech intelligibility and outweigh any advantage gained by restoring loudness. Therefore, it may be advantageous to use the same gain for brief and long signals, even in cases of abnormal adaptation and unusually small amounts of temporal integration. Certainly, a duration-independent gain function ought to provide a reasonable approximation to restoration of normal loudness because the gains required for brief and long signals of equal SPL usually differ by only a few dB. However, the optimal gain for listeners with abnormal adaptation may be slightly lower than the gain derived from loudness measurements with long signals.

#### IV. SUMMARY

The present study obtained measurements of loudness balance between 5- and 200-ms tones at various frequencies to assess how temporal integration for loudness varies with level in 15 listeners with hearing impairments of predominantly cochlear origin and in a control group of seven age-matched listeners with normal hearing. The main conclusions are:

- (1) The amount of temporal integration for loudness, defined as the level difference between equally loud 5- and 200-ms tones, varies nonmonotonically with level and shows a maximum at moderate SPLs in all normal listeners and in most impaired listeners.
- (2) The effect of level on the amount of temporal integration varies considerably among normal listeners and varies even more among impaired listeners.
- (3) The amount of temporal integration for loudness near

threshold generally approaches that for detection and tends to be smaller in impaired listeners than in normal listeners.

- (4) The amount of temporal integration for loudness at high SPLs usually is about the same or even slightly larger in impaired listeners than in normal listeners tested at the same SPLs.
- (5) When evaluated at equal SLs, the impaired listeners tend to show normal or less-than-normal amounts of temporal integration, except when the hearing loss is abrupt—i.e., increases more than 50 dB per octave above the test frequency.
- (6) The amounts of temporal integration for loudness in listeners with abrupt losses are much larger than normal for frequencies at or somewhat below the highest frequency of near-normal thresholds.
- (7) The large amounts of temporal integration obtained for listeners with abrupt losses are consistent with these listeners' flatter-than-normal loudness functions and the hypotheses that the loudness ratio between equal-SPL long and short tones is independent of SPL and is the same in normal and most impaired listeners.
- (8) Overall, the results for impaired listeners appear consistent with expectations based on knowledge of the general properties of their loudness-growth functions and the equal-loudness-ratio hypothesis. The large variation of the impaired listeners' temporal-integration functions may reflect large variation in their loudness-growth functions resulting from differing balances between inner- and outer-hair-cell dysfunction.
- (9) Application of the equal-loudness-ratio hypothesis to the present data indicates that the hearing-aid gain necessary to restore normal loudness is independent of stimulus duration for most impaired listeners. However, some impaired listeners have abnormally small amounts of temporal integration, which may result from abnormally fast adaptation of auditory activity. For these listeners, the gain required for restoring normal loudness may be higher for long than for brief stimuli.

## ACKNOWLEDGMENTS

The authors are grateful for the efficient and helpful collaboration with Nikolai Bisgaard, Ole Dyrland, and Jesper Rye Nielsen at GN Danavox, who provided the impaired listeners for this study and who provided loudness-scaling data for listener MF3. The reviewers, Walt Jesteadt and Brian C. J. Moore, provided very helpful comments to an earlier version of this paper. We are especially grateful to Brian C. J. Moore for providing results of model calculations for MF3. Part of this research was carried out at the Technical University of Denmark. It was supported by NIH-NIDCD Grant No. RO1 DC02241.

- Bacon, S. P., and Gleitman, R. M. (1992). "Modulation detection in subjects with relatively flat hearing losses," *J. Speech Hear. Res.* **35**, 642–653.
- Bacon, S. P., and Viemeister, N. F. (1985). "Temporal modulation transfer functions in normal-hearing and hearing-impaired subjects," *Audiology* **24**, 117–134.

- Barfod, J. (1978). "Multichannel compression hearing aids: Experiments and considerations on clinical applicability," in *Sensorineural Hearing Impairment and Hearing Aids*, edited by C. Ludvigsen and J. Barfod, *Scand. Audiol. Suppl.* **6**, 315–339.
- Buus, S., and Florentine, M. (1989). "Simulated hearing loss as a baseline for the assessment of auditory function in cochlearly impaired listeners," in *Proceedings of the 15th Annual Bioeng. Conference*, edited by S. Buus (IEEE, Piscataway, NJ), pp. 19–20.
- Buus, S., Florentine, M., and Poulsen, T. (1997). "Temporal integration of loudness, loudness discrimination, and the form of the loudness function," *J. Acoust. Soc. Am.* **101**, 669–680.
- Buus, S., Florentine, M., and Poulsen, T. (1997). "Temporal integration of loudness in listeners with normal hearing and sensorineural hearing loss," in *Proceedings of the 2nd Biennial Hearing Aid Research and Development Conference* (NIDCD/VA, Bethesda, MD), p. 42.
- Buus, S., Florentine, M., and Poulsen, T. (1997). "Loudness of brief tones in listeners with normal hearing and sensorineural hearing loss," *ASHA Leader* **2**(15), 143.
- Buus, S., Müsch, H., and Florentine, M. (1998). "On loudness at threshold," *J. Acoust. Soc. Am.* **104**, 399–410.
- Carlyon, R. P., Buus, S., and Florentine, M. (1990). "Temporal integration of trains of tone pulses by normal and by cochlearly impaired listeners," *J. Acoust. Soc. Am.* **87**, 260–268.
- Delgutte, B. (1990). "Physiological mechanisms of psychophysical masking: Observations from auditory-nerve fibers," *J. Acoust. Soc. Am.* **87**, 791–809.
- Fastl, H. (1975). "Loudness and masking patterns of narrow noise bands," *Acustica* **33**, 266–271.
- Florentine, M., and Buus, S. (1986). "Simulated hearing loss: a tool for comparing normal and impaired hearing," *Proceedings of the 12th International Congress of Acoust.*, Vol. 1, B11–3.
- Florentine, M., Buus, S., and Hellman, R. P. (1997). "A model of loudness summation applied to high-frequency hearing loss," in *Modeling Sensorineural Hearing Loss*, edited by W. Jesteadt (Erlbaum, Mahwah, NJ), pp. 187–198.
- Florentine, M., Buus, S., and Poulsen, T. (1996). "Temporal integration of loudness as a function of level," *J. Acoust. Soc. Am.* **99**, 1633–1644.
- Florentine, M., Buus, S., and Robinson, M. (1998). "Temporal integration of loudness under partial masking," *J. Acoust. Soc. Am.* **104**, 999–1007.
- Florentine, M., Fastl, H., and Buus, S. (1988). "Temporal integration in normal hearing, cochlear impairment, and impairment simulated by masking," *J. Acoust. Soc. Am.* **84**, 195–203.
- Florentine, M., Reed, C. M., Rabinowitz, W. M., Braida, L. D., Durlach, N. I., and Buus, S. (1993). "Intensity perception. XIV. Intensity discrimination in listeners with sensorineural hearing loss," *J. Acoust. Soc. Am.* **94**, 2575–2586.
- Florentine, M., and Zwicker, E. (1979). "A model of loudness summation applied to noise-induced hearing loss," *Hearing Res.* **1**, 121–132.
- Hall, J. W., and Grose, J. H. (1989). "Spectro-temporal analysis and cochlear hearing impairment: Effects of frequency selectivity, temporal resolution, signal frequency and rate of modulation," *J. Acoust. Soc. Am.* **85**, 2550–2562.
- Hellman, R. P. (1974). "Effect of spread of excitation on the loudness function at 250 Hz," in *Sensation and Measurement*, edited by H. R. Moskowitz, B. Scharf, and J. C. Stevens (Reidel, Dordrecht, Holland), pp. 241–249.
- Hellman, R. P. (1994). "Relation between the growth of loudness and high-frequency excitation," *J. Acoust. Soc. Am.* **96**, 2655–2663.
- Hellman, R. P., and Meiselman, C. H. (1991). "Loudness relations for individuals and groups in normal and impaired hearing," *J. Acoust. Soc. Am.* **88**, 2596–2606.
- Hellman, R. P., and Meiselman, C. H. (1993). "Rate of loudness growth for pure tones in normal and impaired hearing," *J. Acoust. Soc. Am.* **93**, 966–975.
- Jesteadt, W. (1980). "An adaptive procedure for subjective judgments," *Percept. Psychophys.* **28**, 85–88.
- Killion, M. C. (1993). "The *K*-amp hearing aid: An attempt to present high fidelity for persons with impaired hearing," *Am. J. Audiol.* **2**, 52–74.
- Levitt, H. (1971). "Transformed up-down methods in psychoacoustics," *J. Acoust. Soc. Am.* **49**, 467–477.
- McDermott, H. J., Lech, M., Kornblum, M. S., and Irvine, D. R. F. (1998). "Loudness perception and frequency discrimination in subjects with steeply sloping hearing loss: Possible correlates of neural plasticity," *J. Acoust. Soc. Am.* **104**, 2314–2325.

- Miskolczy-Fodor, F. (1960). "Relation between loudness and duration of tonal pulses. III. Response in cases of abnormal loudness function," *J. Acoust. Soc. Am.* **32**, 486–492.
- Moore, B. C. J. (1995). *Perceptual Consequences of Cochlear Damage* (Oxford University Press, Oxford).
- Moore, B. C. J. (1998). Personal communication.
- Moore, B. C. J., and Glasberg, B. R. (1986). "The role of frequency selectivity in the perception of loudness, pitch and time," in *Frequency Selectivity in Hearing*, edited by B. C. J. Moore (Academic, London), pp. 251–308.
- Moore, B. C. J., and Glasberg, B. R. (1997). "A model of loudness perception applied to cochlear hearing loss," *Aud. Neurosci.* **3**, 289–311.
- Moore, B. C. J., Launer, S., Vickers, D. A., and Baer, T. (1998). "Loudness of modulated sounds as a function of modulation rate, modulation depth, modulation waveform, and overall level," in *Psychophysical and Physiological Advances in Hearing*, edited by A. R. Palmer, A. Rees, A. Q. Summerfield, and R. Meddis (Whurr, London), pp. 465–472.
- Moore, B. C. J., and Oxenham, A. J. (1998). "Psychoacoustic consequences of compression in the peripheral auditory system," *Psychol. Rev.* **105**, 108–124.
- Moore, B. C. J., Peters, R. W., and Glasberg, B. R. (1996). "Detection of decrements and increments in sinusoids at high overall levels," *J. Acoust. Soc. Am.* **99**, 3669–3677.
- Moore, B. C. J., Shailer, M. J., and Schooneveldt, G. P. (1992). "Temporal modulation transfer functions for band-limited noise in subjects with cochlear hearing loss," *Br. J. Audiol.* **26**, 229–237.
- Oxenham, A. J., and Moore, B. C. J. (1997). "Modeling the effects of peripheral nonlinearity in normal and impaired hearing," in *Modeling Sensorineural Hearing Loss*, edited by W. Jesteadt (Erlbaum, Mahwah, NJ), pp. 273–288.
- Oxenham, A. J., Moore, B. C. J., and Vickers, D. A. (1997). "Short-term temporal integration: Evidence for the influence of peripheral compression," *J. Acoust. Soc. Am.* **101**, 3676–3687.
- Patuzzi, R. B. (1993). "Otoacoustic emissions and the categorization of cochlear and retrocochlear hearing losses," *Br. J. Audiol.* **27**, 91–95.
- Pedersen, C. B., and Poulsen, T. (1973). "Loudness of brief tones in hearing-impaired ears," *Acta Oto-Laryngol.* **76**, 402–409.
- Port, E. (1963). "Über die Lautstärke einzelner kurzer Schallimpulse," *Acustica* **13**, 212–223.
- Poulsen, T. (1981). "Loudness of tone pulses in a free field," *J. Acoust. Soc. Am.* **69**, 1786–1790.
- Richards, A. M. (1973). "Loudness growth under masking: relation to true sensorineural impairment," *J. Speech Hear. Res.* **16**, 597–607.
- Richards, A. M. (1977). "Loudness perception for short-duration tones in masking noise," *J. Speech Hear. Res.* **20**, 684–693.
- Robles, L., Ruggero, M. A., and Rich, N. C. (1986). "Basilar membrane mechanics at the base of the chinchilla cochlea I. Input–output functions, tuning curves, and phase responses," *J. Acoust. Soc. Am.* **80**, 1364–1374.
- Ruggero, M. A., and Rich, N. C. (1991). "Furosemide alters organ of Corti mechanics: Evidence for feedback of outer hair cells upon the basilar membrane," *J. Neurosci.* **11**, 1057–1067.
- Scharf, B. (1961). "Loudness summation under partial masking," *J. Acoust. Soc. Am.* **64**, 1036–1040.
- Scharf, B. (1970). "Critical bands," in *Foundations of Modern Auditory Theory, Vol. I*, edited by J. V. Tobias (Academic, New York), pp. 157–202.
- Scharf, B. (1978). "Loudness," in *Handbook of Perception: Vol. 4. Hearing*, edited by E. C. Carterette and M. P. Friedman (Academic, New York), pp. 187–242.
- Steinberg, J. C., and Gardner, M. B. (1937). "The dependence of hearing impairment on sound intensity," *J. Acoust. Soc. Am.* **9**, 11–23.
- Stephens, S. D. G. (1973). "Auditory temporal integration as a function of intensity," *J. Sound Vib.* **37**, 235–246.
- Stevens, S. S. (1955). "The measurement of loudness," *J. Acoust. Soc. Am.* **27**, 815–827.
- Yates, G. K. (1990a). "The basilar membrane input-output function," in *The Mechanics and Biophysics of Hearing*, edited by P. Dallos, C. D. Geisler, J. M. Matthews, M. A. Ruggero, and C. R. Steele (Springer, New York), pp. 106–113.
- Yates, G. K. (1990b). "Basilar-membrane nonlinearity and its influence on auditory-nerve rate-intensity functions," *Hearing Res.* **50**, 145–162.
- Zhang, C., and Zeng, F. (1997). "Loudness of dynamic stimuli in acoustic and electric hearing," *J. Acoust. Soc. Am.* **102**, 2925–2934.
- Zwicker, E. (1961). "Subdivision of the audible frequency range into critical bands (Frequenzgruppen)," *J. Acoust. Soc. Am.* **33**, 248.
- Zwicker, E. (1977). "Procedure for calculating loudness of temporally variable sounds," *J. Acoust. Soc. Am.* **62**, 675–682.
- Zwicker, E. (1984). "Dependence of post-masking on masker duration and its relation to temporal effects in loudness," *J. Acoust. Soc. Am.* **75**, 219–223.
- Zwicker, E., and Fastl, H. (1990). *Psychoacoustics—Facts and Models* (Springer, Berlin).
- Zwislocki, J. J. (1965). "Analysis of some auditory characteristics," in *Handbook of Mathematical Psychology*, edited by R. D. Luce, R. R. Bush, and E. Galanter (Wiley, New York), pp. 1–97.

# System identification of feedback in hearing aids

Johan Hellgren,<sup>a)</sup> Thomas Lunner, and Stig Arlinger

Department of Neuroscience and Locomotion, Division of Technical Audiology, Linköping University, 581 85 Linköping, Sweden

(Received 28 May 1998; revised 18 November 1998; accepted 4 March 1999)

The feedback problems of behind the ear (BTE), in the ear (ITE), and in the ear canal (ITEC) hearing aid categories have been investigated. All possible feedback paths (acoustical via vent, via tubing wall, mechanical, etc.) were converted to a single transfer function from the ear canal to the hearing aid microphone, here called the acoustic feedback equivalent (AFE). The attenuation of the AFE represents the maximum gain that can be used without the hearing aid starting to howl. Magnitude and phase responses of the AFE were identified on ten human subjects and on a Knowles ear manikin (KEMAR). The acoustic feedback via vent and leak between earmould and ear canal dominated the AFE. The transfer function from a reference point under the ear to the position of microphone of the different hearing aid categories was identified and used together with the AFE to calculate the maximum real ear aided gain (REAG) for the hearing aid categories. A model of the AFE, consisting of a fourth-order filter together with a delay, showed good agreement with the measured data. © 1999 Acoustical Society of America. [S0001-4966(99)02506-0]

PACS numbers: 43.66.Ts, 43.64.Ha [JWH]

## INTRODUCTION

Feedback is often used in control systems and amplifiers to improve the performance. In hearing aids, the feedback from the receiver to the microphone is an undesired effect that is hard to avoid due to the proximity of the receiver and microphone. The feedback often causes the hearing aid to howl and limits the gain that can be used without instability. It is a clinical experience that feedback is one of the most common complaints about hearing aids. The feedback will cause a reduction of the sound quality when the gain is set close to the limit. Users often describe the effect on the sound quality as echo, ringing, or being metallic.

Feedback often forces the dispenser to make compromises in the fitting of a hearing aid. The choice of vent size and gain is then a tradeoff that includes the gain that can be achieved, degree of occlusion, and how much feedback is acceptable. For subjects with little or no hearing loss for low frequencies, a large vent is desired to avoid occlusion effects. If the same subject has a pronounced hearing loss in the high frequencies, a reduced vent size is often needed to avoid feedback. The gain desired by subjects with severe hearing losses is often so high that it can be hard to achieve even without vent because of leaking earmould. Oscillation due to feedback often occurs at high frequencies. To avoid oscillations the high-frequency gain is often reduced (Dyrlund and Lundh, 1990).

### A. How the feedback affects the gain of the hearing aid

Figure 1 shows a schematic picture of the different transfer functions that affect the signal on the way from the source to the eardrum.  $G_1(f)$  is the (acoustic) transfer function from a sound source to the hearing aid microphone. It

should be noted that  $G_1(f)$  depends on the hearing aid category since the position of the microphone, and thus the transfer function to it, differs between hearing aid categories (Fikret-Pasa and Revit, 1992).  $G_2(f)$  is the electroacoustical transfer function from the inlet of the hearing aid microphone to the tip of the earmould inside the ear canal.  $G_2(f)$  is the open loop gain, i.e., the gain through the hearing aid when there is no feedback.  $G_3(f)$  is the feedback path from the ear canal back to the hearing aid microphone inlet. In reality the feedback is caused by a number of different feedback paths. This is further discussed below.  $G_4(f)$  is the transfer function from the tip of the earmould to the point of interest in the ear canal (e.g., tympanic membrane). The closed loop gain is the gain with feedback present, given by Eq. (1).  $G_2(f)G_3(f)$  is the loop response. The system will start to oscillate when the gain of the hearing aid is increased to make at some frequency, the amplitude of the loop response 0 dB and the phase a multiple of  $360^\circ$ .

$$G(f) = \frac{G_2(f)}{1 - G_2(f)G_3(f)}. \quad (1)$$

Real ear aided gain (REAG) (=transfer function from a reference point to the point of interest in the ear canal) of a hearing aid with feedback present is given by Eq. (2):

$$\text{REAG}(f) = \frac{G_1(f)G_2(f)G_4(f)}{1 - G_2(f)G_3(f)}. \quad (2)$$

With the controls (e.g., volume and tone controls) of the hearing aid it is possible to adjust  $G_2(f)$ . At low gain settings of the volume control (large gain margin) the  $\text{REAG}(f)$  will be very close to  $G_1(f)G_2(f)G_4(f)$  since  $|G_2(f)G_3(f)| \ll 1$ . As the gain in the hearing aid is increased, the feedback will increasingly influence REAG.  $\text{REAG}(f)$  will vary markedly depending on the phase of  $G_2(f)G_3(f)$ , if  $|G_2(f)|$  is adjusted so that  $|G_2(f)G_3(f)| = 1$ . For frequencies with  $\angle G_2(f)G_3(f) = 180^\circ$  the magni-

<sup>a)</sup>Electronic mail: johan.hellgren@oto.liu.se



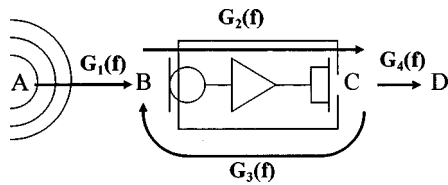


FIG. 1. A schematic figure of the signal paths from a sound source to the tympanic membrane. A is the sound source and  $G_1(f)$  is the transfer function from the sound source (A) to the hearing aid microphone (B).  $G_2(f)$  is the open loop gain of the hearing aid from the microphone to the tip of the earmould (C).  $G_3(f)$  is the feedback from the tip of the earmould back to the hearing aid microphone.  $G_4(f)$  is the transfer function from the earmould tip to the tympanic membrane (D).

tude of REAG will be  $|G_1(f)G_2(f)G_4(f)|/2$  and for frequencies with  $\angle G_2(f)G_3(f) = 0^\circ$  it will approach infinity. The system will then be marginally stable and the gain has to be lowered in order to avoid oscillations. A practical upper limit of the gain of the hearing aid is to set  $|G_2(f)G_3(f)| \leq 0.5$ , i.e., a gain margin of 6 dB. REAG is then limited by Eq. (3):

$$\frac{2}{3}|G_1(f)G_2(f)G_4(f)| \leq |\text{REAG}(f)| \leq 2|G_1(f)G_2(f)G_4(f)|. \quad (3)$$

The variations of REAG caused by the phase of the loop response are then within 9.5 dB [REAG is within  $-3.5$  to  $+6$  dB relative to the open loop gain  $G_1(f)G_2(f)G_4(f)$ ]. The phase of the loop response will vary with frequency. Thus, the shape of the REAG is altered as the gain of the hearing aid is increased. This appears as peaks on the frequency response occurring at the frequencies where the phase is a multiple of  $360^\circ$ . Cox (1982) measured the effect of feedback on the frequency response of the hearing aid. The change of the frequency response showed peaks and notches when the gain of the hearing aid was set close to the limit.  $1/G_3(f)$  can be used as a practical upper limit of the closed loop gain from the hearing aid microphone to the earmould tip. Higher gain can be achieved at some frequencies, but then the deviation from open loop gain will be outside the  $-3.5$ - to  $6$ -dB range. The upper limit of REAG( $f$ ) is then given by

$$|\text{REAG}(f)| \leq \left| \frac{G_1(f)G_4(f)}{G_3(f)} \right|. \quad (4)$$

Another argument supporting Eq. (4) as an upper limit of the gain is to look at the maximum gain of a hearing aid with negative feedback. Assume that it was possible to adjust the phase of the loop to  $180^\circ$  (or  $180^\circ + n360^\circ$ ) for all frequencies that can cause oscillation. Then the hearing aid would not start to oscillate independently of how much the open loop gain was increased. However, the REAG would approach an asymptote as the open loop gain was increased due to the negative feedback. The negative phase of the loop response will give that

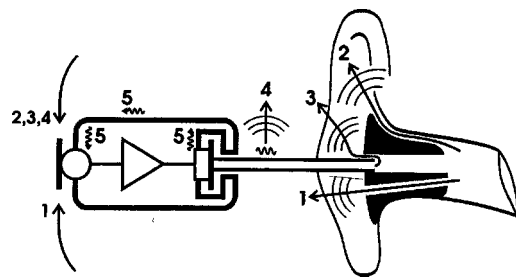


FIG. 2. Feedback paths of a hearing aid.

$$\begin{aligned} |\text{REAG}(f)| &= \left| \frac{G_1(f)G_2(f)G_4(f)}{1 - G_2(f)G_3(f)} \right| \\ &= \frac{|G_1(f)G_2(f)G_4(f)|}{1 + |G_2(f)G_3(f)|} \\ &< \left| \frac{G_1(f)G_4(f)}{G_3(f)} \right|. \end{aligned} \quad (5)$$

This can be considered as a more theoretical argument, since the phase of the loop response is usually not considered when fitting hearing aids.

## B. Feedback paths of a hearing aid

The feedback of a hearing aid consists of several pathways from the receiver to the hearing aid microphone as shown in Fig. 2. The principal feedback paths are (G. Naylor, personal communication)

- (1) the vent,
- (2) imperfect sealing of the earmould,
- (3) imperfect sealing in the joints of the tubing (behind the ear, BTE, hearing aids),
- (4) emission from the tubing walls (BTE hearing aids),
- (5) structural transmission within the hearing aid,
- (6) emission from the shell of the hearing aid (not shown in Fig. 2),
- (7) acoustic transmission in the hearing aid (not shown in Fig. 2), and
- (8) electrical feedback (not shown in Fig. 2).

The feedback path via the vent is an obvious path. The imperfect sealing of the earmould will work like an extra vent in parallel with the vent in the earmould. The mean leakage, according to Angaard Johansen (1975b), corresponds to a vent with a diameter of 1.6 mm and a length of 22 mm. Imperfect gluing of the earmould hook and old hard tubes may cause leakage in the tubing between the receiver and the earmould. Lundh (1982) found imperfect gluing of the earmould hook in 25% of the tested cases. Flack *et al.* (1995) investigated the attenuation of different tubing walls and found that the least attenuation occurred at 2.5–3 kHz. The attenuation at these frequencies was 70–75 dB for different tubes. The last four types of feedback paths are closely connected to the design of the hearing aid.

There are a number of parameters that affect the feedback. The four first and the sixth types of feedback paths in the list above each include an acoustic path outside the hearing aid to the microphone. The position of the microphone

differs between hearing aid categories—BTE, in the ear hearing aid (ITE), or in the ear canal hearing aid (ITEC). Thus, the hearing aid category will be a parameter that affects the feedback. The BTEs give, in general, less feedback than ITEs and canal aids since the microphone is placed at a greater distance from the ear canal. The acoustics outside the earmould will also affect the feedback of these paths. An example of this is that hearing aids often start to howl when a hand or a telephone set is placed close to the ear. The size and type of the vent will also affect the feedback. Cox (1982) reported discrepancies in the feedback of a parallel vent and a diagonal (side branch) vent. Angaard Johansen (1975a) reported 25 dB more damping with a parallel vent than with a side branch vent at 5 kHz. The shape of the earmould will affect the leak and thus also the feedback. Further, the shape of the ear canal will also affect the leak. Jaw movements will cause changes in the shape of the ear canal in a timeframe of seconds/minutes, cerumen production and disposal in weeks or months, and aging in years (Oliviera, 1997). It can be assumed that the fit of the individual earmoulds will vary over subjects. It can also be assumed that the leak within a subject and the individual earmould will vary over insertions of the earmould.

Some of the above parameters are set at the dispensing and are constant thereafter (e.g., category of hearing aid, and type and size of vent). Others vary when the hearing aid is used (e.g., the acoustics outside the ear and the shape of the ear canal).

### C. Previous studies

Three different types of measurement of the feedback of hearing aids have been described in the literature. In all three, the hearing aid receiver was used as the sound source and the level at the hearing aid microphone relative to the level in the ear canal was the desired measure.

The first type is measurements on KEMAR or couplers, where the sound pressure level in the coupler is easily measured and used as a reference to the level at the hearing aid microphone. In the second, the level of the feedback at the hearing aid microphone is measured on human subjects, but the reference is achieved by attaching the hearing aid to a coupler and then measuring the level in the coupler. In the third type, the measurements are performed *in situ* with a probe microphone in the ear canal to obtain the reference. In some of the studies, the hearing aid microphone itself was used to measure the level at the hearing aid microphone, while a separate microphone that was not in contact with the hearing aid was used in others. All feedback paths are measured with the hearing aid microphone, while the internal feedback paths are excluded when a separate microphone is used. The benefit with measuring both internal and external feedback is that it yields the combined feedback, which controls the maximum gain that can be used. The drawback is that the results are valid only for hearing aids with the same amount of internal feedback. The external feedback will, in general, have lower magnitude than the combined. The maximum gain calculated with the external feedback can then only be achieved if the hearing aid has no internal feedback.

Lundh (1982) and Lybarger (1975) are examples of measurements on couplers. Lundh (1982) performed measurements of the combined feedback in BTE on KEMAR. Lybarger (1975) measured the external acoustic feedback to a position 2.5 cm from the vent opening of a BTE hearing aid attached to a coupler.

Angaard Johansen (1975a) measured the combined feedback of BTEs where the reference was found by measurements on a coupler. Lundh (1982) and Dyrland (1989) can also be interpreted as this type of measurement. These papers presented the 2-cc coupler gain of BTE hearing aids that gives 0-dB loop gain at all frequencies on human subjects. The reciprocal of this measure is equal to SPL of the feedback at the hearing aid microphone measured on the human subject, relative to the SPL measured in 2 cc coupler. The combination of internal and external feedback was measured in all three papers.

Grover and Martin (1974) measured the feedback of BTE instruments by measuring the level at the hearing aid microphone and in the ear canal *in situ*. Angaard Johansen (1975a) performed some measurements of the acoustic feedback on human subjects where a probe microphone positioned in the vent was used. The reference level in the ear canal was calculated from measurements of the sound pressure level in the vent.

Gatehouse (1989), Erickson and Van Tassell (1991), and Kuk (1994) measured the maximum real ear insertion gain (REIG) that could be achieved by specific hearing aids. These can be seen as indirect studies of the feedback. Gatehouse (1989) and Kuk (1994) studied BTE instruments while Erickson and Van Tassell (1991) studied ITE instruments. In the study by Kuk (1994) the effect of different earmould design was investigated. In these three studies REIG was measured when the gain of the hearing aids was set as high as possible without the hearing aid starting to howl. Feedback limited the maximum gain in only four of the nine cases studied by Erickson and Van Tassell. The results of these studies are influenced by the frequency response of the hearing aids used. When oscillation occurs at one frequency, it will limit the gain at all other frequencies. It would, however, be possible to increase the gain at these other frequencies if the gain (and phase) at the frequency that causes howling were kept unchanged. The real ear unaided response (REUR) of the subjects did also influence the results.

Some calculations and simulations of the feedback path have also been performed. Egolf *et al.* (1985) calculated the feedback path from the outside opening of the vent to the microphone position of an eyeglass-type hearing aid. Kates (1988) simulated the feedback as a part of the simulation of the hearing aid response of an ITE instrument. Angaard Johansen (1975a) and Lundh (1982) also performed calculations of the feedback.

In the present study the transfer functions of the feedback of three hearing aid categories, BTE, ITE, and ITEC, have been identified. It was then possible to calculate and to compare the maximum gain that can be achieved with the different hearing aid categories. The hearing aid microphone itself was used to measure the signal at the hearing aid microphone. The combination of external and internal feedback

was thus measured. The signal in the ear canal was measured *in situ* with a probe microphone. The transfer function of the feedback was identified in terms of both magnitude and phase. Filters with approximately the same transfer functions can then be identified. The filters can be used for simulation of the feedback and to explain the feedback.

## I. MATERIAL

### A. Test subjects and KEMAR

Measurements were performed on ten human subjects and on KEMAR, except for the ITEC where measurements were made only on five subjects. The measurements were performed on the left ear of the test subjects and on the right ear of KEMAR. Some of the subjects were hearing-impaired (sensorineural loss) and others had normal hearing. The hearing-impaired subjects used hearing aids on the ear that was used for tests. The KEMAR was equipped with a Brüel & Kjær ear simulator type 4157 (IEC 711, 1981).

### B. Hearing aids

Two types of BTEs, one type of ITE, and one type of ITEC were used. The BTEs were Oticon Personic 425 and Oticon 390 PL. They are in the following called BTE A and BTE B, respectively. They were chosen to represent a wide-band BTE and a high power BTE. The ITE was Oticon I24. Oticon manufactured the ITEC hearing aid specially for the purpose of this study.

Five samples of BTE A and BTE B were used. Test subjects 1–5 used the corresponding number of sample. Test subjects 6–10 all used sample 1. Individual ITE and ITEC instruments were manufactured for the subjects.

All hearing aids were modified by cutting the signal path. The microphone signal was fed via a signal conditioner to the AD converter of the data acquisition system. The output of the DA converter was fed via the signal conditioner to the hearing aid where the signal path was cut.

The gain of the hearing aids was measured using an ear simulator (IEC 711, 1981) and Rastronics PortaREM 2000 real ear measurement equipment. The gain of the ITE and BTE hearing aids was adjusted to give a specified gain at 1000 Hz on the ear simulator before the volume control was locked with wax to assure that the gain was kept constant throughout the measurements. The specified values were 44, 44, and 28 dB for BTE A, BTE B, and ITE, respectively. The ITECs did not have a volume control. Figure 3 shows the gain of the BTE A, BTE B, and the ITE hearing aid on the ear simulator. The gain of the hearing aid and thus the sound of the test signal in the ear canal were not crucial since the measurements were performed as relative measurements (SPL at hearing aid microphone relative to SPL in ear canal).

New shell earmoulds were used with the BTE hearing aids. The ITEC hearing aids were new but the ITE hearing aids were about 3 years old. The mean lengths of the vent, measured as the shortest distance between the two openings, were 17.6, 17.5, and 18.6 mm for the BTE, ITE, and ITEC instruments, respectively. The mean distance between vent opening and microphone, measured as the shortest distance

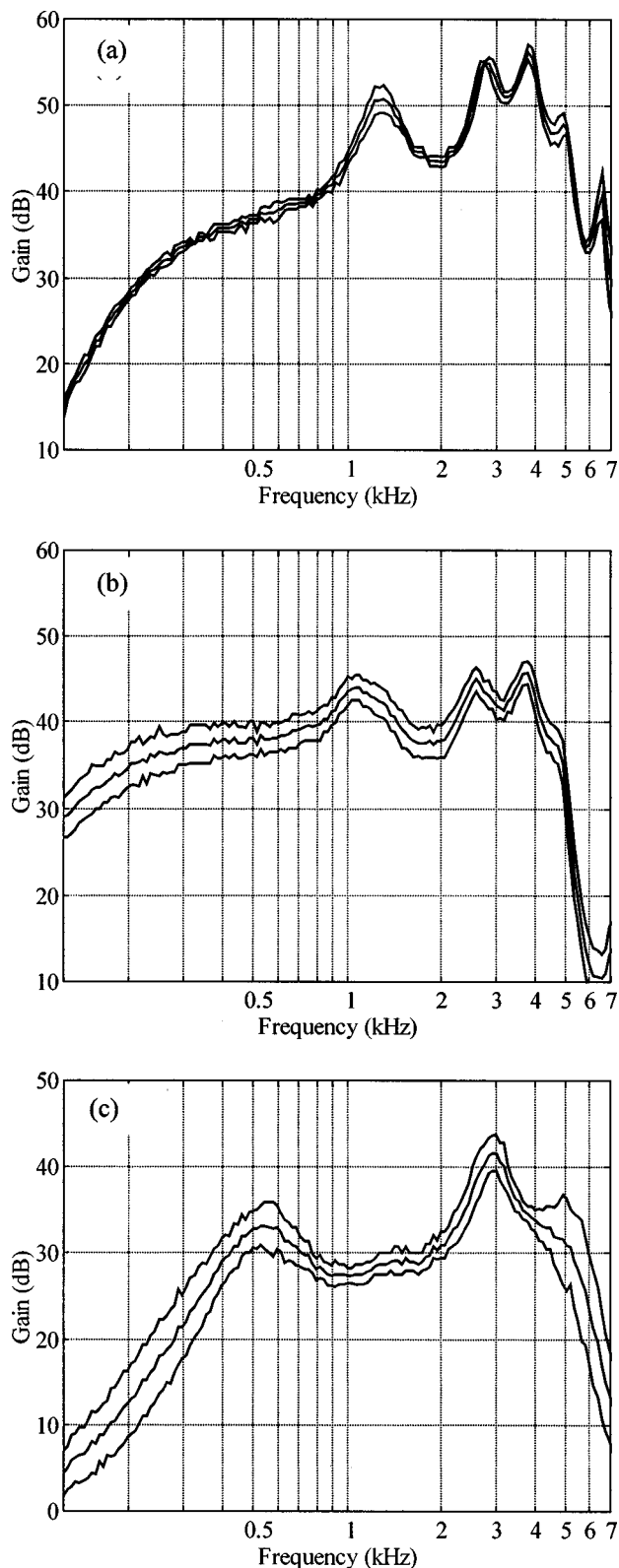


FIG. 3. Gain of the hearing aids measured on PortaREM 2000 with ear simulator, mean and mean  $\pm 1$  s.d. (a) BTE A,  $n=5$ . (b) BTE B,  $n=5$ . (c) ITE,  $n=10$ .

from the vent opening to the nearest inlet to the microphone, was 20.4 mm for the ITE and 11.6 mm for the ITEC.

The vent of the earmould used with KEMAR and BTE instruments had a varying cross section. The dimensions of the vent were approximately the same as of a tube with a

diameter of 2.2 mm and a length of 15.5 mm. The ITE tested on KEMAR had O-rings to seal the fit to the ear simulator. The length of the vent in the KEMAR ITE was 25.1 mm and the distance between vent and microphone was 16.5 mm, i.e., the KEMAR ITE had a longer vent than the average used by the human subjects but the distance to the microphone was smaller.

Skeleton earmoulds were used for the BTEs on subjects. The vent diameter was 3.2 mm for the ITEC and 2.2 mm for BTE and ITE. The 0.8-mm diameter probe tube was placed through the vent so the effective vent diameters were 3.1 and 2.1 mm, respectively.

### C. Microphones to measure sound in the ear canal

A probe microphone with a probe tube placed in the ear canal was used to measure the signal in the ear canal of the subjects. The combined probe and reference microphone PM-12 from Rastronics was used. A plastic tube with an outer diameter of 0.8 mm and an inner diameter of 0.5 mm was used as probe tube. The probe tube was inserted through the vent of the hearing aid. The reference microphone of PM-12 was used to measure the signal at the reference point below the ear. The KEMAR was equipped with a Brüel & Kjær ear simulator type 4157 (IEC 711, 1981) and a Brüel & Kjær microphone type 4134. This microphone is located at the position that corresponds to the tympanic membrane while the probe microphone that was used on the human subjects was placed in the ear canal, a couple of mm from the tip of the earmould.

A sound field was used to analyze the difference between the characteristics of the two microphones used (probe microphone and hearing aid microphone, coupler microphone and hearing aid microphone, or reference microphone and hearing aid microphone). The two microphones were placed close together in the sound field. The wide band noise of a Madsen OB 822 audiometer was fed via an amplifier to a loudspeaker inside an anechoic room. The absolute level and spectrum of this signal was not critical since it was only used to analyze the difference between two microphone signals.

### D. Data acquisition system

The data acquisition system was based on a Loughborough DSP board in a host PC. It was used to convert the signals of the microphones to digital signals and to store them on the hard disk of the PC. The DSP was also used to generate the test noise signal applied to the receiver via the DA converter. The test signal and the microphone signals were stored on the hard disk for off line analysis. The DSP board has two 16-bit AD converters and two 16-bit DA converters of sigma delta type. A sampling frequency of 15.75 kHz was used throughout the measurements. The data acquisition system used antialiasing filters and reconstruction filters with a cutoff frequency of  $0.45f_s$ , where  $f_s$  is the sampling frequency. The cutoff frequency of the data acquisition system was thus 7.1 kHz. The signals from the microphones were amplified by the signal conditioner to fit the sensitivity of the AD converters.

### E. Measurement setup

The test subject was seated in a chair in the center of an anechoic room to provide low ambient noise levels during the measurements. The modified hearing aid was connected via a cable to the signal conditioner placed on the back of the subject's chair in the anechoic room. The computer with the acquisition system was placed in the adjoining control room.

## II. METHODS

### A. Spectral analysis

The Matlab command "spa" of the System Identification Toolbox was used to calculate transfer functions (both amplitude and phase) between recorded signals. When spectral analysis is used, the auto covariance of the input signal and the cross covariance between the input signal and the output signal are estimated. The covariances are then windowed and Fourier transformed. A Hanning window that was nonzero for 255 samples ( $-128 < \tau < 128$ ) was used. Thus, correlation in the time window from  $-8.1$  to  $8.1$  ms was considered in the analysis. An estimate of the transfer function is formed by the ratios between the Fourier transforms. The windowing causes smoothing of the transfer function. Ljung and Glad (1994) states that the frequency resolution with a Hanning window of width  $\gamma$  will be approximately

$$\frac{\pi}{\gamma\sqrt{2}} \text{ radians/time unit.} \quad (6)$$

The window used thus gives a frequency resolution of about 44 Hz.

A sequence of 5000 samples (about 0.3 s) was used for each case that was analyzed. The theory of this spectral analysis is described in Ljung and Glad (1994).

### B. Identification of the AFE

All feedback paths discussed above contribute to the feedback in a hearing aid. To deal with these combined paths a new path was defined, the acoustic feedback equivalent (AFE). The AFE was defined as the equivalent sound pressure level caused by all feedback paths, and referred to the hearing aid microphone relative to the sound pressure level in the ear canal. In this way, the different feedback paths of Fig. 2 are transformed to one single common feedback path ( $G_3$ ) as of Fig. 1. The AFE can also be expressed as the loop response, caused by all the feedback paths present, divided by the open loop gain from the microphone to the earmould tip ( $G_2$  in Fig. 1).

Not all feedback paths from the receiver to the hearing aid microphone pass through the ear canal. Internal feedback and transmission through the tubing wall are examples of this. By the introduction of the AFE we can treat the feedback as if it was one (acoustic) path from inside the ear canal to the hearing aid microphone. The effect on the closed loop gain will be the same with AFE as the actual feedback paths. A possible problem is if there is internal feedback in the hearing aid and a notch in the transfer function from the receiver to the ear canal. The AFE will then tend to approach infinity at this frequency. However, it can be assumed that

this will not occur since the hearing aids do not have complete notches within the frequency range of interest.

To measure the AFE, a broadband signal was applied to the receiver. The signals in the ear canal and at the hearing aid microphone were then recorded. Spectral analysis was used to obtain a spectral representation (both amplitude and phase) of the transfer function from the probe microphone (coupler microphone in case of KEMAR) to the hearing aid microphone. After correction for differences between the two microphones, the AFE was obtained.

### C. Transfer function from common reference point to hearing aid microphone

The transfer functions from a reference microphone under the ear to the microphone of the different hearing aids on the human subjects were identified. This was used to calculate the maximum REAG with a common reference point. The measurements and calculations were made in the same way as when the AFE was identified, but the reference microphone was used instead of the probe microphone and the signal was the wideband noise from the clinical audiometer presented in a free sound field.

### D. Estimation of impact of feedback paths other than via vent and leakage

The AFE is the feedback from the ear canal to the hearing aid microphone that together with the open loop response of the hearing aid gives the same transfer function as the actual loop response caused by all feedback paths present. In an ideal situation only feedback via the vent would be present, while without vent no feedback would exist. With a vent, the acoustical feedback from the ear canal to the hearing aid microphone and the AFE would then be identical. In reality there will be some feedback caused by unwanted feedback paths such as internal mechanical vibrations, electrical feedback, and emission from tubing walls and the shell of the hearing aid. To identify the impact of these unwanted factors, the loop response was identified when the hearing instrument was attached to the ear simulator. No vent was used in these measurements. This (complex) loop response was subtracted from the loop response measured with the hearing aid on the subjects. This gave an estimate of the loop response caused by feedback via the vent and the leakage around the earmould.

The loop response was identified with spectral analysis. The AD and DA converter introduced a 54-sample group delay. This gave problems in the spectral analysis as a portion of the correlation between the input and output could be found outside the window used. To overcome this problem the input signal (signal to the DA) was delayed 64 samples. The results of the analysis were then corrected for the delay. This was not a problem when the other transfer functions were identified, as the group delay was much smaller.

## III. RESULTS

### A. AFE with BTE A

Figure 4 shows the AFE of BTE A measured on KEMAR. The magnitude function is fairly flat up to 2 kHz.

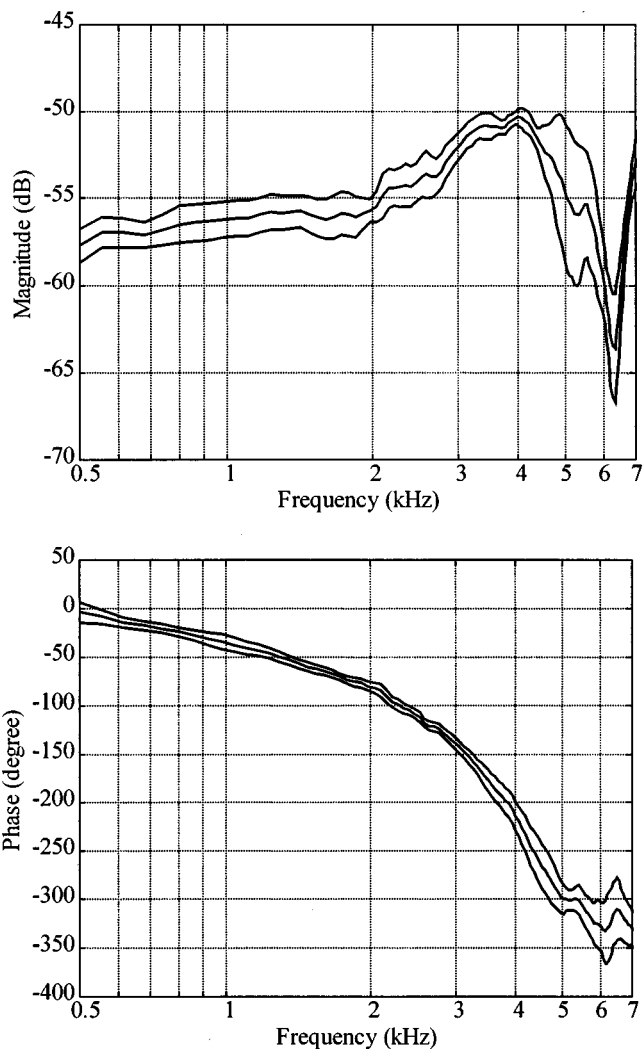


FIG. 4. The AFE on KEMAR measured with five different samples of BTE A. Mean and mean  $\pm$  s.d. of the magnitude and phase.

Then there is a peak with low Q-value at about 4 kHz. The standard deviation across measurements is on the order of 1 dB for frequencies below 4 kHz. At higher frequencies the standard deviation increases. The phase is close to linear for frequencies below 5 kHz. The change of phase is about  $-60^\circ$  to  $-70^\circ$  per kHz up to 5 kHz.

Figure 5 shows the mean and mean  $\pm 1$  standard deviation of the AFE of human subjects with BTE A together with the mean result of measurements on KEMAR. The mean AFE increases from about  $-64$  dB at 1 kHz to about  $-42$  dB at 7 kHz. The large amplitude of the AFE at 7 kHz could be due to a resonance in the vent, but it is hard to make any conclusions since it is on the limit of the bandwidth of the acquisition system.

Figure 6 shows the standard deviation for the magnitude of the AFE on each subject when the measurement is repeated (hearing aid removed between measurements). This was calculated using five measurements from each of the five subjects and assuming that they had different means but the same standard deviation. Each subject used an individual sample of BTE A. The standard deviation was close to 2 dB

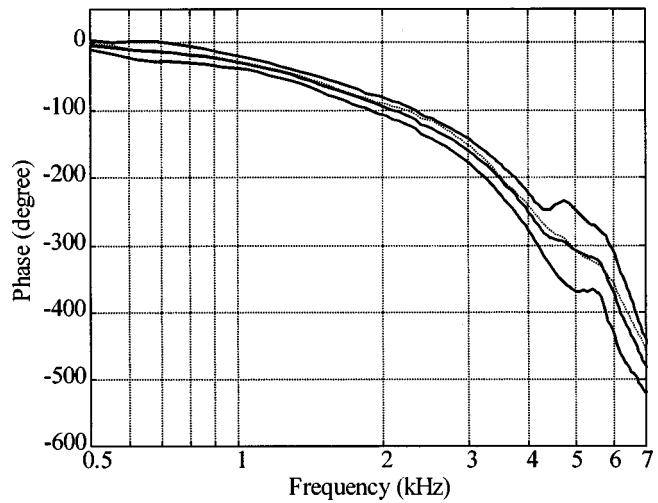
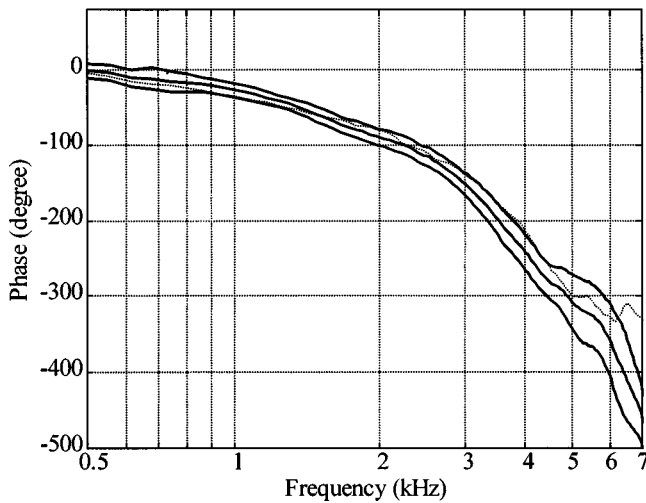
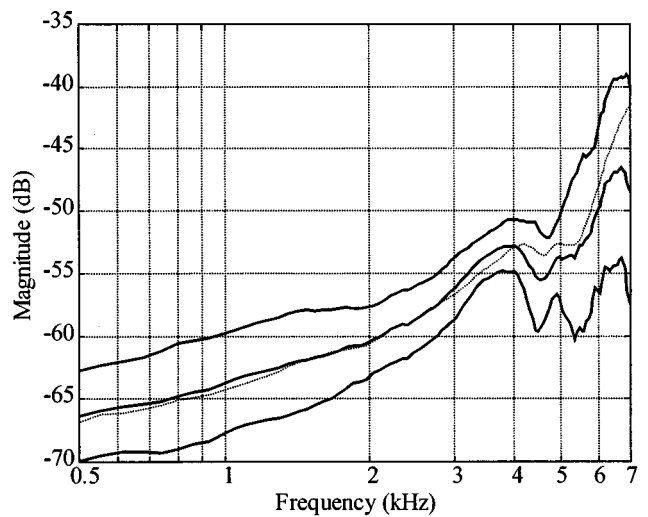
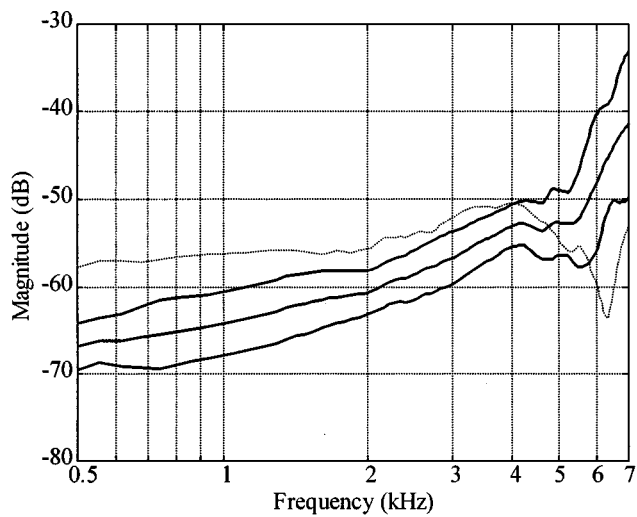


FIG. 5. (a) The AFE on human subjects measured with BTE A. Mean and mean  $\pm$  1 s.d. of the magnitude and phase,  $n=10$ . Dashed line is mean AFE of KEMAR,  $n=5$ .

FIG. 7. The AFE measured on ten human subjects with BTE B,  $n=10$ . Mean and mean  $\pm$  1 s.d. of the magnitude and phase. Dashed lines are mean measured on human subjects with BTE A.

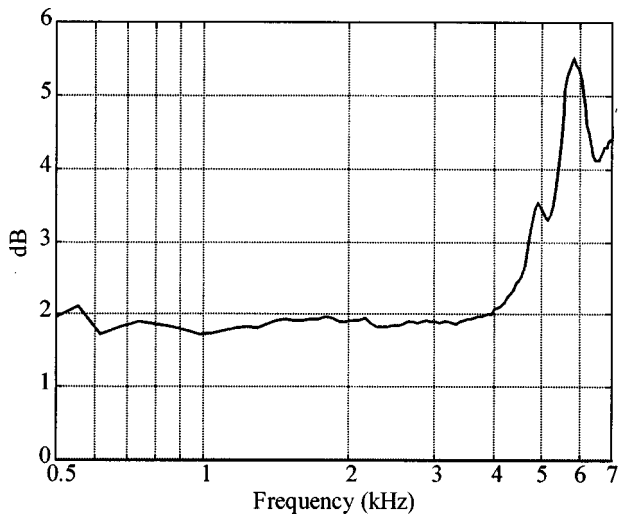


FIG. 6. Standard deviations of measurements of AFE when earmould is reinserted. Measured on five subjects and five insertions of the earmould for each subject. Each subject used a specified sample of BTE A,  $n=25$ .

for frequencies up to 4 kHz and increased at higher frequencies.

Figure 7 shows the AFE on the human subjects using BTE B. The mean AFE measured with BTE B is comparable to the mean AFE obtained with BTE A, which is also included in Fig. 7.

### B. AFE with ITE

Figure 8 shows the AFE with the ITE hearing aid. The solid lines are data from human subjects and the dotted lines are data from KEMAR. Only one measurement was performed on KEMAR since only one ITE was manufactured for KEMAR and variations between insertions can be assumed to be small as the fitting in the ear canal was sealed with O-rings. The large standard deviation can to some extent be explained by an extreme case with a magnitude about 20 dB higher than the mean. The change of phase on the

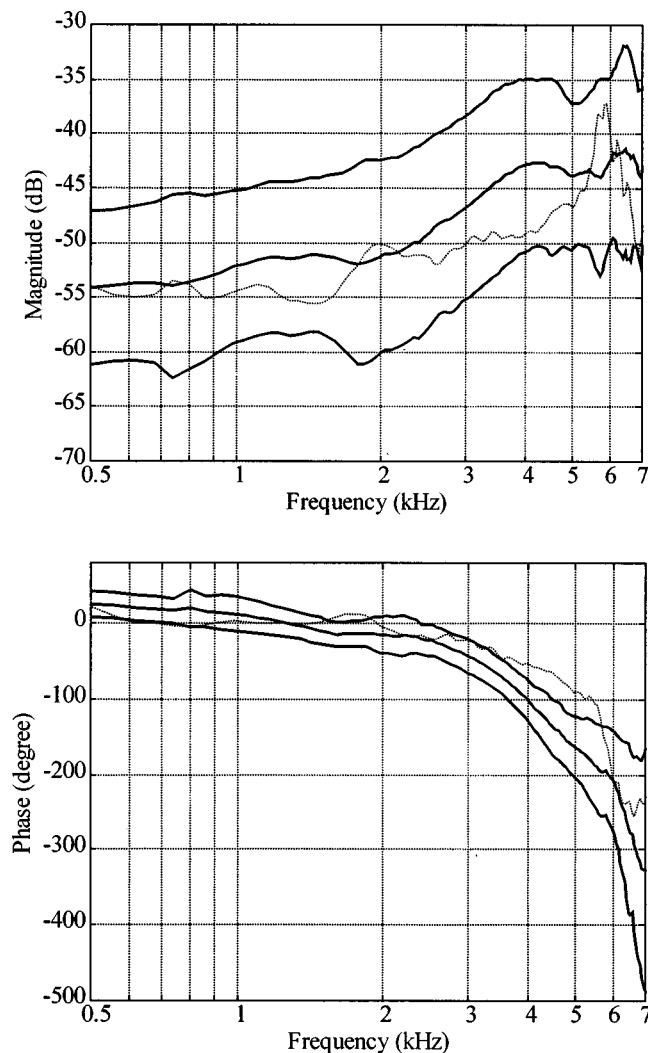


FIG. 8. The AFE measured with ITE. Solid lines are mean and mean  $\pm$  s.d. of magnitude and phase measured on human subjects. Dashed lines are data from KEMAR with ITE.

human subjects is quite linear at about  $-40^\circ$  for each kHz. On KEMAR the change of phase versus frequency is less linear.

Figure 9 shows the standard deviation of the acoustic path when the same ITE is inserted five times on subject 1. The standard deviation was on the order of 1 dB for frequencies up to 4 kHz.

### C. AFE with ITEC

Figure 10 shows the AFE of five human subjects (1–5) with the ITEC hearing aids. No measurements on KEMAR with ITEC were performed. The change of phase is about  $-100^\circ$  from 1 to 4.5 kHz. The standard deviation is quite small for frequencies up to 5 kHz. At 6.5 kHz the phase of some measurements differed by approximately  $360^\circ$ , which caused a large standard deviation.

Figure 11 shows the standard deviation of the AFB when the same ITEC was inserted five times on subject 1.

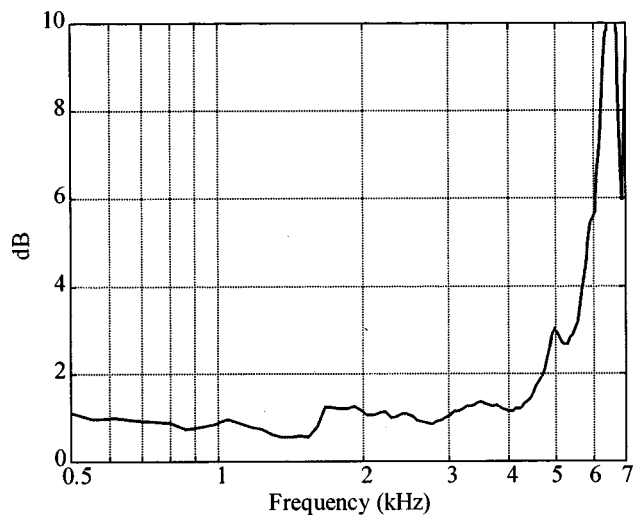


FIG. 9. Standard deviation of magnitude of AFE when the ITE hearing aid was inserted five times in the same ear canal. Measured on subject 1.

### D. Reference microphone to hearing aid microphone transfer function

Figure 12 shows the sound level at the hearing aid microphone relative to the sound level at the reference microphone for BTE, ITE, and ITEC, respectively. Some data were excluded from the analysis due to technical failures during the measurements.

### E. Maximum REAG

The mean curves of Fig. 12(a)–(c) were used together with the corresponding AFEs to calculate the maximum REAG given in Fig. 13. These curves are the mean upper limit of REAG (with reference point under the ear) for the three hearing aids.

### F. Impact of feedback paths other than the acoustic from ear canal

The mean magnitude of the estimate of the loop response caused by the vent and leakage of the earmould was, in the frequency range from 1 to 4.5 kHz, within 1 dB of the measured loop response for all four types of tested hearing aids. Outside this frequency range the discrepancy was greater than 1 dB for some of the tested hearing aids, but less than 3 dB. The same discrepancy will be found between the AFE and the transfer functions of the acoustic paths from the ear canal to the hearing aid microphone, under the assumption that the feedback caused by the other feedback paths are the same on the coupler and on the subjects. The largest discrepancies were found in cases with small magnitude of the measured loop response. The conclusion is that the feedback via the vent and leak around the earmould was the dominating feedback path, and that the measured AFE is a good estimate of the acoustic feedback path, especially in the frequency range from 1 to 4.5 kHz.

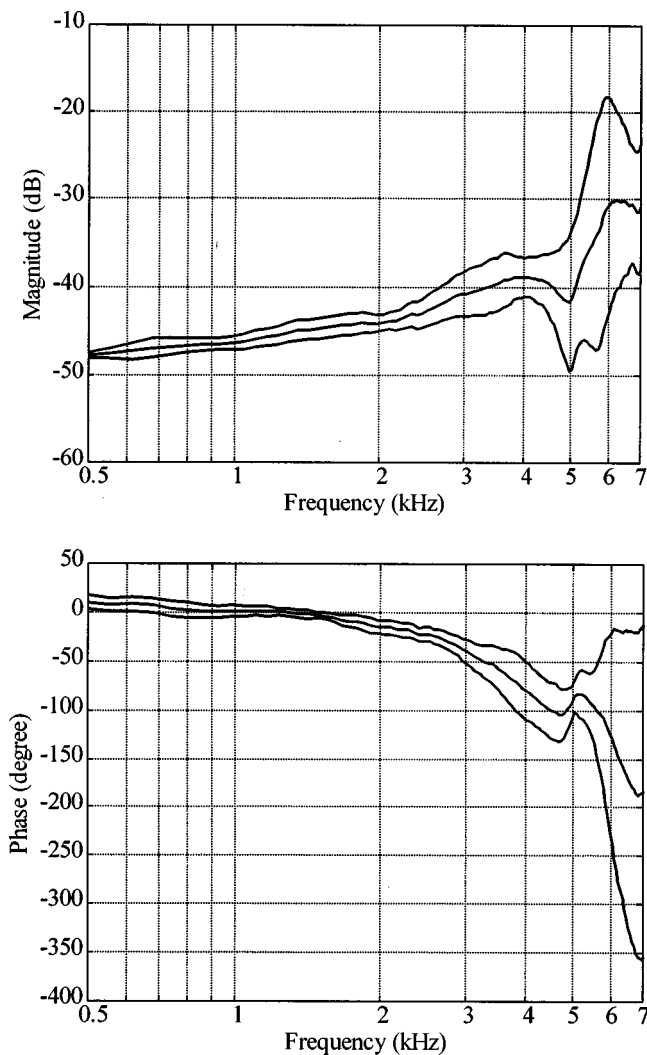


FIG. 10. The AFE measured with ITEC on five human subjects. Mean and mean  $\pm$  s.d. of magnitude and phase. [One of the measurements of the difference in frequency response between the probe microphone (with individual probe tube) and the hearing aid microphone could not be used since it by mistake was made with a narrow-band noise instead of a broadband noise. For this subject the mean value of the difference between the microphones was used when the AFE was calculated.]

#### IV. DISCUSSION

##### A. Comparison of results with BTE A and BTE B

The AFE with BTE A at 2 kHz was found to have mean magnitudes of  $-55.6$  and  $-60.6$  dB measured on the KEMAR and human subjects, respectively. The corresponding figures for BTE B were  $-54.5$  and  $-60.4$  dB. Two-sided  $t$ -tests were used to evaluate if the difference between the results measured with the two BTE instruments was significant. The  $t$ -test for dependent samples used to test the difference measured on human did not reveal a significant difference ( $p=0.71$ ). The  $t$ -test for dependent samples was used since it can be assumed that the individual ears and earmoulds affect the acoustic feedback. On KEMAR the ear as well as the earmould were held constant. Thus, a  $t$ -test for independent samples was used. The difference of the AFE measured with the two hearing instruments on KEMAR was not found to be significant ( $p=0.41$ ). Such differences, had

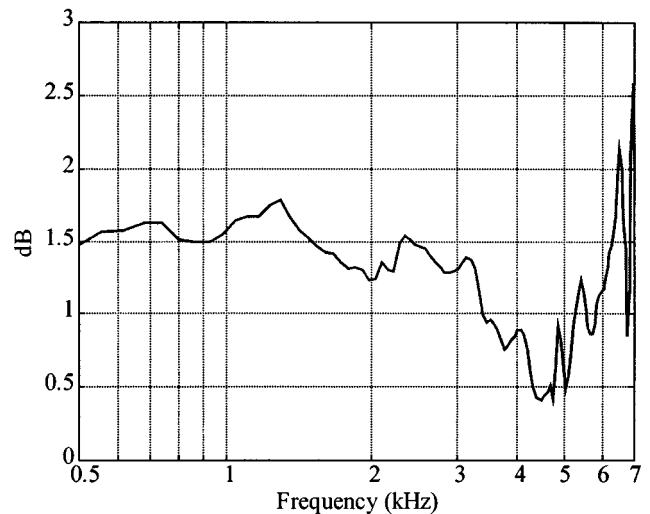


FIG. 11. Standard deviation of AFE when an ITEC hearing aid was inserted five times. Measured on subject 1.

they been present, could be caused by differences in the internal feedback or in the directivity of the microphones.

##### B. Comparison of results measured on human subjects and KEMAR

The magnitudes of the AFE at 2 kHz measured on the human subjects and on KEMAR were compared. The mean AFE at 2 kHz measured with the BTE hearing aids was  $-60.5$  dB ( $n=20$ ) and  $-55.1$  dB ( $n=10$ ) for human subjects and KEMAR, respectively. The corresponding values for the ITE hearing aid were  $-51.2$  dB ( $n=10$ ) and  $-50.2$  dB ( $n=1$ ). Data from both BTE A and BTE B were used together in this analysis. A two-sided  $t$ -test for independent samples revealed a significant difference ( $p<0.001$ ) for the BTE hearing aids, while no significant difference was found for the ITE ( $p=0.91$ ). The phase of AFE on KEMAR agreed fairly well with data from human subjects for both ITE and BTE.

Possible reasons for differences in the data measured on human subjects and KEMAR are differences in position of reference, leak around the earmould, the residual ear canal, vent, and pinnae.  $O$ -rings were used to seal the fit of the ITE hearing aid used on KEMAR, while a soft earmould were used with the BTE hearing aids. This could cause some more feedback via leak for the BTE instrument. The difference in position of the reference is one reason for the difference, which is further discussed below.

##### C. Variances in the AFE

The standard deviation between subjects with ITEC, found in Fig. 10, is on the order of 1–2 dB for frequencies up to 3 kHz. This is lower than the standard deviation with BTE A, BTE B, and ITE as shown in Figs. 5, 7, and 8, respectively. At 2 kHz the standard deviations are 2.6, 2.8, 8.9, and 1.0 dB for BTE A, BTE B, ITE, and ITEC, respectively. The large vent of the ITEC can explain the small standard deviation, since it provides a well-defined and dominating feedback path through the vent. Presumably a smaller vent would



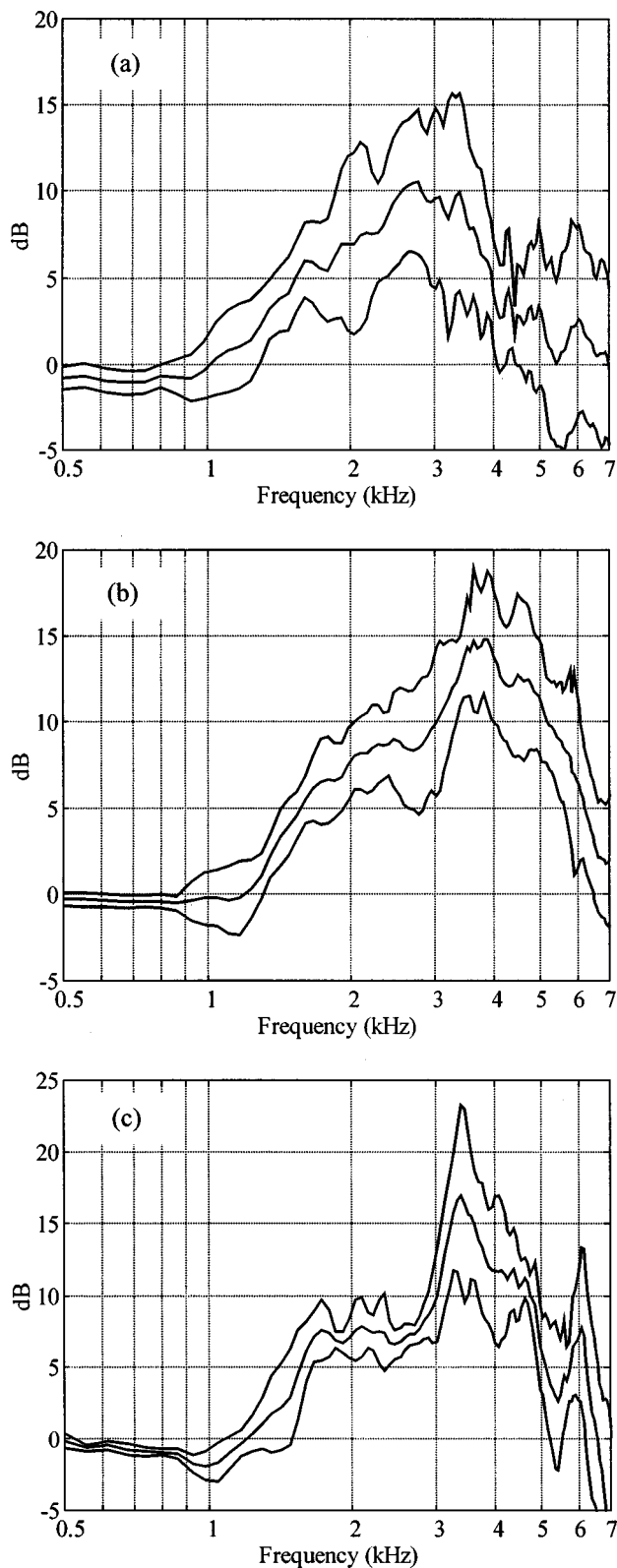


FIG. 12. Transfer function from reference point under the ear to position of hearing aid microphone. Mean and mean  $\pm$  s.d. (a) BTE A,  $n=9$ . (b) ITE,  $n=8$ . (c) ITEC,  $n=3$ .

give larger variations because then the leakage due to improper fit will have larger relative influence on the AFE. The largest standard deviation was found with ITE hearing aids. No significant correlation with the length of the vent ( $p=0.59$ ), nor with the distance between vent opening and

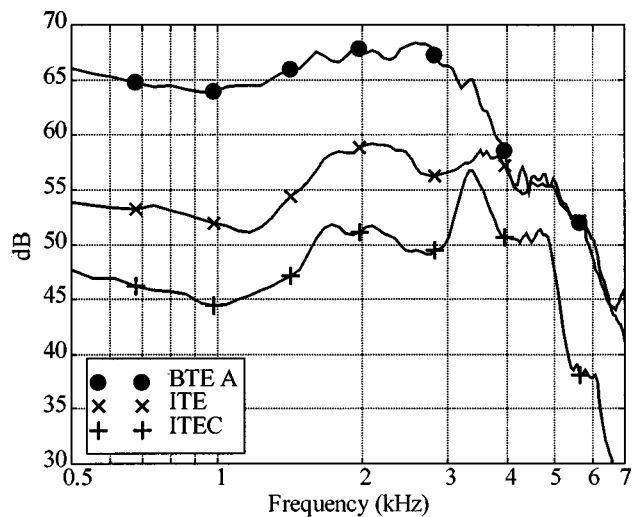


FIG. 13. Maximum REAG for BTE, ITE, and ITEC. BTE and ITE used a 2.1 mm vent while the ITEC used a 3.1 mm vent. The reference point is under the ear.

microphone ( $p=0.19$ ), was found. There is one extreme case with an AFE at 2 kHz of 18 dB over the mean that contributed much to the large variance. The most likely reason for the large standard deviation with ITE is variation in the leak around the earmould.

The standard deviation of the attenuation when the BTE hearing aids are inserted repeatedly on the same subject is about 2 dB for frequencies up to 4 kHz, as can be seen in Fig. 6. The corresponding values for ITE and ITEC are about 1 dB (Fig. 9) and about 1.5 dB (Fig. 11), respectively. The larger vent diameter of the ITEC, and thus the lower sensitivity to leak, is a possible explanation for the lower value of the ITEC relative to the BTE. An explanation for the low value of the ITE relative to ITEC and BTE has not been found. It can be assumed that variations in the leak is the primary reason for the variance, and that larger variations would have been found if the measurements had been separated more in time.

#### D. Comparison to feedback of tube in an infinite baffle

The acoustic feedback from the tip of the earmould to the hearing aid microphone can be approximated with a tube with an open end in an infinite baffle. The theoretical transfer function of this approximation was compared with the measured AFE. Equations (A1) and (1) of Kates (1988) were used to calculate the pressure transfer function across the tube. The pressure transfer function from the opening in the baffle to the position of the hearing aid microphone was calculated using Eq. (4) of Egolf *et al.* (1985). Figure 14 shows the calculated transfer function and the mean AFE measured on human subjects for BTE A, ITE, and ITEC. The length,  $l$ , and diameter,  $\Phi$ , of the tube and the distance between opening in the baffle and position of interest,  $R$ , were set to correspond to the means in the measured cases. An angle of  $90^\circ$  from the vent centerline to the point of interest was used in all three cases (microphone close to the baffle).

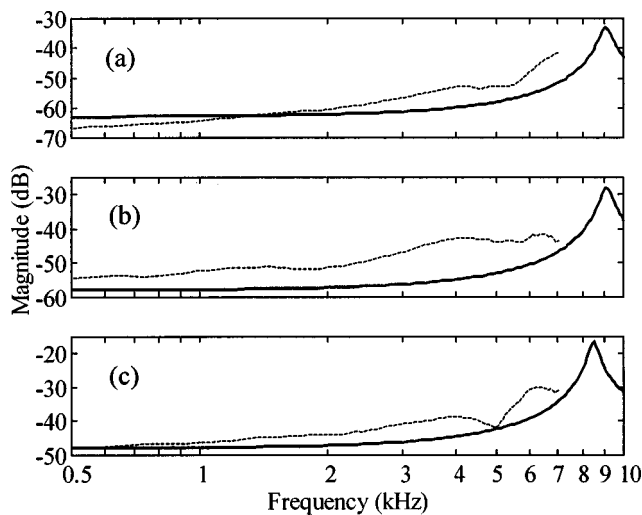


FIG. 14. Calculated transfer function of feedback from tip of earmould to hearing aid microphone (solid line), mean AFE measured on subjects (dotted line). (a) BTE A:  $\Phi = 2.1$  mm,  $l = 17.6$  mm,  $R = 40$  mm. (b) ITE:  $\Phi = 2.1$  mm,  $l = 17.5$  mm,  $R = 22.4$  mm. (c) ITEC:  $\Phi = 3.1$  mm,  $l = 18.6$  mm,  $R = 13.6$  mm.

In all three cases the calculations give similar shapes of the magnitude function, with a close to constant magnitude for frequencies up to 2 kHz and a resonance with a peak at approximately 9 kHz. The mean AFE measured on human subjects shows a more sloping characteristic. The results with BTE A on KEMAR show a fairly constant attenuation for frequencies below 2 kHz, a peak with low Q-value at 4 kHz, and a dip at 6.3 kHz. At 4 kHz, both the results from KEMAR and human subjects show greater magnitude of the feedback than the calculations. The discrepancy between measured feedback on human subjects and the feedback with an infinite baffle is discussed in Egolf *et al.* (1985), Kates (1988), and Angaard Johansen (1975a). In the last work a discrepancy of about 5–10 dB was found for frequencies between 2 and 4 kHz. In Egolf *et al.* (1985), the measured feedback from the vent opening to the position of the microphone of an eyeglass-type hearing aid on a manikin was at 5.7 kHz found to be about 6 dB higher than calculated with an infinite baffle model. The effect of the pinnae is, in all three of these papers, suggested as a probable reason for the discrepancy in the frequency region above 2 kHz.

Our calculations of the feedback with tube and infinite baffle were modified to incorporate the pinnae. This was done by attaching the tube corresponding to the vent to a second tube that ended in the baffle. The second tube had a diameter of 20 mm and a length of 7 mm. These values were chosen to correspond to the dimensions of the residual concha. The feedback at low frequencies was then close to unaltered, while it increased by about 5.5 dB at 5.5 kHz. This supports the hypothesis that the discrepancy around 5 kHz between the measured AFE and the model with a single tube and an infinite baffle is caused by the absence of the pinnae in the model.

One factor affecting the large difference in magnitude of measured AFE with ITE compared to the model is that there was one extreme case in the measurements, which had about

20 dB higher AFE than the mean and thus increased the mean AFE by the order of 2 dB.

The ITE and ITEC hearing aids used had a vent opening positioned below the tragus, while the microphone was placed in the cavity given by the residual concha. It can thus be argued whether the model with an infinite baffle is suitable for these ITE and ITEC hearing aids.

### E. Impact of reference position in the ear canal

The acoustic system from a position in the ear canal to the tympanic membrane can be assumed to be passive and linear. Thus, the pressure transfer function from a position in the ear canal to the tympanic membrane is independent of the sound source. Sesterhenn *et al.* (1998) measured the sound pressure level at different distances from the tympanic membrane in a model of the outer ear. At a distance of 15 mm from the tympanic membrane the sound pressure level of a 6-kHz tone was about 28 dB lower than at 2 mm from the tympanic membrane. The corresponding figures at 3 and 8 kHz were about 1 and 6 dB, respectively. This shows that there is a peak in the transfer function from the earmould tip to the tympanic membrane. The explanation for the peak is standing waves caused by reflection in the tympanic membrane. The frequency of the peak will depend on the distance to the tympanic membrane. The peak will show up as a dip in the acoustic feedback when the reference is the tympanic membrane. This is the probable reason for the dip at 6.3 kHz in the AFE measured on KEMAR with BTE A, found in Fig. 4. This dip can also be found in Fig. 7.2.1 of Lundh (1982), where calculations of the feedback at the hearing aid microphone relative to the coupler microphone are presented.

A dip at approximately 5 kHz was also found on the individual AFE measured on human subjects. The frequency and depth of the dip differed among subjects. The effect of the dip is barely noticeable in the mean AFE, except in the case of ITEC hearing aids. The probe microphone used on human subjects was placed a couple of millimeters from the earmould tip. The effect on AFE of different probe positions is given by the inverse of the transfer function from tip of earmould to the probe. This can be calculated from the data by Sesterhenn *et al.* (1998). Figure 15 shows the estimated effect on the measured AFE if the probe is placed 6 mm from the tip of the earmould instead of at the tip. The tip of the earmould was assumed to be 16 mm from the tympanic membrane. The figure shows a dip at 5.5 kHz and an increase at higher frequencies. The peak in the transfer function from the earmould tip to the tympanic membrane causes the dip, and the peak in the transfer function from the probe microphone to the tympanic membrane causes the increase at higher frequencies.

It can thus be assumed that the dip found in the AFE on human subjects was caused by the position of the probe. Further, it can be assumed that the sloping characteristics at the highest frequencies in the AFE on human subjects partly were caused by the position of the probe.

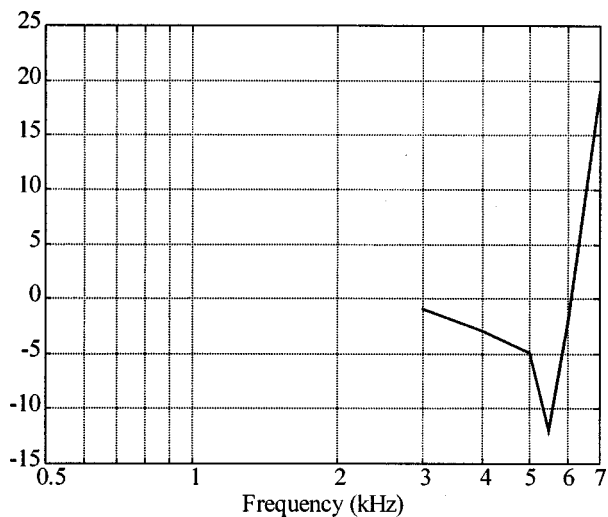


FIG. 15. The effect on AFE of placing the probe 6 mm from the earmould tip relative to the earmould tip. The earmould was supposed to be 16 mm from the tympanic membrane. Data from Fig. 7 of Sesterhenn *et al.* (1998).

## F. Comparison with previous studies

Table I shows the measured feedback at 2 kHz on BTE instruments of this and previous studies. The attenuation of the feedback ranges from 42 to 65 dB. Vent size, leak between earmould and ear canal wall, position of reference, and dimensions of the pinnae are parameters that vary between studies and will affect the feedback. To compare the effect of the vent in the different studies, an approximate normalization was performed.

The normalization was based on the model of the acoustic feedback path with a tube with an end in an infinite baffle. The transfer function across the tube was calculated with Eqs. (A1) and (1) of Kates (1988). The difference in transfer function with a vent of length 17.6 mm and diameter 2.1 mm, and with the vents of the different studies were used to normalize the results. Normalization of the impact of distance from vent opening to the microphone was not performed. Angaard Johansen (1975a) identified the mean leak

TABLE I. Measurements of feedback of BTE hearing aids. In the column labeled "Normalized" the results were converted to results at 2 kHz with a vent with a length of 17.6 mm and a diameter of 2.1 mm.

Study	Subjects	Position of reference	External or combination of external and internal feedback	Mean vent length (mm)	Vent diameter (mm)	Mean magnitude of feedback at 2 kHz (dB)	Normalized (dB)	Note
Present	Ten human subjects	Probe microphone in ear canal	Combined	17.6	2.1	-60.5	-60	a
Grover and Martin (1974)	Human subjects	Probe microphone in ear canal	External	No vent	No vent	-61	-55	b
Angaard Johansen (1975b)	One human subject	Ear canal	External	~22	~3	-54	-59	c
Angaard Johansen (1975a)	Four human subjects	2 cc coupler	Combined		Bass tube No vent	-56 -61	-61 -55	d
Dyrlund (1989)	Human subjects	2 cc coupler	Combined	No vent	No vent	-64	-58	
Lundh (1982)	Human subjects	2 cc coupler	Combined	No vent	No vent	-64 -61	-58 -55	e
Lybarger (1975)	Human subject	2 cc coupler	External		3.2 0.79 no vent	-42 -46 -60	-66 -51 -54	f
Present study	KEMAR	IEC 711 ear simulator	Combined	15.5	2.2	-55.1	-57	
Lundh (1982)	KEMAR	Coupler	Combined	11.5	1.65	-65	-63	g

<sup>a</sup>Mean of results with the two types of used instruments.

<sup>b</sup>Both forward facing and downward facing as well as skeleton earmoulds as well as shell earmoulds gave about the same feedback at 2 kHz.

<sup>c</sup>Reference found by measurements with a probe microphone in the vent and correcting the results by the ratio of the sound pressure level in the vent and in "ear canal" measured on a coupler.

<sup>d</sup>Results corrected for ratio between sound pressure level in 2 cc coupler and real ears, stated to be 8 dB at 2 kHz.

<sup>e</sup>Results with top-placed and bottom-placed microphone.

<sup>f</sup>Positive venting valve. Correction factor calculated for a vent with a length of 1 mm.

<sup>g</sup>Both top-placed and bottom-placed microphone.

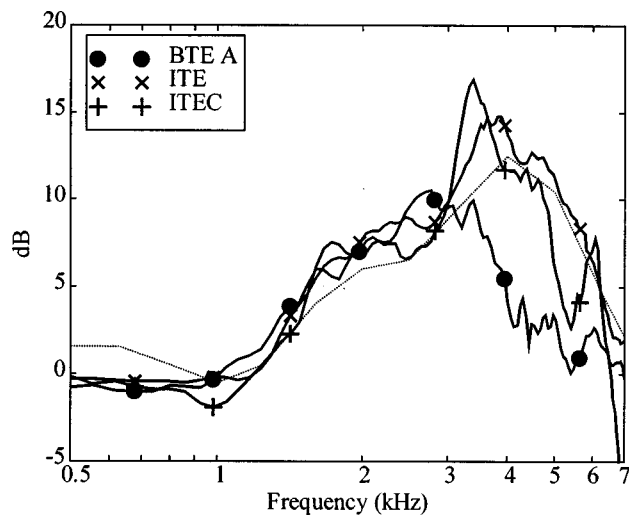


FIG. 16. Mean transfer functions from reference microphone under the ear to the position of the hearing aid microphone for different categories of hearing aid. The dotted line is free field frequency response for blocked entrance of the ear canal given in ISO/WD11904-1,  $n = 12$ .

of a BTE earmould to correspond to a vent with a diameter of 1.58 mm and a length of 22 mm. These dimensions of the vent were used for the cases where no vent was used.

It can be argued whether the results by Lybarger (1975), where so-called positive venting valves were used, can be normalized by the model with homogenous cross section of the vent. If the results of Lybarger are omitted, the normalized results are within an 8-dB range (−63 to 55 dB). The results of the present study (−60 dB with human subjects and −57 dB with KEMAR) agree reasonably with those of the other studies.

### G. Reference microphone to hearing aid microphone transfer function

Figure 16 shows the mean transfer functions from the reference point under the ear to the position of the hearing aid microphone of BTE A, ITE, and ITEC from Fig. 12(a)–(c). All three curves are fairly close together up to 3 kHz. Above 3 kHz the sound level at the BTE hearing aid microphone is lower than the level at ITE and ITEC hearing aid microphones. The mean curves of ITE and ITEC resemble the free field frequency response for the blocked entrance of the ear canal given in ISO/WD11904-1 (ISO, 1997) which is shown by a dotted line in Fig. 16. It should be noted, however, that the reference point of the latter is the midpoint of the line connecting the subject's ear canal openings (with test subject absent). Figure 16 can be compared to Fig. 10 of Fikret-Pasa and Revit (1992) where the effects of ITE and ITEC microphone location relative to the over-the-ear location were presented. The sound levels at the ITE and ITEC microphones were about the same as at over-the-ear location for frequencies up to 2 kHz, but about 10 dB higher at 4 kHz.

### H. Maximum gain of a hearing aid

The mean maximum REAG with the reference point under the ear presented in Fig. 13 shall be interpreted as the upper limit of REAG for BTE A with 2.1-mm vent, the ITE

with 2.1-mm vent, and the ITEC with 3.1-mm vent. If the internal feedback is neglected, the curves of Fig. 13 can be generalized to be valid for the three categories of hearing aids with the specified vent sizes. The limit is independent of the frequency response of the hearing aid. Thus, this limit cannot be exceeded by selection of another hearing aid of the same category. Changing earmould, vent dimensions, and category of hearing aid can change the AFE and the limit of the gain. There will be large individual variations of the maximum REAG. This is due to individual variations in reference microphone to hearing aid microphone transfer functions, as seen in Fig. 12, as well as individual variations of the AFE as seen in Figs. 5, 7, 8, and 10.

The BTE A had a higher maximum REAG than the ITE for frequencies below 3 kHz, but they were about the same for frequencies above 4 kHz (Fig. 13). The mean difference was 10.5 dB for frequencies below 3 kHz, and 0.2 dB for frequencies above 4 kHz. The difference at low frequencies is mainly caused by differences in the AFE, while these differences are canceled by the transfer function to the hearing aid microphone at higher frequencies. The difference between maximum REAG of ITEC and ITE was more constant. The mean difference over the entire frequency range was −8.3 dB. The main reasons for the difference between these two can be assumed to be the difference in vent size and distance between vent opening and microphone.

### 1. The effect of the transfer function of the hearing aid

The gain that can be provided by a hearing aid is highly dependent on the hearing aid transfer function. The gain of the hearing aid can be increased until at some frequency the limit determined by feedback is reached. The margin to this limit at other frequencies cannot be utilized unless the frequency response is altered. The hearing loss, and thus the prescribed gain, often increases with frequency. The measured AFEs show a decreasing attenuation with increasing frequency. Thus, problems with oscillations will mainly occur at high frequencies. This agrees well with clinical experience (Dyrlund and Lundh, 1990). In some cases it will be beneficial for a hearing aid fitting to reduce the gain at the frequency of oscillation to allow for increased gain at other frequencies.

Oscillation will only occur at frequencies where the phase of the loop response is zero (or  $n360^\circ$ ). The phase of the loop is controlled by the AFE as well as the hearing aid transfer function. The AFE of the BTE, ITE, and ITEC has from 500 Hz to 6 kHz phase shifts of about  $360^\circ$ ,  $200^\circ$ , and  $120^\circ$ , respectively. Assuming that the hearing aid does not delay the signal from the microphone to the earmould tip, not more than one zero crossing would be found in this frequency region. Thus, there would be at most one frequency where oscillation could occur. In reality there are delays in the signal processing of the hearing aid as well in the acoustic path from the receiver to the earmould tip. The loop responses of the hearing aids used are shown in Fig. 17. The phase presented in Fig. 17 is the phase of the loop without the 54-sample (3.4 ms) group delay caused by the AD and DA converters that were considered parts of the acquisition system. The loop responses show larger amounts of phase

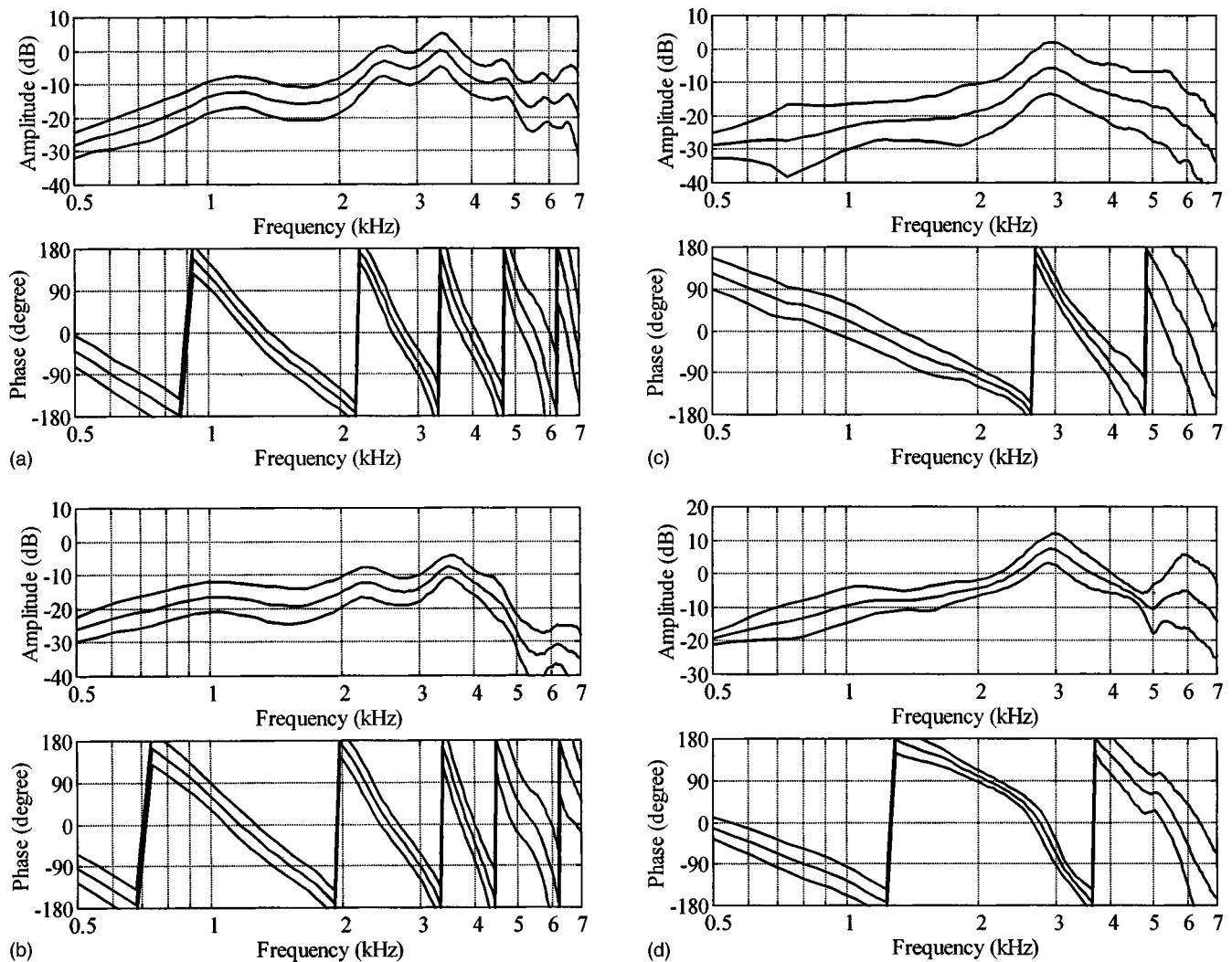


FIG. 17. Transfer function of loop. Phase adjusted to ignore the delay introduced by the AD and DA converter. Mean and mean  $\pm$  s.d. of magnitude and phase. (a) BTE A. (b) BTE B. (c) ITE. (d) ITEC.

shift than the AFE, indicating that the hearing aids give some group delay. The loop response of both BTE A and BTE B give a phase shift of about  $1440^\circ$  and four zero crossings in the frequency region from 500 Hz to 6 kHz. The ITE and the ITEC gave a phase shift of about  $840^\circ$  and  $760^\circ$ , respectively. The phase responses of the ITE and ITEC at low frequencies are about the same except for a constant difference of  $180^\circ$  that could be caused by a shift of polarity of the receiver. The distances between the zero crossings are about 1.4, 2.4, and 2.6 kHz for the BTEs, the ITE, and the ITEC, respectively. This corresponds to a group delay of the loop of about 0.7, 0.4, and 0.4 ms for the BTEs, the ITE, and the ITEC, respectively. The main reason for the difference is the difference in distance of the acoustic path from receiver to microphone via the ear canal. The gain margin is likely to be smaller on a system with a large number of zero crossings since it is the minimum difference between unity gain and the loop gain at frequencies with a phase of  $n360^\circ$ . A larger slope of the phase (larger number of zero crossings) will also increase the impact of the feedback on the closed loop response, with a larger number of peaks and notches at small gain margins.

Digital signal processing in hearing aids will, in most cases, introduce additional group delay. The delay can originate from the AD and DA converters, as well as the signal processing algorithm. All algorithms that are based on the discrete Fourier transform will cause substantial delay due to the inherent block processing. A block size of 16 samples at a sampling frequency of 20 kHz will introduce a 1.6-ms group delay. This will increase the group delay, and the number of zero crossings, by a factor of 5 for an ITE or ITEC. Digital signal processing can, however, be used to reduce the group delay, and thus the number of zero crossings, of the loop response by canceling the part of the loop response that is a minimum phase system (Wang and Harjani, 1993).

The dip found in the AFE that was caused by the probe placement was, in many cases, accompanied by a dip in the loop response. The dip in the loop response could not, however, be caused by the placement of the probe, since the probe microphone signal was not used in the calculations of the loop response. The load impedance at the earmould tip is the likely cause of the dip in the loop response. The load impedance will have a dip at the same frequency as the reso-

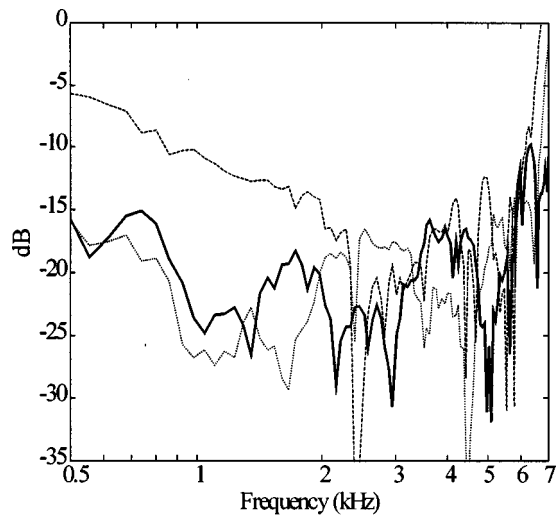


FIG. 18. Mean magnitude of the error function relative to the measured AFE: BTE A (solid line), ITE (dotted line), and ITEC (dashed line).

nance of the transfer function. The dip in the load impedance at the earmould tip will cause a dip in the transfer function from receiver to earmould tip. The dips in the AFE and loop response will thus have the same frequency.

### I. Model of the AFE

Continuous time models of the individual AFE were identified. These models were manually fitted to the measured AFE. The Laplace transform of the model that gave the best fit is given in Eq. (7), which has two complex conjugate pole pairs, a complex conjugate zero pair, a single zero, a delay defined by  $T$ , and a gain factor  $k$ :

$$AFE(s) = ke^{-sT} \frac{(s-z_1)(s-z_2)(s-z_2)}{(s-p_1)(s-p_1)(s-p_2)(s-p_2)}. \quad (7)$$

The single zero and a pole pair were used to obtain the sloping characteristic around 4 kHz. The second pole pair was used to obtain a peak at higher frequencies. These poles were placed closer to the imaginary axis to give a higher Q-value of the peak. The complex zero pair gave the dip found in the AFE. The delay in the model was used since the propagation of sound from the ear canal to the hearing aid microphone will introduce a delay. Figure 18 shows the mean magnitude of the error function, defined as the difference between the model and the measured AFE. The mean values over frequencies of the results of Fig. 18 were  $-19.9$ ,  $-15.7$ , and  $-19.8$  dB for BTE A, ITE, and ITEC, respectively. The poorer fit with ITE was primarily found at frequencies less than 2 kHz and greater than 6 kHz. The data of Fig. 18 shows that the model gave a reasonable fit to the measured data.

### V. SUMMARY

The feedback in hearing aids has been analyzed by introducing the AFE. The complex transfer function of the AFE has been identified for BTE, ITE, and ITEC hearing aids on both human subjects and KEMAR (BTE and ITE). The AFE of BTE, ITE, and ITEC on human subjects had a

mean attenuation at 2 kHz of 60, 51, and 46 dB, respectively. The BTE and ITE used a 2.1-mm vent, while the ITEC used a 3.1-mm vent. A general observation was that the AFE showed less attenuation at high frequencies than at low. The mean estimated transfer function of the feedback via vent and leak was found to be within 1 dB of the mean measured AFE in the frequency range from 1 to 4.5 kHz.

Data measured with ITE on KEMAR agreed reasonably well with data from human subjects. The BTE on KEMAR showed less attenuation than on human subjects, while the phase showed good agreement. The difference in attenuation was greatest at low frequencies.

The AFE with ITEC showed the smallest variations between subjects. This can be explained by the larger vent used in the ITECs.

The placement of the probe influenced the results.

The mean maximum REAG with the reference point under the ear was calculated and at 2 kHz was 68, 59, and 51 dB for the BTE with 2.1-mm vent, ITE with 2.1-mm vent, and ITEC with 3.1-mm vent, respectively.

The AFE could be modeled by a fourth-order continuous time filter together with a delay.

### ACKNOWLEDGMENTS

The financial support from the Swedish Research Council for Engineering Sciences is gratefully acknowledged. Many thanks to Erica Billermark at Linköping University Hospital for performing measurements and to Mette Boel Pedersen, Peter Lundh, Kim Vejlbj Hansen, and Leif Hansen at Oticon A/S for advice and sharing knowledge and for production and modifications of hearing aids.

- Angaard Johansen, P. (1975a). "An evaluation of the acoustic feedback damping for behind the ear hearing aids," Research Laboratory for Technical Audiology, Odense, Denmark.
- Angaard Johansen, P. (1975b). "Magnitude of the acoustical feedback as a function of leakage in earmoulds," *Scand. Audiol. Suppl.* **5**, 271–279.
- Cox, R. M. (1982). "Combined effects of earmould vents and suboscillatory feedback on hearing aid frequency response," *Ear Hear.* **3**, 12–17.
- Dyrlund, O. (1989). "Acoustical feedback associated with the use of post aural hearing aids for profoundly deaf children," *Scand. Audiol.* **18**, 237–241.
- Dyrlund, O., and Lundh, P. (1990). "Gain and feedback problems when fitting behind-the-ear hearing aids to profoundly hearing impaired children," *Scand. Audiol.* **19**, 85–95.
- Egolf, D. P., Howell, H. C., Weaver, K. A., and Barker, D. S. (1985). "The hearing aid feedback path: Mathematical simulations and experimental verification," *J. Acoust. Soc. Am.* **78**, 1578–1587.
- Erickson, F. N., and Van Tassel, D. J. (1991). "Maximum Real-Ear gain of in-the-ear hearing aids," *J. Speech Hear. Res.* **34**, 351–359.
- Fikret-Pasa, S., and Revit, L. J. (1992). "Individualized correction factors in the preselection of hearing aids," *J. Speech Hear. Res.* **35**, 384–400.
- Flack, L., White, R., Tweed, J., Gregory, D. W., and Qureshi, M. Y. (1995). "An investigation into attenuation by earmould tubing," *Br. J. Audiol.* **29**, 237–245.
- Gatehouse, S. (1989). "Limitations on insert gains with vented earmoulds imposed by oscillatory feedback," *Br. J. Audiol.* **23**, 133–136.
- Grover, B. C., and Martin, M. C. (1974). "On the practical gain limit for post-aural hearing aids," *Br. J. Audiol.* **8**, 121–124.
- IEC (1981). IEC 711, "Occluded ear simulator for the measurement of earphones coupled to the ear by ear inserts" (International Electrotechnical Commission, Geneva).

- ISO (1997). ISO/WD11904-1, "Acoustics—Determination of noise immisions from sound sources placed close to the ears—Part 1: Technique using microphones in real ears (MIRE-technique)" (International Organization of Standardization, Geneva).
- Kates, J. M. (1988). "A computer simulation of hearing aid response and the effects of ear canal size," *J. Acoust. Soc. Am.* **83**, 1952–1963.
- Kuk, F. K. (1994). "Maximum Usable Real-Ear Insertion gain with ten Earmould Designs," *J. Am. Acad. Audiol.* **5**, 44–51.
- Ljung, L., and Glad, T. (1994). *Modeling of Dynamic Systems* (Prentice-Hall, Englewood Cliffs, NJ).
- Lundh, P. (1982). "Acoustical feedback problems associated with hearing aids," Oticon internal report No 19-8-6, Oticon, Copenhagen, Denmark.
- Lybarger, S. F. (1975). "Sound leakage from vented earmolds," *Scand. Audiol. Suppl.* **5**, 260–270.
- Oliviera, R. J. (1997). "The active ear canal," *J. Am. Acad. Audiol.* **8**, 401–410.
- Sesterhenn, G., Arold, R., and Boppert, R. (1998). "Wie zuverlässig sind Sondermessungen," *Zeitschrift für Audiologie* **37**(2), 75–83.
- Wang, R., and Harjani, R. (1993). "Suppression of acoustic oscillation in hearing aids using minimum phase techniques," *Proc. IEEE Int. Symp. Circ. Syst.* 818–821.

# Missing-data model of vowel identification<sup>a)</sup>

Alain de Cheveigné<sup>b)</sup>

Laboratoire de Linguistique Formelle, CNRS/Université Paris 7, 2 place Jussieu, case 7003, 75251, Paris, France and ATR Human Information Processing Research Laboratories, 2-2 Hikaridai, Seika-cho, Soraku-gun, Kyoto 619-0288, Japan

Hideki Kawahara

Media Design Informatics Group, Design Information Science Department, Faculty of Systems Engineering, CREST/Wakayama University, Sakaedani, Wakayama 640-8510, Japan and ATR Human Information Processing Research Laboratories, 2-2 Hikaridai, Seika-cho, Soraku-gun, Kyoto 619-0288, Japan

(Received 9 October 1998; revised 15 January 1999; accepted 5 February 1999)

Vowel identity correlates well with the shape of the transfer function of the vocal tract, in particular the position of the first two or three formant peaks. However, in voiced speech the transfer function is *sampled* at multiples of the fundamental frequency ( $F_0$ ), and the short-term spectrum contains peaks at those frequencies, rather than at formants. It is not clear how the auditory system estimates the original spectral envelope from the vowel waveform. Cochlear excitation patterns, for example, resolve harmonics in the low-frequency region and their shape varies strongly with  $F_0$ . The problem cannot be cured by smoothing: lag-domain components of the spectral envelope are aliased and cause  $F_0$ -dependent distortion. The problem is severe at high  $F_0$ 's where the spectral envelope is severely undersampled. This paper treats vowel identification as a process of pattern recognition with *missing data*. Matching is restricted to available data, and missing data are ignored using an  $F_0$ -dependent weighting function that emphasizes regions near harmonics. The model is presented in two versions: a frequency-domain version based on short-term spectra, or tonotopic excitation patterns, and a time-domain version based on autocorrelation functions. It accounts for the relative  $F_0$ -independency observed in vowel identification. © 1999 Acoustical Society of America.

[S0001-4966(99)00906-6]

PACS numbers: 43.71.An, 43.72.Ar, 43.66.Ba [JMH]

## INTRODUCTION

In voiced speech, the vocal tract resonator is excited with a regular train of glottal pulses, due to opening and closing of the glottis at a rate equal to the fundamental frequency ( $F_0$ ). According to the acoustic theory of speech production (Fant, 1970), speech is the result of filtering this train by the resonator. The shape of glottal pulses depends on the mode of phonation and characteristics of the speaker. This shape can be included mathematically within the vocal tract impulse response, and the vocal tract is then seen as excited by a train of pulses, infinitely narrow in time. In the following, the term “vocal tract” implies this resonator-cum-glottal-pulse-shape equivalent. In the frequency domain, if  $F_0$  is constant, the spectrum of the excitation is a series of equal-amplitude peaks at multiples of  $F_0$ . The speech spectrum therefore also consists of peaks with amplitudes determined by the amplitude of the transfer function. In other words, the transfer function of the vocal tract is *sampled* at multiples of  $F_0$  (Fig. 1).

The timbre and identity of a sustained vowel are determined by the shape of the vocal tract transfer function, particularly the positions of the first two or three formants. However, the listener has no access to this shape, but only to the waveform or auditory representations derived from it.

Figure 2 shows the rms output of a bank of gammatone filters in response to the Japanese vowel /a/. This pattern can be taken as approximating the activity evoked by the vowel over a tonotopic dimension within the auditory system (excitation pattern). At  $F_0 = 50$  Hz (top), the pattern is smooth with two clear peaks corresponding to the formants. At  $F_0 = 200$  Hz (middle), these peaks are still present, but slightly shifted and there are many other smaller peaks. At  $F_0 = 216$  Hz (bottom), the peaks at  $F_1$  and  $F_2$  of /a/ are no more prominent than other peaks, and it is not clear what aspect of the excitation pattern might be used to characterize the vowel. Upon listening, the vowel's timbre or identity do not change strikingly between 200 and 216 Hz.

Despite certain interactions (see the Discussion), vowel quality is on the whole remarkably independent of  $F_0$ . One could make the hypothesis that the auditory system, by some process that is yet to be understood, forms an internal representation that is invariant over variations of  $F_0$ . For example, summation of activity of converging nerve fibers might smooth out the ripples visible in Fig. 2. Indeed, the figure of 3.5 bark has been proposed as an appropriate integration range for vowel spectrum matching. In this paper, we argue against smoothing for several reasons: (a) Undersampling implies a genuine *loss of data*: information about the transfer function at frequencies other than multiples of the  $F_0$  is lost, and smoothing cannot retrieve it. (b) Smoothing and interpolation attempt to guess missing samples based on an *a priori* model of what they should look like, and this may

<sup>a)</sup>Part of this work was presented in an ATR technical report (de Cheveigné and Kawahara, 1998).

<sup>b)</sup>Electronic mail: cheveign@ircam.fr



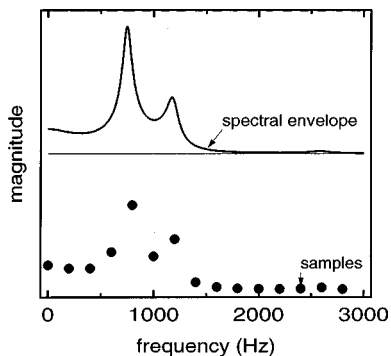


FIG. 1. Line: spectral envelope of vowel /a/. Dots: spectral envelope sampled at intervals of  $F_0=200$  Hz.

be misleading. (c) In particular, aliasing due to undersampling may produce  $F_0$ -dependent distortion that interferes with the recognition process. (d) Aliasing is avoided if the auditory system puts a nonuniform weight on the unsmoothed representations that it derives from the waveform. This weighting function requires an estimate of the  $F_0$ .

### A. Sampling of the spectral envelope

The shape of a spectral envelope can be described in the Fourier domain along a dimension of *lag* (time interval or inverse frequency). This dimension is also known as *spatial frequency*, or *quefrecy* in the context of cepstrum analysis (Rabiner and Schafer, 1978). Components at short lags represent gross features of the spectral envelope that vary gradually along the frequency axis, whereas components with larger lag values represent finer details.<sup>1</sup> We use the term *sampling lag* to designate the inverse  $T_0$  of the spacing  $F_0$  between samples of the spectral envelope. From the sampling theorem, we know that the envelope is adequately represented by the sample points if its shape contains no components beyond half the sampling lag. This limit,  $T_0/2$ , will be referred to as *Nyquist lag*.

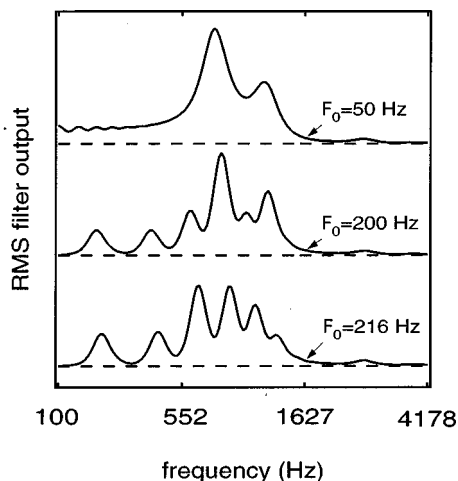


FIG. 2. Magnitude of output of gammatone filter bank as a function of channel frequency (excitation pattern). The filterbank had 150 channels uniformly spaced on a scale of equivalent rectangular bandwidth (ERB) from 100 to 4178 Hz. Each curve is for a different  $F_0$ . Note the peaks at harmonics for the higher two  $F_0$ 's, and the lack of unambiguous evidence for  $F_1$  and  $F_2$  at  $F_0=216$  Hz.

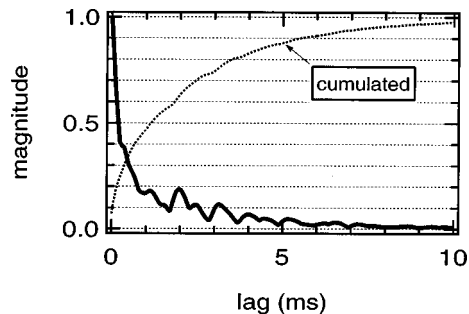


FIG. 3. Full line: magnitude of the lag-domain components of the vowel spectral envelope as a function of lag, averaged over the five Japanese vowels (/a/, /e/, /i/, /o/, /u/). Dotted line: same, cumulated over lags (proportion of lag-domain components with lags smaller than a given lag). The zero-lag value (dc component) is not included in the cumulated-distribution calculation, nor in the plots.

Figure 3 shows the distribution of spectral envelope lag-domain components, averaged over synthetic envelopes of the five Japanese vowels /a/, /e/, /i/, /o/, /u/.<sup>2</sup> The figure was obtained by calculating the magnitude of the Fourier transform of the envelope of each vowel, and averaging over vowels. The low-lag region dominates, but components are also present at larger lags. The cumulated distribution is plotted as a dotted line (the zero-lag point was excluded from both distributions because it represents a dc offset rather than the shape). On average, about 10% of lag-domain components lie beyond 5 ms which is the Nyquist lag for  $F_0=100$  Hz. These are not adequately represented in the sampled spectrum when  $F_0=100$  Hz. At 200 Hz (2.5 ms), the proportion is about 25%, and at 300 Hz (1.67 ms), about 40%.

Consider the short-term spectrum of a synthetic vowel /a/ at  $F_0=100$  Hz [Fig. 4(a), full line at top]. It was calculated by taking the Fourier transform of a 100-ms portion of the waveform. The sampling rate was 40 960 Hz. The short-term spectrum resembles the spectral envelope (dotted line) sampled along the frequency axis at intervals of 100 Hz. If the spectral envelope contained no lag-domain components beyond the Nyquist lag, its shape could be accurately reconstructed by smoothing the short-term spectrum (filtering in the lag domain) to remove components beyond that lag. The result of such smoothing is shown as the lower line in Fig. 4(a). Smoothing was performed by taking the Fourier transform of the short-term magnitude spectrum, setting values beyond the Nyquist lag to zero, then applying the inverse transform. The slight difference between this and the original spectral envelope (dotted line at top) implies that the original did in fact contain components beyond the Nyquist lag. The difference is small, suggesting that little was lost by sampling the envelope at 100-Hz intervals.

At  $F_0=200$  Hz [Fig. 4(b)], the peaks in the smoothed spectrum (bottom) are wider and there is a strong ripple with a period inverse of the Nyquist lag (2.5 ms). At  $F_0=300$  Hz [Fig. 4(c)], the reconstructed envelope (full line at bottom) is severely distorted, indicating that spectral envelope components beyond the Nyquist lag (1.67 ms) were necessary to describe the original shape. The significance of the dotted line is discussed in the next paragraph.

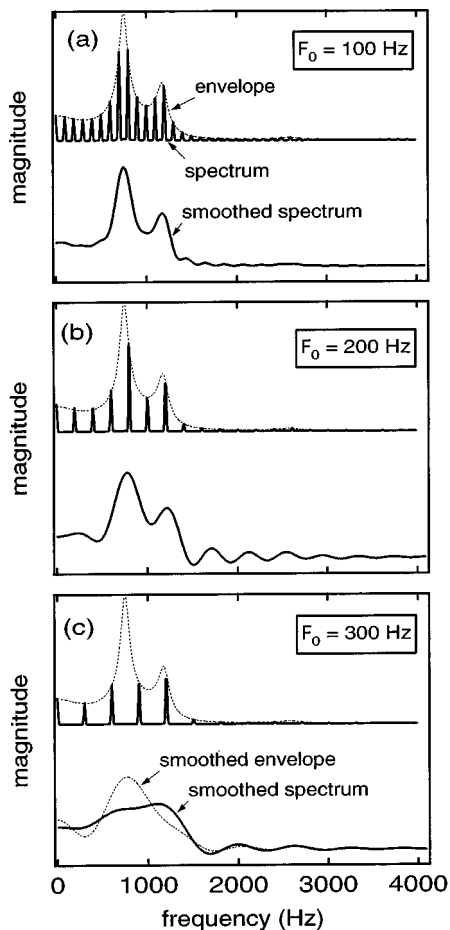


FIG. 4. (a) Top: envelope (dotted line) and short-term magnitude spectrum of /a/ at  $F_0=100$  Hz (full line). Bottom: smoothed short-term spectrum. Smoothing was performed by taking the Fourier transform of the magnitude spectrum, setting it to zero for lags larger than the Nyquist lag  $T_0=1/2F_0$  (5 ms), and taking the inverse Fourier transform. The smoothed spectrum consists entirely of components below the Nyquist lag. (b) Same, at  $F_0=200$  Hz. Note the ripples with a period corresponding to the inverse of the Nyquist lag (2.5 ms), that indicate that aliasing is taking place. (c) Same, at  $F_0=300$  Hz, with the addition of the smoothed spectral envelope (dotted curve at bottom). The spectral envelope was smoothed by removal of lag components beyond the Nyquist lag. The difference between smoothed envelope and smoothed spectrum is the result of aliasing.

## B. Smoothing considered harmful

Undersampling produces data that are incomplete, but correct in the sense that each sample corresponds to a value of the spectral envelope. This is no longer the case if the sampled spectrum is smoothed. Consider the smoothed spectra of Fig. 4. They differ from the original spectral envelope partly because of the absence of high-lag components, and partly because components beyond the Nyquist lag are *aliased*; that is, reflected with respect to the Nyquist lag and reintroduced into the smoothed envelope. The respective contribution of both factors is illustrated in Fig. 4(c). The lower dotted line represents the Nyquist-smoothed spectral envelope, that differs from the spectral envelope (upper dotted line) merely because it lacks components beyond the Nyquist lag. The lower full line (Nyquist-smoothed short-term spectrum) differs from it by the additional factor of aliasing. The effects of aliasing are highly  $F_0$ -dependent and may interfere with the identification process, particularly at

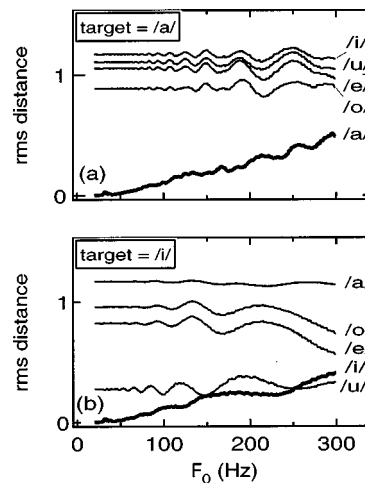


FIG. 5. Simple vowel identification model. (a) Distance between reference templates for vowels /a/, /e/, /i/, /o/, and /u/, and the smoothed short-term spectrum calculated from a target /a/ waveform, as a function of  $F_0$ . Smoothing was performed by removing all components beyond the Nyquist lag ( $1/2F_0$ ). (b) Same as (a), for a target /i/ waveform.

high  $F_0$ 's. In other words, smoothing replaces data that are incomplete, but correct, with data that are complete but incorrect.

To illustrate the difficulties that this causes for identification, a simple vowel-identification model was implemented. The reference template for each vowel was the spectral envelope of that vowel. Target vowels were synthesized at  $F_0$ 's ranging from 20 to 300 Hz in 1-Hz steps. Their short-term spectra were smoothed by removing components above the Nyquist lag (the value of which depended on the target's  $F_0$ ), and compared to each template by scaling target and template to the same rms value (1.0) and calculating their rms difference. Target-template distances for target =/a/ are plotted in Fig. 5(a). The distance from competing templates /e/, /i/, /o/, and /u/ remains relatively large, despite some fluctuations. The distance from the "correct" template /a/ is smaller, but it increases steadily with  $F_0$ , indicating that the estimated envelope is less and less faithful to the original.

A similar plot for the target /i/ is shown in Fig. 5(b). Here, at high  $F_0$ 's the estimated envelope is actually closer to the incorrect /u/ template than to the correct /i/ template. The model thus fails. It should be stressed that the task is comparatively easy: the vowel set is small, its members well separated in  $F_1-F_2$  space, and there is no variability. In more realistic conditions, the vowel space might be more densely populated and there would be many sources of variability, making the task even more difficult. This model is simple and allows ample room for improvement, but it serves to deliver a message that should be clear:  $F_0$ -dependent effects of spectral sampling can be severe, particularly at high  $F_0$ , and they cannot be eliminated by smoothing.

## I. MISSING-DATA VOWEL IDENTIFICATION MODEL

The model acknowledges that important spectral information was lost due to undersampling. Instead of interpolating or otherwise trying to estimate the missing data, pattern

matching proceeds using the available data only, giving zero weight to missing data in the pattern-matching process. A similar idea underlies “missing data” or “missing feature” techniques that have been proposed recently in speech recognition to cope with deleted spectro-temporal features (Cooke *et al.*, 1994, 1997; Lippmann, 1997; Morris *et al.*, 1998). Two versions of the model are proposed: one works in the frequency domain, and the other in the time domain. Both require an estimate of the  $F_0$  of the vowel.

### A. Frequency-domain version

The frequency-domain model is straightforward. Spectral templates are assumed to be available for all vowel classes. The following steps occur when a vowel is recognized: (a) its short-term spectrum is estimated, (b) its  $F_0$  is estimated, (c) based on  $F_0$ , a spectral weighting function is calculated that emphasizes regions near multiples of  $F_0$ , and (d) the short-term spectrum is compared to all templates using the weighting function. The template that yields the smallest distance determines the vowel that is “recognized.” Templates are defined over the whole spectrum, but comparison is restricted to certain frequencies depending on the  $F_0$ .

The weighting function  $W(f)$  and spectral distance  $D(T, T_i)$  from target  $T$  to template  $T_i$  might be defined as

$$W(f) = \sum_{n=0}^{\infty} \delta(f - n\hat{F}_0), \quad (1)$$

$$D(T, T_i) = \int (T(f) - T_i(f))^2 W(f) df, \quad (2)$$

where  $\delta()$  is the Dirac delta function,  $\hat{F}_0$  is the fundamental frequency estimate,  $T(f)$  is the short-term spectrum measured from the waveform, and  $T_i(f)$  is the spectral envelope of the  $i$ th vowel. The infinitely narrow peaks of  $W(f)$  in Eq. (1) are satisfactory in theory. In practice, the peaks might be made to widen gradually with  $f$  to accommodate inevitable inaccuracy in  $F_0$  estimation. With a square shape and relative width of 3%, such a weighting function is equivalent to the harmonic sieve of Duifhuis *et al.* (1982), that has been proposed as a mechanism to select information in the context of pitch perception (Moore *et al.*, 1984, 1985; Darwin *et al.*, 1992) and concurrent vowel identification (Scheffers, 1983).

The model requires an  $F_0$  estimate. This is impossible to obtain for stimuli that are too short, whispered, or otherwise nonstationary. However, in those cases no special processing is called for: the short-term spectrum is unbiased, and the sampling operation is not necessary. A complete formulation of the model must explain how it switches from one mode to the other. For example, the weighting function might be uniform by default, and replaced by a gradually sharpening spectral comb as reliable  $F_0$  information is obtained. Sharpness of the spectral comb could be under the control of a “periodicity measure” (many  $F_0$  estimation methods produce such a measure as a by-product). When periodicity is good,  $F_0$ -estimation errors are unlikely; the model is therefore robust.

As an illustration, a simple missing-data vowel identification model was built, similar to that of the Introduction, subsection B but with smoothing replaced by nonuniform

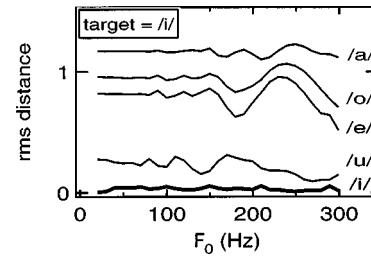


FIG. 6. Simple missing-data vowel identification model. Weighted distance between reference templates for vowels /a/, /e/, /i/, /o/, and /u/, and the unsmoothed short-term spectrum estimated from a target /i/ waveform, as a function of  $F_0$ . The weighting function was 1.0 at multiples of  $F_0$ , and 0 elsewhere. The nonzero values of the distance to the /i/ template are due to sampling error in the spectrum calculations.

weighting. The weighting function was 1.0 at multiples of  $F_0$ , and 0.0 elsewhere. Figure 6 shows the weighted distance of a synthetic /i/ target vowel to each of the templates, as a function of  $F_0$ . The distance to the correct template /i/ is not quite zero, because of limited sampling resolution in the implementation, but it remains smaller than the distances to incorrect templates. The vowel /i/ is identified correctly at all  $F_0$ 's [compare with Fig. 5(b)].

Physiologically, one can imagine that short-term spectral estimation occurs in the cochlea, and that the harmonic sieve is applied along a tonotopic dimension at some point in the auditory system, based on an  $F_0$  estimate, itself possibly derived from tonotopic information. The main difficulty, apart from the issue of frequency resolution, is to imagine how the variable-pitch harmonic sieve is constructed based on the  $F_0$  estimate, and how it is deployed across frequency channels. In the following section, we consider an alternative model based on autocorrelation that might be implemented physiologically using time-domain processing in the auditory system.

### B. Autocorrelation version

The Fourier-domain reasoning of the Introduction, subsection A that was applied to the *magnitude* of the vocal-tract transfer function can be applied equally well to its *squared magnitude*. The Fourier transform of the squared magnitude is the autocorrelation of the vocal-tract impulse response,  $ACF_{\text{tract}}$ , plotted as a thin line in Fig. 7(a). When a vowel is produced with a constant  $F_0$ , the squared magnitude vocal-tract transfer function is sampled at multiples of  $F_0$ . The sampling theorem tells us that the samples describe uniquely a function bandlimited to lags smaller than  $T_0/2$ . In other words, the information about the vocal tract available in the samples is also represented in the  $\tau < T_0/2$  portion of the autocorrelation function  $ACF_{\text{tract}}$  [Fig. 7(a), thick line].

In the vowel identification model of Sec. IA, magnitude spectra could be replaced by squared-magnitude spectra. Parseval's theorem tells us that the Euclidean distance between such spectra is the same as that between the corresponding autocorrelation functions. One can thus in turn replace squared-magnitude spectra by autocorrelation functions in the vowel identification model. The spectral

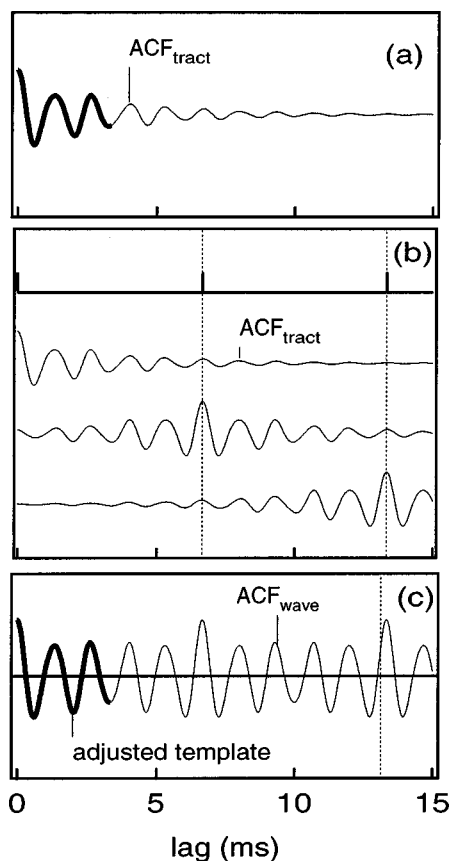


FIG. 7. (a) Autocorrelation of vocal tract transfer function ( $ACF_{tract}$ ). (b) Illustration of the convolution process by which  $ACF_{wave}$  is derived from  $ACF_{tract}$  in the case where  $F_0 = 150$  Hz. Copies of  $ACF_{tract}$  are shifted and added to obtain  $ACF_{wave}$  [thin line in (c)]. The vertical dotted lines indicate multiples of the period, 6.67 ms. The thick line in (c) is an adjusted template (see the text).

weighting function is replaced by a lag-domain weighting function restricted to lags  $\tau < T_0/2$ .

A difficulty remains. The autocorrelation function of the impulse response of the vocal tract ( $ACF_{tract}$ ) is not directly observable. Observable is the autocorrelation of the waveform ( $ACF_{wave}$ ), that differs from  $ACF_{tract}$  but is related to it by the following relation:

$$ACF_{wave}(\tau) = ACF_{tract}(\tau) \circ \sum_{k=-\infty}^{\infty} \delta(\tau - kT), \quad (3)$$

where  $\circ$  represents convolution. As illustrated in [Fig. 7(b)], copies of ( $ACF_{tract}$ ) are shifted to multiples of  $T_0$ , and added up to obtain  $ACF_{wave}$ . Because of overlap between the shifted functions,  $ACF_{wave}$  differs from  $ACF_{tract}$  in the important region  $\tau < T_0/2$ . The difference depends on  $F_0$ .

For this reason,  $ACF_{wave}$  cannot make a perfect match to the templates, even if it is restricted to  $\tau < T_0/2$ . However, if  $F_0$  is known, it is possible to *adjust* the templates to obtain a perfect match. This is done by adding up appropriately shifted versions of the templates, exactly as in the convolution illustrated in Fig. 7(b). In this way, an accurate match is obtained between the correct vowel template and the observed  $ACF_{wave}$  [Fig. 7(c)]. We can thus formulate an autocorrelation version of the missing-data vowel perception model. The following steps occur when a vowel is recog-

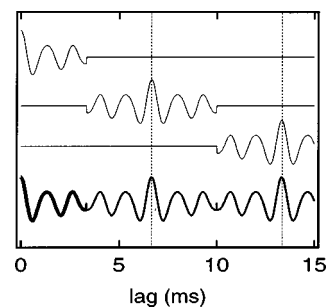


FIG. 8. Illustration of the hypothetical case of a vocal tract with a transfer function band-limited to lags smaller than 3.33 ms.  $ACF_{wave}$  is derived from  $ACF_{tract}$  by convolution, but the  $\tau < 3.33$  ms portion is unaffected by the convolution and remains equal to  $ACF_{tract}$ . Template adjustment would be unnecessary in this hypothetical case.

nized: (a)  $ACF_{wave}$  is estimated from the waveform, (b)  $F_0$  is estimated, (c) based on  $F_0$ , the set of  $ACF_{tract}$  templates are adjusted, and (d) each one is compared to  $ACF_{wave}$  over the  $\tau < T_0/2$  range of lags. The closest match indicates the “recognized” vowel.

The  $F_0$  estimate is involved in two places: template adjustment and determination of the range of lags to be matched. The adjustment step would be unnecessary in the hypothesis that  $ACF_{tract}$  is limited to  $T_0/2$ , as illustrated in the top of Fig. 8. If  $ACF_{wave}$  is zero beyond  $T_0/2$ ,  $ACF_{wave}$  and  $ACF_{tract}$  are equal for  $\tau < T_0/2$  (thick line at the bottom of Fig. 8), and adjustment is unnecessary. Omitting the adjustment stage is thus equivalent to putting faith in the assumption that the squared-magnitude spectrum is bandlimited to  $\tau < T_0/2$  in the lag domain. This assumption is all the more incorrect as  $F_0$  is high.

In the squared-magnitude spectrum, high-amplitude parts of the spectrum are emphasized at the expense of others [Fig. 9(a)]. The first formant is well represented, and this accounts for the ripple that dominates  $ACF_{tract}$ . The second formant is less well represented, and higher formants hardly at all. The magnitude spectrum of the Introduction, subsection B is a slightly more balanced representation [Fig. 9(b)]. The *log magnitude spectrum* represents both peaks and valleys in equal detail, whatever their amplitude [Fig. 9(c)], and its inverse Fourier transform, the *cepstrum*, is widely used in speech processing. The success of the cepstrum in speech-

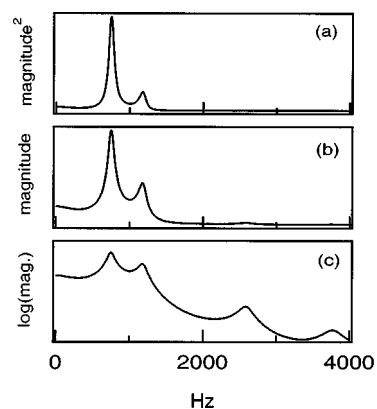


FIG. 9. (a) Squared-magnitude transfer function of /a/. (b) Magnitude transfer function of /a/. (c) Log-magnitude transfer function of /a/.

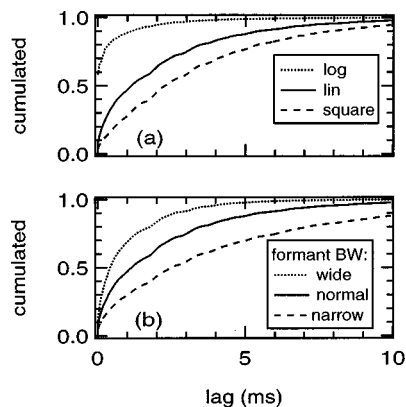


FIG. 10. (a) Cumulated magnitude of lag-domain components of the spectral envelope, for a linear (full line), log (dotted line), or squared (dashed line) spectral envelope. The smoother log spectrum has fewer high-lag components, whereas the sharper squared spectrum has more. (b) Cumulated magnitude of lag-domain components of the spectral envelope for vowels with formant bandwidths that are standard (full line), twice standard (dotted line), or half standard (dashed line). Narrow formants imply more high-lag components; wide formants imply less.

processing applications suggests that the log magnitude spectrum (and cepstrum) are more effective substrates for pattern matching than the squared-magnitude spectrum (and autocorrelation). The excessive emphasis on  $F_1$  of the autocorrelation function is alleviated in the model of Sec. IC, where autocorrelation functions are calculated within channels of a basilar-membrane/hair-cell model.

As illustrated in Fig. 10(a), the lag-domain composition of the spectral envelope differs from that of its squared and log transforms. The squared version is richer in high-lag components, whereas the log version is poorer. As illustrated in Fig. 10(b), the lag-domain composition also depends on formant bandwidth. Narrow formants require more high-lag components, wide formants less. The severity of the aliasing effects described in this paper thus depend on other factors in addition to  $F_0$ . The relative lack of high-lag components of the log spectrum is another possible reason for its success (and that of the cepstrum) in speech processing. The perceptual severity of aliasing would depend on the representation used by the auditory system. Autocorrelation functions are *a priori* highly sensitive, but this sensitivity might be alleviated by compressive mechanisms in a physiological representation (next section).

Both the autocorrelation model and the spectral model of Sec. IA are perfectly accurate. In a task lacking variability (other than due to  $F_0$ ), both would perform perfectly at any  $F_0$ . However, it is clear from Fig. 7 that identification at higher  $F_0$  has less information to go on, and is likely to degrade as soon as variability is introduced into the task.

### C. Multichannel autocorrelation version

In Sec. IA we suggested that the spectral version of the model might be implemented by frequency-domain processing within the auditory system, based on a tonotopic representation. Here, we describe how the autocorrelation version might be implemented by *time-domain* processing within the auditory system, based on the temporal structure of nerve-fiber discharge patterns.

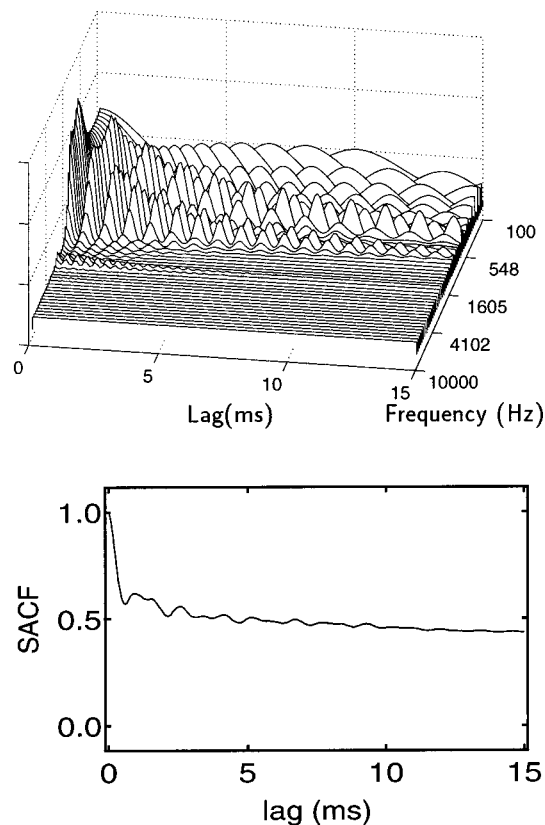


FIG. 11. Top: Array of autocorrelation functions of auditory-nerve fiber discharge probability indexed by channel frequency, in response to a single impulse of the vowel /a/. Probabilities were produced by a model of peripheral filtering and hair-cell transduction, with 50 channels uniformly spaced between 100 and 10 000 Hz on a scale of equivalent rectangular bandwidth (ERB). Bottom: summary autocorrelation function (SACF).

Autocorrelation of nerve-fiber discharge patterns has been suggested as a basis for pitch perception (Licklider, 1951; Meddis and Hewitt, 1991a, b; Cariani and Delgutte, 1997a, b), and cross correlation is an accepted mechanism for localization (Jeffress, 1948; Yin *et al.*, 1987). In the pitch model of Meddis and Hewitt (1991a, b), autocorrelation functions (ACF) of auditory-nerve discharge probability were calculated within each channel of a model of basilar-membrane filtering and hair-cell transduction, and ACFs for all channels were added up to form a summary autocorrelation function (SACF). The pitch was derived from the position of the highest peak in the SACF. Many aspects of pitch phenomena are well accounted for by that model or others that are closely related (de Cheveigné, 1998a). The SACF was also proposed as a substrate for *vowel identification* in the concurrent vowel-identification model of Meddis and Hewitt (1992). In their model, vowels were identified by matching the “low-lag” portion of the SACF ( $\tau < 4$  ms), a scheme that was also used with success by de Cheveigné (1997).

Figure 11 (top) shows an array of autocorrelation functions calculated from the instantaneous discharge probability functions produced by a model of peripheral filtering and hair-cell transduction (Slaney, 1993). The model had 40 channels uniformly distributed on a scale of equivalent rectangular bandwidth (ERB) between 100 and 10 000 Hz. The stimulus was a single-impulse response of the vocal tract

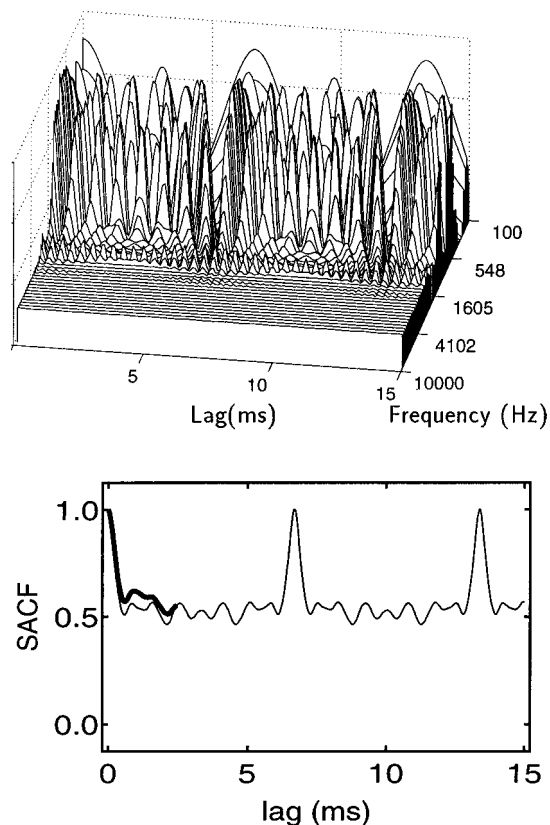


FIG. 12. Same as Fig. 11, in response to vowel /a/ at  $F_0=150$  Hz.

corresponding to the vowel /a/. The combination of vocal tract and basilar-membrane filter leads to a much longer impulse response than the vocal tract alone, which explains the slow decay of the ACFs in Fig. 11 compared to Fig. 7(b), particularly in channels tuned to low frequencies and/or near a formant. Because of the rectifying property of the hair-cell model, ACFs are never negative.

Figure 11 (bottom) shows the corresponding summary autocorrelation function (normalized by dividing by the value at zero lag). The SACF decays to a relatively high baseline due to the summation of non-negative ACFs for different channels. Compared to the autocorrelation of the waveform [Fig. 7(b)], the SACF lacks the strong ripple at the period of  $F_1$ . This is probably a consequence of the saturating properties of the basilar-membrane model, which equalize contributions of different channels. Compared to waveform autocorrelation, the SACF is affected relatively less by  $F_1$ , and more by other components.

Figure 12 shows the response of the model to the vowel /a/ at  $F_0=150$  Hz. Each channel shows a peak at multiples of  $T_0$ , as does the SACF. Over the interval  $[0-T_0/2]$ , the SACF resembles the SACF obtained in response to the impulse response (thick line), justifying the choice of Meddis and Hewitt (1992) of the low-lag portion of the SACF as a substrate for vowel identification. There are, nevertheless, differences between the two SACFs, as we observed previously for the waveform-based autocorrelation functions. Analysis of these differences is complicated by the presence of the nonlinear hair-cell transduction stage. Because of the nonlinearity, within-channel ACFs in response to the peri-

odic stimulus cannot accurately be calculated as a convolution, as in Sec. I B. However, this does not prevent postulating a model similar to that of Sec. I B, in which templates are tabulated rather than calculated.

It is customary to use the SACF instead of the full ACF array as a basis for pitch (Meddis and Hewitt, 1991a, b) or vowel identification (Meddis and Hewitt, 1992), but this is not mandatory. Matching could just as well be performed on the full ACF array, or better still, on an array of “sub-SACFs” calculated over sub-bands wide enough to avoid gaps between harmonics at high  $F_0$ . One advantage might be better discrimination, as the two-dimensional array is a richer pattern than the one-dimensional SACF. Another is the possibility to factor out spectral tilt and other transmission channel characteristics (thanks to within-channel compression or automatic gain control, as performed by the hair-cell model). A third is that parts of the array may be weighted differently in the spectral domain to handle “missing features,” due to bandlimited noise or filtering (Cooke *et al.*, 1994, 1997; Lippmann, 1997; Morris *et al.*, 1998). The full ACF or sub-SACF array is a representation in which correlates of sources and interference can be weighted differentially in both the lag and the frequency domain. It is thus a flexible starting point for sophisticated models of identification and segregation.

## II. DISCUSSION

The identity of a vowel depends on the shape of its spectral envelope, essentially the position of the first two or three formants. However, when a vowel is produced at a constant pitch, its spectral envelope is sparsely sampled, implying that details of this shape may not be well represented. Spectral representations derived from the waveform (short-term spectra, auditory excitation patterns, etc.) have a harmonic structure that interferes with the determination of the spectral envelope. Harmonics masquerade as formant peaks, and at high  $F_0$ 's the short-term spectrum bears little resemblance to the vowel's spectral envelope. Nevertheless, to a first approximation, the vowel's identity does not depend on the fundamental frequency.  $F_0$  typically varies from 90 to 200 Hz for male speakers, or 150 to 310 Hz for women (Howard, 1991), although much wide ranges have been reported, in particular for children (Fairbanks, 1940; Keating and Buhr, 1978). Sundberg (1982) states that steady-state sung vowels retain a degree of intelligibility up to 520 Hz (or even 1000 Hz in consonant–vowel–consonant (CVC) context). The problem of vowel perception at high  $F_0$  has been recognized by many authors (Carlson *et al.*, 1975; Carré and Lancia, 1975; Traunmüller, 1981, 1990; Bladon, 1982; Ryalls and Lieberman, 1982; Klatt, 1982; Sundberg, 1982; Darwin and Gardner, 1985; Gottfried and Chew, 1986; Assmann and Neary, 1987; Beddor and Hawkins, 1990; Perc, 1991; Assmann, 1991; Hirahara and Kato, 1992; Hirahara, 1993; Rosner and Pickering, 1994; Hirahara *et al.*, 1996).

Two qualifications must be made. The first is that in natural speech vowels are rarely sustained. Harmonics in the spectrum of a nonstationary sound are less sharp, and the bias due to undersampling is less severe. In the limit, a single vowel period has no harmonic structure, yet it is sufficient

for vowel identification (McKeown and Patterson, 1995; Robinson and Patterson, 1995). Summerfield *et al.* (1984) remarked that a long-duration synthetic vowel sounds vowel-like at onset, but may then lose its identity. The identity is partially regained at offset. It has been suggested that modulation of the  $F_0$  (e.g., *vibrato*) improves identification by giving samples of the *derivative* of the spectral envelope at  $F_0$  multiples, in addition to the magnitude. However, Sundberg (1982) reported that *vibrato* did not improve identification of high-pitched sung vowels. Hillenbrand and Gayvert (1993) found that vowels were more intelligible when synthesized with a falling rather than static  $F_0$  contour, but the advantage was small. McAdams and Rodet (1988) suggested that FM-induced envelope tracing might enhance identification of concurrent vowels, but Marin and McAdams (1991) failed to find support for the idea. In natural speech, dynamic cues such as consonant–vowel transitions play an important role in vowel identification (Strange *et al.*, 1976, 1998; Hillenbrand and Neary, 1999). Intelligibility of high- $F_0$  sung vowels is better when they are in CVC context (Sundberg, 1982; Gottfried and Chew, 1986), and in the extreme, there is evidence that a vowel may be identified from the unvoiced portion of a consonant with which it is articulated (Bonneau, 1996). In natural speech, the spectral undersampling problem is thus less severe than it might seem from consideration of monotone sustained vowels.

The second qualification is that, to a second approximation, vowel quality *is* dependent on  $F_0$ . From a production point of view, there are several sources of correlation between spectral envelope and  $F_0$ . Within a population of speakers, women and children tend to have higher-pitched voices and shorter vocal tracts than men (Peterson and Barney, 1952). For a given speaker, a change in  $F_0$  may imply a change in glottal pulse shape. If the glottal shape is mathematically included in the “vocal tract” transfer function (Introduction), the latter would vary with  $F_0$  even if the shape of the vocal tract did not. Variation in larynx height concurrent with  $F_0$  variations may also affect the spectral envelope, and in the case of singing, professional singers are known to intentionally raise the  $F_1$  of a high-pitched vowel to ensure that it does not fall below the  $F_0$ , in which case it would not be adequately excited (Sundberg, 1982). It is also possible that different vowels are systematically given different  $F_0$ 's (“intrinsic  $F_0$ ”).

From a perceptual point of view, numerous experiments have found interactions between  $F_0$  and vowel quality. In an experiment that investigated both musical timbre and vowel quality, Slawson (1967) found that, for an  $F_0$  increase of one octave,  $F_1$  and  $F_2$  should be increased by about 10% to minimize the change in quality. Carlson *et al.* (1975) measured the perceptual  $F_1$  boundary between Swedish /i/ and /e/, and found that it increased from about 300 to 350 Hz for  $F_0$ 's varying from 100 to 160 Hz. Ainsworth (1975) determined that a one-octave increase in  $F_0$  had to be accompanied by a 3.4% increase of  $F_1$  and a 1.2% increase of  $F_2$  (figures that he argued might be underestimated by a factor of 2). Neary (1989) likewise found a 7%–9% shift for  $F_1$  and a 1.5% shift for  $F_2$ . Assmann *et al.* (1982) found that inclusion of an  $F_0$ -based parameter improved discriminant

analysis of sets of acoustic parameters of vowels, and Hirahara and Kato (1992) obtained data points that were better clustered (both for production and for perception) when expressed in a plane of  $(F_1 - F_0) \times (F_2 - F_0)$  bark differences, rather than an  $F_1 \times F_2$  plane. Miller (1989) proposed a vowel perception model in which formant frequencies were normalized by the cubic root of  $F_0$ , but several authors have argued that this implies an exaggerated  $F_0$  dependency (Neary, 1989; Rosner and Pickering, 1994). Other effects are reported by Traunmüller (1981), Fahey *et al.* (1996), Hoemke and Diehl (1994), Kewley-Port (1996), Kewley-Port *et al.* (1996). Overall, effects of  $F_0$  are rather small. However, to the extent that an orderly relation exists between  $F_0$  and vowel quality, and that listeners exploit it, complete insensitivity to  $F_0$  is not necessarily desirable in a vowel perception model.

Intelligibility usually gets worse at high  $F_0$ . Ryalls and Liebermann (1982) found that vowels synthesized with formants appropriate for male and female voices were less intelligible at 250 Hz than at 185 Hz (female), 135 Hz (male), or 100 Hz (both). With synthetic vowels, Sundberg *et al.* (1982) found a decrease in intelligibility between 260 and 700 Hz. With natural vowels sung by a soprano, Benolken and Swanson (1990) also found a steady decrease from 262 to 1047 Hz. Sundberg (1982) notes that intelligibility is better when vowels are pronounced in a CVC context. Gottfried and Chew (1986) asked a counter tenor to pronounce ten vowels in “ $h \wedge d$ ” context (“ $\wedge$ ” signifying the vowel), in both a full voice (130 to 330 Hz) and a head voice (220 to 520 Hz). For both voices the error rate increased with  $F_0$ , but in the region of overlap (220 to 330 Hz) the head voice was more intelligible. Whatever the mechanism of vowel perception, a decrease in intelligibility with increases in  $F_0$  is understandable from the progressively sparser sampling of the spectral envelope. Apart from a few exceptions (see the histograms of Benolken and Swanson, 1990), the decrease is gradual and one does not see the sort of irregularities that might be expected as the peaks of the spectral envelope are swept by an expanding comb of voice harmonics.

As a counterpoint, it is interesting to note that Hillenbrand and Neary (1999) found only minor differences in identification rate between vowels pronounced by male (M), female (F), or child (C) speakers, despite  $F_0$  differences on the order of an octave. Moreover, the ranking was inconsistent according to whether the vowels were natural [rates ranked as (M,F) > C] or resynthesized with formant trajectories that were natural [M > (F,C)] or static [(M,C) > F]. Resynthesized vowels used the original vowels'  $F_0$  contour. If identifiability decreased with  $F_0$ , one should expect instead the ranking M > F > C to prevail uniformly. Identification was better for natural than for static formant contours, which confirms the importance of dynamic cues. However, it was better still for natural vowels, suggesting that cues available in natural vowels are not completely exhausted by dynamic formant and  $F_0$  values.

Attempts have been made to relate the observable harmonic peaks of a short-term spectrum with the underlying (but not directly observable) formant peaks. The problem is different for front vowels, for which  $F_1$  is relatively low and

well separated from other formants, and back vowels, for which  $F_1$  and  $F_2$  are close. For front vowels, the difficulty is that of estimating  $F_1$ , given that harmonic spacing is relatively wide compared to formant bandwidth and peripheral resolution. The second formant is less of an issue because that formant is wider, peripheral filtering is less sharp, and  $F_2$  is less in need of resolution from other formants, given that the assumption that higher formants are grouped to form an “ $F_2'$ ” works quite well (Carlson *et al.*, 1975). It has been proposed that  $F_1$  is determined from the position of the single most prominent harmonic (Mushnikov and Chistovich, 1972), or a combination of the two or three most prominent harmonics in the  $F_1$  region (Carlson *et al.*, 1975). Assmann and Neary (1987) found that a weighted sum of two harmonics was a better predictor than either one harmonic or three. However, Darwin and Gardner (1985) found evidence that more remote harmonics also influenced the perceived  $F_1$  position.

The question of whether one, two, or three harmonics affect  $F_1$  is a good empirical question, but an awkward basis for building models of vowel perception. A model exploiting the weighted sum of two harmonics, for example, must locate them. Criteria might be spectral peaks, or a harmonic series, but the model must solve the limit case where  $F_0$  is low and individual harmonics are not well resolved, as well as the opposite case when only one harmonic (possibly none) belongs to the formant. “Return on investment” is limited because the model cannot be applied to vowels for which  $F_1$  and  $F_2$  are close (see below). Finally, such a model is likely to predict irregular variations of  $F_1$  boundaries when  $F_0$  varies. A one-harmonic model predicts a saw-toothed variation of the  $F_1$  estimate: as the most prominent harmonic shifts down the slope of the  $F_1$  peak, it is suddenly replaced by a different harmonic. A two-harmonic model has a similar problem (its severity depends on the weight of the weaker harmonic relative to the stronger). Hirahara (1993) did in fact find an irregularity in the variation of the  $F_1$  boundary of a Japanese /i/-/e/ continuum. The boundary shifted from 320 to 400 Hz as  $F_0$  increased from 100 to 150 Hz, stayed at 400 Hz as  $F_0$  increased further to 250 Hz, then increased to about 520 Hz as  $F_0$  increased to 450 Hz. This behavior is indeed irregular, but not quite what one expects from a harmonic tracking mechanism. One expects instead a general decrease in  $F_1$  boundary to compensate for the upward shift of the harmonics, with a local increase at each switch between harmonics. To summarize, schemes based on individual harmonics or their weighted sums are an incomplete answer to the problems posed by harmonic structure.

For back vowels the problem is yet more complex, as pointed out by Assmann (1991).  $F_1$  and  $F_2$  are close and harder to relate individually to the set of harmonics than in the case of an isolated  $F_1$ . It has been proposed that when  $F_1$  and  $F_2$  are closely spaced (less than 3 to 3.5 bark), they are “merged” into a single spectral prominence, characterized by its center of gravity (Chistovich and Lublinskaya, 1979). The position of the center of gravity (COG) should be affected by both the frequencies and the amplitudes of both formants. Assmann (1991) tested this hypothesis with negative results. In one of his experiments, spectral integration

effects compatible with the COG hypothesis were found at  $F_0=250$  Hz, but not at 125 Hz. However, contrary to the criteria of Chistovich and Lublinskaya, they were *larger* for formants spaced more than 3–3.5 bark rather than smaller. Correct or not, the COG hypothesis is an awkward basis for a vowel perception model. If the COG is derived from formant estimates (with the convention that formants are grouped if closer than 3.5 bark), the problem of formant estimation remains entire. If the COG is derived from a smoothed spectrum, then at least three questions arise. First, how does one distinguish a peak reflecting an isolated formant ( $F_1$ ) from that reflecting a closely spaced pair ( $F_1-F_2$ )? Second, what is the appropriate form of smoothing? Third, is such smoothing not prone to the aliasing effects described earlier?

Klatt (1982) points out that resolution of the auditory periphery is a poor guide: it is much too sharp in the low-frequency region, except for the lowest  $F_0$ 's. The *ad hoc* value of 3.5 bark handles  $F_0$ 's up to 350 Hz, but no further, yet for lower  $F_0$ 's this smoothing may be unnecessarily severe. Logically, the best amount of smoothing depends on  $F_0$ , and indeed there is some evidence that the auditory system applies wider spectral integration at higher  $F_0$ 's (Assmann, 1991). The “simple vowel classifier” of Scheffers (1983) and the PEAK procedure of Assmann and Summerfield (1989) both imply  $F_0$ -dependent smoothing. Both reconstruct an envelope by linear interpolation between samples of the spectrum at harmonics, an operation equivalent to smoothing by convolution with a triangular window of width  $2F_0$  [similar smoothing is employed in the STRAIGHT analysis system of Kawahara (1997)]. Such precisely tuned  $F_0$ -dependent smoothing is certainly more effective than the fixed smoothing assumed by the COG hypothesis (or the “second integration” of Rosner and Pickering). However, two things are worth noting. First, the ambiguity between a single formant and two closely spaced formants is not resolved [consider for example interpolating between peaks of Fig. 2(b) or (c)]. Second, as other smoothing schemes, this one is susceptible to the aliasing problems analyzed in the Introduction.

The “missing-data” model is based on spectral samples that do not necessarily coincide with formant peaks. In this sense, it resembles the “whole spectrum” model of Bladon (1982). In that model, the spectrum was smoothed by convolution with an “auditory filter” (Bladon and Lindblom, 1981) which makes it susceptible to the aliasing problems pointed out in the Introduction. The missing-data model avoids them, and can thus be seen as an  $F_0$ -insensitive implementation of Bladon's ideas. The model belongs to the “top-down” or “analysis-by-synthesis” variety (Bell *et al.*, 1961; Rabiner and Schafer, 1978), in the sense that it does not try to extract an  $F_0$ -invariant representation from the waveform. Instead, it synthesizes a pattern ( $F_0$ -sampled spectrum, or  $F_0$ -adjusted autocorrelation function) to be compared with incoming patterns.

Bladon's model does not account for the insensitivity to formant bandwidth, formant amplitude, or spectral tilt observed by Klatt (1982), and the present model in its spectral version (Sec. IA) is prone to the same criticism. The



autocorrelation-based version offers some flexibility to deal with this problem. Differences in formant bandwidth affect mainly the high-lag part of the autocorrelation function, and sensitivity to this parameter can be reduced by ignoring features beyond a certain fixed lag. It is tempting to treat spectral tilt and formant amplitude in the same way, by ignoring lag-domain features *below* a certain lag, a technique that works well with the Fourier transform of the log magnitude spectrum, or cepstrum, and is commonly used in speech processing (e.g., Tohkura, 1987). Unfortunately this idea makes less sense applied to a transform of the square-magnitude spectrum (autocorrelation). Nevertheless, the filter-bank implementation of Sec. IC does offer a degree of insensitivity to gross spectral features, thanks to amplitude compression in the hair-cell model. A whole-spectrum model does not account for the fact that spectral peaks (formants) are known to carry a stronger weight than spectral valleys. However, the autocorrelation function puts a strong weight on spectral peaks (essentially  $F_1$  and  $F_2$ ), that are emphasized in the square-magnitude spectrum. In the multichannel version of Sec. IC, this rather extreme emphasis is softened by the automatic gain control (AGC) properties of the hair-cell model, that allow weaker features to be better represented.

A key idea is that the auditory system applies *variable* weights to different parts of incoming evidence according to their reliability. For example, relatively fine features of the spectral envelope may be used at low  $F_0$  but ignored at high  $F_0$ . Assmann (1991) found stronger evidence for the center-of-gravity hypothesis (that emphasizes gross spectral features) at 250 Hz than at 125 Hz. He also suggested that the reduced amplitude of the higher formant region ( $>1$  kHz) might serve as a cue for the presence of two closely spaced low formants ( $<1$  kHz), but only *when  $F_0$  is high*. When  $F_0$  is low, evidence of the presence of two formants would be derived instead from the detailed spectral shape in the low-frequency region. Different sources of evidence would thus be weighted differently according to their reliability. The same principle can account for the fact that in general (for low  $F_0$ ) the amplitude balance between widely spaced parts of the spectrum has little effect on vowel identity (Chistovich, 1985; Klatt, 1982). It can also explain the finding of Beddor and Hawkins (1990) that overall spectral shape was important for vowel matching when spectral peaks were broad (nasal or wide formants), whereas detailed spectral shape (formant position) was important when formants were narrow.

Versnel and Shamma (1998) presented evidence for a Fourier-transform representation of spectral shape in the primary auditory cortex of the ferret. Responses to vowels could be predicted from responses to sine-wave ripple stimuli. The principle is related to that of our autocorrelation scheme, the greatest difference being that their representation used a *logarithmic* axis for frequency. In other words, ripples had a constant period in octaves rather than Hz. Vowel stimuli were either voiced, extracted from the TIMIT database with  $F_0$ 's in the range 100–130 Hz, or unvoiced (synthesized with a 20 component/octave carrier). The authors noted little difference between voiced and unvoiced responses.

The spectral-domain version of the model is similar to

the “harmonic sieve” of Duifhuis *et al.* (1982). The harmonic sieve was proposed by Duifhuis as a means to select the components of a sound to be included in the calculation of its pitch. Scheffers (1983) used it to assign components of a mixed-speech spectrum to each voice, and Moore *et al.* (1984, 1985), and Darwin and Ciocca (1992) showed that a harmonic sieve with a width of about 3% determined which components of a sound contribute to its pitch. Darwin and Gardner (1986) found that mistuning a component of a vowel by 3% reduced its contribution to the vowel's quality. In those cases, the harmonic sieve played an important role in *segregating* the harmonic sound from competing components. Here, we suggest that it also plays a role in handling the bias due to  $F_0$  in the identification of isolated vowels. Note that this proposition seems to contradict results that show that the identification of a member of a concurrent vowel pair is no better when that vowel is harmonic rather than inharmonic (de Cheveigné *et al.*, 1997).

### III. CONCLUSIONS

- (1) Because of spectral undersampling, the information available to describe the spectral envelope of a steady-state voiced vowel is incomplete. It is limited to a series of  $F_0$ -spaced samples in the frequency domain, or in the lag domain to lag components below the Nyquist lag ( $1/2F_0$ ).
- (2) If the sampled spectrum is smoothed, lag-domain components beyond the Nyquist lag are aliased and contribute spurious  $F_0$ -dependent components to the smoothed spectrum. Aliasing is more severe if  $F_0$  is high, and if the spectral envelope is rich in high-lag components (narrow formants).
- (3) *Sampling* is evident in the harmonic structure of representations such as the short-term spectrum or auditory excitation pattern. *Aliasing* is evident in the fact that smoothing or interpolation of these patterns does not produce an  $F_0$ -invariant representation.
- (4) In a vowel identification model, effects of aliasing can be eliminated by (a) avoiding spectral smoothing, and (b) restricting pattern matching to the available samples. The model can be implemented in the spectral domain using a harmonic sieve based on an estimate of  $F_0$ .
- (5) The model can also be implemented in the lag domain based on the autocorrelation function of the waveform. Pattern matching is restricted to lags smaller than the Nyquist lag. Templates must be adjusted based on the  $F_0$  estimate to compensate for lag-domain aliasing.
- (6) Both versions of the model ensure  $F_0$ -independent pattern matching. They do not address other known sources of  $F_0$  dependency of vowel production or perception. They also do not deal with the loss of information due to sampling. Sparse sampling at high  $F_0$  should lead to progressively degraded identification, as is indeed observed experimentally.
- (7) The autocorrelation model might be implemented physiologically by neural spike coincidence counting circuits within channels of the auditory nerve, in a manner similar to Licklider's pitch-perception model. Nonlinearity of hair-cell transduction makes template adjustment more

difficult, but the extra tonotopic dimension provides a rich substrate for pattern matching, information weighting, and source segregation.

## ACKNOWLEDGMENTS

Part of this work was carried out at ATR Human Information Processing Research Laboratories, within a research agreement with the Center National de la Recherche Scientifique and the University of Paris 7. The author thanks ATR for its kind hospitality, and the CNRS for leave of absence. Thanks to Malcolm Slaney for the AuditoryToolbox software, and to Minoru Tszuzaki, Hiroaki Kato, Tatsuya Hirahara, Erik McDermott, Peter Assmann, Chris Darwin, and Kuldip Paliwal for comments on previous versions of this manuscript. Thanks also to Hartmut Traunmüller, two anonymous reviewers, and the editor for their useful criticism.

<sup>1</sup>Here, we consider the Fourier transform of the magnitude spectrum. Later on in the paper, we consider the Fourier transform of the *squared* magnitude spectrum, or autocorrelation function. It is also common to consider the Fourier transform of the *log* magnitude spectrum, or cepstrum. The principle is the same in each case. The representations differ, however, in their lag-domain composition, and the result of filtering or smoothing varies between representations.

<sup>2</sup>Frequencies of formants  $F_1$ ,  $F_2$ ,  $F_3$ ,  $F_4$  were taken from Hirahara and Kato (1992). They were, respectively (750, 1187, 1595, 3781 Hz) for /a/, (469, 2031, 2687, 3375 Hz) for /e/, (281, 2281, 3187, 3781 Hz) for /i/, (468, 781, 2656, 3281 Hz) for /o/, and (312, 1219, 2469, 3375 Hz) for /u/. Formants  $F_5$  of all vowels were set to 4200 Hz. Bandwidths of formants had identical values for all vowels: 90, 110, 170, 250, and 300 Hz, respectively, for formants  $F_1$  to  $F_5$ .

Ainsworth, W. A. (1975). "Intrinsic and extrinsic factors in vowel judgments," in *Auditory Analysis and Perception of Speech*, edited by G. Fant and M. A. A. Tatham (Academic, London), pp. 103–113.

Assmann, P. F., Neary, T. M., and Hogan, J. T. (1982). "Vowel identification: orthographic, perceptual, and acoustic effects," *J. Acoust. Soc. Am.* **71**, 975–989.

Assmann, P. F., and Neary, T. M. (1987). "Perception of front vowels: the role of harmonics in the first formant region," *J. Acoust. Soc. Am.* **81**, 520–534.

Assmann, P. F., and Summerfield, Q. (1989). "Modeling the perception of concurrent vowels: Vowels with the same fundamental frequency," *J. Acoust. Soc. Am.* **85**, 327–338.

Assmann, P. (1991). "The perception of back vowels: centre of gravity hypothesis," *Q. J. Exp. Psychol.* **43A**, 423–448.

Beddor, P. S., and Hawkins, S. (1990). "The influence of spectral prominence on perceived vowel quality," *J. Acoust. Soc. Am.* **87**, 2684–2704.

Bell, C. G., Fujisaki, H., Heinz, J. M., Stevens, K. N., and House, A. S. (1961). "Reduction of speech spectra by analysis-by-synthesis techniques," *J. Acoust. Soc. Am.* **33**, 1725–1736.

Benolken, M. S., and Swanson, C. E. (1990). "The effect of pitch-related changes on the perception of sung vowels," *J. Acoust. Soc. Am.* **87**, 1781–1785.

Bladon, R. A. W., and Lindblom, B. (1981). "Modeling the judgment of vowel quality differences," *J. Acoust. Soc. Am.* **69**, 1414–1422.

Bladon, A. (1982). "Arguments against formants in the auditory representation of speech," in *The Representation of Speech in the Peripheral Auditory System*, edited by R. Carlson and B. Granström (Amsterdam, Elsevier), pp. 95–102.

Bonneau, A. (1996). "Identification of vowels from french stop bursts," in *Proceedings of the ESCA Workshop on the Auditory Basis of Speech Perception*, pp. 133–136.

Cariani, P. A., and Delgutte, B. (1996a). "Neural correlates of the pitch of complex tones. I Pitch and pitch salience," *J. Neurophysiol.* **76**, 1698–1716.

Cariani, P. A., and Delgutte, B. (1996b). "Neural correlates of the pitch of complex tones. II Pitch shift, pitch ambiguity, phase-invariance, pitch circularity, rate-pitch and the dominance region for pitch," *J. Neurophysiol.* **76**, 1717–1734.

Carlson, R., Fant, G., and Granström, B. (1975). "Two-formant models, pitch and vowel perception," in *Auditory Analysis and Perception in Speech*, edited by G. Fant and M. A. A. Tatham (Academic, London), pp. 55–82.

Carré, R., and Lancia, R. (1975). "Perception of vowel amplitude transients," in *Auditory Analysis and Perception of Speech*, edited by G. Fant and M. A. A. Tatham (Academic, London), pp. 83–90.

Chistovich, L. A., and Lublinskaja, V. V. (1979). "The 'center of gravity' effect in vowel spectra and critical distance between the formants: psychoacoustical study of the perception of vowel-like stimuli," *Hearing Res.* **1**, 185–195.

Chistovich, L. A. (1985). "Central auditory processing of peripheral vowel spectra," *J. Acoust. Soc. Am.* **77**, 789–805.

Cooke, M., Green, P., Anderson, C., and Abberley, D. (1994). "Recognition of occluded speech by hidden markov models," University of Sheffield Department of Computer Science Technical Report, TR-94-05-01.

Cooke, M., Morris, A., and Green, P. (1997). "Missing data techniques for robust speech recognition," *Proceedings of the ICASSP*, pp. 863–866.

Darwin, C. J., and Gardner, R. B. (1985). "Which harmonics contribute to the estimation of first formant frequency?," *Speech Commun.* **4**, 231–235.

Darwin, C. J., and Gardner, R. B. (1986). "Mistuning of a harmonic of a vowel: Grouping and phase effects on vowel quality," *J. Acoust. Soc. Am.* **79**, 838–845.

Darwin, C. J., and Ciocca, V. (1992). "Grouping in pitch perception: effects of onset asynchrony and ear of presentation of a mistuned component," *J. Acoust. Soc. Am.* **91**, 3381–3390.

de Cheveigné, A., McAdams, S., and Marin, C. (1997). "Concurrent vowel identification II. Effects of phase, harmonicity and task," *J. Acoust. Soc. Am.* **101**, 2848–2856.

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. (1998). "Cancellation model of pitch perception," *J. Acoust. Soc. Am.* **103**, 1261–1271.

de Cheveigné, A., and Kawahara, H. (1998). "A model of vowel perception based on missing feature theory," ATR-HIP Technical Report, TR-H-252.

Duifhuis, H., Willems, L. F., and Sluyter, R. J. (1982). "Measurement of pitch in speech: an implementation of Goldstein's theory of pitch perception," *J. Acoust. Soc. Am.* **71**, 1568–1580.

Fahey, R. P., Diehl, R. L., and Traunmüller, H. (1996). "Perception of back vowels: effects of varying  $F_1$ – $F_0$  bark distance," *J. Acoust. Soc. Am.* **99**, 2350–2357.

Fairbanks, G. (1940). "Recent experimental investigations of vocal pitch in speech," *J. Acoust. Soc. Am.* **11**, 457–466.

Fant, G. (1970). *Acoustic Theory of Speech Production* (Mouton, The Hague).

Gottfried, T. L., and Chew, S. L. (1986). "Intelligibility of vowels sung by a countertenor," *J. Acoust. Soc. Am.* **79**, 124–130.

Hillenbrand, J., and Gayvert, R. T. (1993). "Identification of steady-state vowels synthesized from the Peterson and Barney measurements," *J. Acoust. Soc. Am.* **94**, 668–674.

Hillenbrand, J. M., and Nearey, T. M. (1999). "Identification of resynthesized /hVd/ utterances: Effects of formant contour," *J. Acoust. Soc. Am.* (in press).

Hirahara, T., and Kato, H. (1992). "The effect of  $F_0$  on vowel identification," in *Speech Perception, Production and Linguistic Structure*, edited by Y. Tohkura, E. Vatikiotis-Bateson, and Y. Sagisaka (Ohmsha, Tokyo), pp. 89–112.

Hirahara, T. (1993). "On the role of relative harmonics level around the  $F_1$  of high vowel identification," *Proceedings of the ARO abstracts (ISSN 0742-3152)*, abstract #258, p. 65.

Hirahara, T., Cariani, P., and Delgutte, B. (1996). "Representation of low-frequency vowel formants in the auditory nerve," *Proceedings of the ESCA Workshop on the Auditory Basis of Speech Perception*, pp. 83–86.

Hoemeke, K. A., and Diehl, R. L. (1994). "Perception of vowel height: the role of  $F_1$ – $F_0$  distance," *J. Acoust. Soc. Am.* **96**, 661–674.

Howard, I. (1991). "Speech fundamental period estimation using pattern classification," London, unpublished doctoral dissertation.

- Jeffress, L. A. (1948). "A place theory of sound localization," *J. Comp. Physiol. Psychol.* **41**, 35–39.
- Kawahara, H. (1997). "Speech representation and transformation using adaptive interpolation of weighted spectrum: vocoder revisited," *Proceedings of the ICASSP*, pp. 1303–1306.
- Keating, P., and Buhr, R. (1978). "Fundamental frequency in the speech of infants and children," *J. Acoust. Soc. Am.* **63**, 567–571.
- Kewley-Port, D. (1996). "Psychophysical studies of vowel formants," *Proceedings of the ESCA Workshop on the Auditory Basis of Speech Perception*, pp. 148–153.
- Kewley-Port, D., Li, X., Zheng, Y., and Neel, A. T. (1996). "Fundamental frequency effects on thresholds for vowel formant discrimination," *J. Acoust. Soc. Am.* **100**, 2462–2470.
- Klatt, D. H. (1982). "Speech processing strategies based on auditory models," in *The Representation of Speech in the Peripheral Auditory System*, edited by R. Carlson and B. Granström (Elsevier, Amsterdam), pp. 181–196.
- Licklider, J. C. R. (1951). "A duplex theory of pitch perception," *Experientia* **7**, 128–134.
- Lippmann, R. P., and Carlson, B. A. (1997). "Using missing feature theory to actively select features for robust speech recognition with interruptions, filtering, and noise," *Proceedings of ESCA Eurospeech*, KN-37–40.
- Marin, C., and McAdams, S. (1991). "Segregation of concurrent sounds. II. Effects of spectral envelope tracing, frequency modulation coherence, and frequency modulation width," *J. Acoust. Soc. Am.* **89**, 341–351.
- McAdams, S., and Rodet, X. (1988). "The role of FM-induced AM in dynamic spectral profile analysis," in *Basic Issues in Hearing*, edited by H. Duifhuis, J. Horst, and H. Wit (Academic, London), pp. 359–369.
- McKeown, J. D., and Patterson, R. D. (1996). "The time course of auditory segregation: Concurrent vowels that vary in duration," *J. Acoust. Soc. Am.* **98**, 1866–1877.
- Meddis, R., and Hewitt, M. J. (1991a). "Virtual pitch and phase sensitivity of a computer model of the auditory periphery. I: Pitch identification," *J. Acoust. Soc. Am.* **89**, 2866–2882.
- Meddis, R., and Hewitt, M. J. (1991b). "Virtual pitch and phase sensitivity of a computer model of the auditory periphery. II: Phase sensitivity," *J. Acoust. Soc. Am.* **89**, 2883–2894.
- Meddis, R., and Hewitt, M. J. (1992). "Modeling the identification of concurrent vowels with different fundamental frequencies," *J. Acoust. Soc. Am.* **91**, 233–245.
- Miller, J. D. (1989). "Auditory-perceptual interpretation of the vowel," *J. Acoust. Soc. Am.* **85**, 2114–2134.
- Moore, B. C. J., Glasberg, B. R., and Shailer, M. J. (1984). "Frequency and intensity difference limens for harmonics within complex tones," *J. Acoust. Soc. Am.* **75**, 550–561.
- Moore, B. C. J., Peters, R. W., and Glasberg, B. R. (1985). "Thresholds for the detection of inharmonicity in complex tones," *J. Acoust. Soc. Am.* **77**, 1861–1867.
- Morris, A. C., Cooke, M. P., and Green, P. D. (1998). "Some solutions to the missing feature problem in data classification, with application to noise robust ASR," *Proceedings of the ICASSP*, pp. 737–740.
- Mushnikov, V. N., and Chistovich, L. A. (1972). "Method for the experimental investigation of the role of component loudnesses in the recognition of a vowel," *Sov. Phys.-Acoust.* **17**, 339–344.
- Neary, T. M. (1989). "Static, dynamic, and relational properties in vowel perception," *J. Acoust. Soc. Am.* **85**, 2088–2113.
- Perec, G. (1991). "Experimental demonstration of the tomatopic organization in the soprano (*Cantatrix sopranica L.*)," in *Cantatrix Sopranica L.*, edited by G. Perec (Paris), Seuil, pp. 11–33.
- Peterson, G. E., and Barney, H. L. (1952). "Control methods in a study of the vowels," *J. Acoust. Soc. Am.* **24**, 175–184.
- Rabiner, L. R., and Schafer, R. W. (1978). *Digital Processing of Speech Signals* (Prentice-Hall, Englewood Cliffs, NJ).
- Robinson, K., and Patterson, R. D. (1995). "The stimulus duration required to identify vowels, their octave, and their pitch chroma," *J. Acoust. Soc. Am.* **98**, 1858–1865.
- Ryalls, J. H., and Lieberman, P. (1982). "Fundamental frequency and vowel perception," *J. Acoust. Soc. Am.* **72**, 1631–1634.
- Rosner, B. S., and Pickering, J. B. (1994). *Vowel Perception and Production* (Oxford University Press, Oxford).
- Scheffers, M. T. M. (1983). "Sifting vowels," Gröningen unpublished doctoral dissertation.
- Slaney, M. (1993). "An efficient implementation of the Patterson–Holdsworth auditory filter bank," Apple Computer Technical Report, pp. 35.
- Slawson, A. W. (1967). "Vowel quality and musical timbre as functions of spectrum envelope and fundamental frequency," *J. Acoust. Soc. Am.* **43**, 87–101.
- Strange, W., Verbrugge, R. R., Shankweiler, D. P., and Edman, T. R. (1976). "Consonant environment specifies vowel identity," *J. Acoust. Soc. Am.* **60**, 213–224.
- Strange, W., and Bohn, O.-S. (1998). "Dynamic specification of coarticulated German vowels: Perceptual and acoustical studies," *J. Acoust. Soc. Am.* **104**, 488–504.
- Summerfield, Q., Haggard, M., Foster, J., and Gray, S. (1984). "Perceiving vowels from uniform spectra: phonetic exploration of an auditory aftereffect," *Percept. Psychophys.* **35**, 203–213.
- Sundberg, J. (1982). "Perception of singing," in *The Psychology of Music*, edited by D. Deutsch (Academic, Orlando, FL), pp. 59–98.
- Sundberg, J., and Gauffin, J. (1982). "Amplitude of the voice fundamental and the intelligibility of super pitch vowels," in *The Representation of Speech in the Peripheral Auditory System*, edited by R. Carlson and B. Granström, pp. 223–228.
- Tohkura, Y. (1987). "A weighted cepstral distance measure for speech recognition," *IEEE Trans. Acoust., Speech, Signal Process.* **35**, 1414–1422.
- Trautmüller, H. (1981). "Perceptual dimension of openness in vowels," *J. Acoust. Soc. Am.* **69**, 1465–1475.
- Trautmüller, H. (1990). "A note on hidden factors in vowel perception experiments," *J. Acoust. Soc. Am.* **88**, 2015–2019.
- Versnel, H., and Shamma, S. (1998). "Spectral-ripple representation of steady-state vowels," *J. Acoust. Soc. Am.* **103**, 5502–2514.
- Yin, T. C. T., Chan, J. C. K., and Carney, L. H. (1987). "Effects of interaural time delays of noise stimuli on low-frequency cells in the cat's inferior colliculus. III Evidence for cross-correlation," *J. Neurophysiol.* **58**, 562–583.

# Identification of resynthesized /hVd/ utterances: Effects of formant contour

James M. Hillenbrand

*Department of Speech Pathology and Audiology, Western Michigan University, Kalamazoo, Michigan 49008*

Terrance M. Nearey

*Department of Linguistics, University of Alberta, Edmonton, Alberta T6G 2E7, Canada*

(Received 15 May 1998; revised 1 November 1998; accepted 10 March 1999)

The purpose of this study was to examine the role of formant frequency movements in vowel recognition. Measurements of vowel duration, fundamental frequency, and formant contours were taken from a database of acoustic measurements of 1668 /hVd/ utterances spoken by 45 men, 48 women, and 46 children [Hillenbrand *et al.*, *J. Acoust. Soc. Am.* **97**, 3099–3111 (1995)]. A 300-utterance subset was selected from this database, representing equal numbers of 12 vowels and approximately equal numbers of tokens produced by men, women, and children. Listeners were asked to identify the original, naturally produced signals and two formant-synthesized versions. One set of “original formant” (OF) synthetic signals was generated using the measured formant contours, and a second set of “flat formant” (FF) signals was synthesized with formant frequencies fixed at the values measured at the steadiest portion of the vowel. Results included: (a) the OF synthetic signals were identified with substantially greater accuracy than the FF signals; and (b) the naturally produced signals were identified with greater accuracy than the OF synthetic signals. Pattern recognition results showed that a simple approach to vowel specification based on duration, steady-state  $F_0$ , and formant frequency measurements at 20% and 80% of vowel duration accounts for much but by no means all of the variation in listeners’ labeling of the three types of stimuli. © 1999 Acoustical Society of America. [S0001-4966(99)04406-9]

PACS numbers: 43.72.Ar, 43.71.Es, 43.72.Ja [JH]

## INTRODUCTION

There is a long tradition of representing vowels by a single spectral cross section taken from the nucleus of the vowel. The essence of this approach was summarized nicely by Tiffany (1953):

It has been commonly assumed or implied that the essential physical specification of a vowel phoneme could be accomplished in terms of its acoustic spectrum as measured over a single fundamental period, or over a short interval including at most a few cycles of the fundamental frequency. That is to say, each vowel has been assumed to have a unique energy vs frequency distribution, with the significant physical variables all accounted for by an essentially cross-sectional analysis of the vowel’s harmonic composition. (p. 290).

The potential limitations of this static approach to vowel quality were recognized by Tiffany, who noted that vowel duration and changes over time in formant frequencies and the fundamental frequency ( $F_0$ ) may play a role in vowel perception (see also similar comments by Potter and Steinberg, 1950; Peterson and Barney, 1952; and Stevens and House, 1963). While the role of  $F_0$  contour and duration received a fair amount of attention in early accounts of vowel recognition (e.g., Ainsworth, 1972; Bennett, 1968; Black, 1949; Stevens, 1959; Tiffany, 1953), it has only been more recently that the role of formant frequency movements has been examined systematically.

Evidence from several studies suggests that formant frequency movements do, in fact, play an important role in

vowel perception. For example, Strange *et al.* (1983) and Nearey and Assmann (1986) showed high identification rates for “silent center” stimuli in which the vowel centers were gated out, leaving only brief onglides and offglides (see also Jenkins *et al.*, 1983; Parker and Diehl, 1984). Nearey and Assmann (1986) showed that it was not simply spectral movement that was required, but a specific pattern of spectral change throughout the course of the vowel. Brief excerpts of naturally produced vowels excised from nucleus and offglide segments were presented to listeners in three conditions: (1) natural order (nucleus followed by offglide); (2) repeated nucleus (nucleus followed by itself); and (3) reverse order (offglide followed by nucleus). Identification error rates for the natural-order signals were comparable to those for the original, unmodified vowels, while the repeated-nucleus and reverse-order conditions produced much higher error rates.

There is also evidence that vowels with static formant patterns are not particularly well identified. Hillenbrand and Gayvert (1993a) synthesized 300-ms monotone vowels with stationary formant patterns from the  $F_0$  and formant measurements of each of the 1520 tokens in the Peterson and Barney (1952) /hVd/ database (2 repetitions of 10 vowels spoken by 33 men, 28 women, and 15 children). The 27% identification error rate for these steady-state synthetic signals was nearly five times greater than the error rate reported by Peterson and Barney for the original utterances. Synthesizing the signals with a falling pitch contour resulted in a highly significant but relatively small drop in the error rate. These results suggest that the duration and spectral change

information that was removed from these stimuli by the steady-state synthesis method plays an important role in vowel recognition. Similarly high identification error rates were reported by Fairbanks and Grubb (1961) for naturally produced sustained vowels.

Evidence implicating a role for spectral change also comes from studies using statistically based pattern classifiers. For example, Assmann *et al.* (1982) trained a linear discriminant classifier with: (a) steady-state  $F_0$  and formant information alone; and (b) steady-state information plus formant slopes and duration. The pattern classification model that incorporated dynamic information provided more accurate predictions of error patterns produced by human listeners (see also Nearey and Assmann, 1986; Parker and Diehl, 1984). Hillenbrand *et al.* (1995) trained a quadratic discriminant classifier on  $F_0$  and formant measurements from /hVd/ utterances spoken by 45 men, 48 women, and 46 children. The pattern classifier was trained on various combinations of acoustic measurements, with formant frequencies sampled either at steady state or at 20% and 80% of vowel duration. The classifier was much more accurate when it was trained on two samples of the formant pattern. For example, with  $F_1$  and  $F_2$  alone, the classification accuracy was 71% for a single sample at steady state and 91% for two samples of the formant pattern. Two-sample parameter sets including either  $F_0$  or  $F_3$  produced classification accuracies approaching those of human listeners. Adding vowel duration to the parameter set also improved classification accuracy, although the effect was relatively small if the formant pattern was sampled twice.

While the pattern classification evidence is clearly relevant, it needs to be kept in mind that demonstrating that a pattern classifier is strongly affected by spectral change is not the same as showing that human listeners rely on spectral change patterns. As Nearey (1992) noted, pattern classification results have only an indirect bearing on perception unless the classification results are compared to human listener data (see also Shankweiler *et al.*, 1977). A particularly clear example of the limitations of pattern classification studies can be seen by comparing pattern classification results using static acoustic measurements with studies in which human listeners identify vowels with static formant patterns. Several pattern classification studies have shown that vowels can be identified with relatively modest error rates in the 12%–14% range based on static acoustic measurements (e.g., Hillenbrand and Gayvert, 1993b; Miller, 1984, 1989; Nearey, 1992; Nearey *et al.*, 1979; Syrdal and Gopal, 1986). However, Hillenbrand and Gayvert (1993a) reported a 27% error rate for human listeners who were asked to identify static vowels synthesized from the Peterson and Barney (1952)  $F_0$  and formant measurements, and Fairbanks and Grubb (1961) reported a 26% error rate for naturally produced static vowels. The Fairbanks and Grubb findings are especially striking since there were just seven talkers, all of them men, and the investigators went to great lengths to ensure the quality and representativeness of their test signals.

In our view, the primary lesson from this mismatch of pattern recognition results and listening tests is not that pattern classification evidence is necessarily irrelevant or mis-

leading, but rather, that pattern recognition studies need to be followed up with appropriately designed perception experiments. The purpose of the present experiment was to follow up on the pattern classification tests reported in Hillenbrand *et al.* (1995) by asking listeners to identify naturally produced /hVd/ signals and two synthetically generated versions. One set of synthesized signals was generated using the

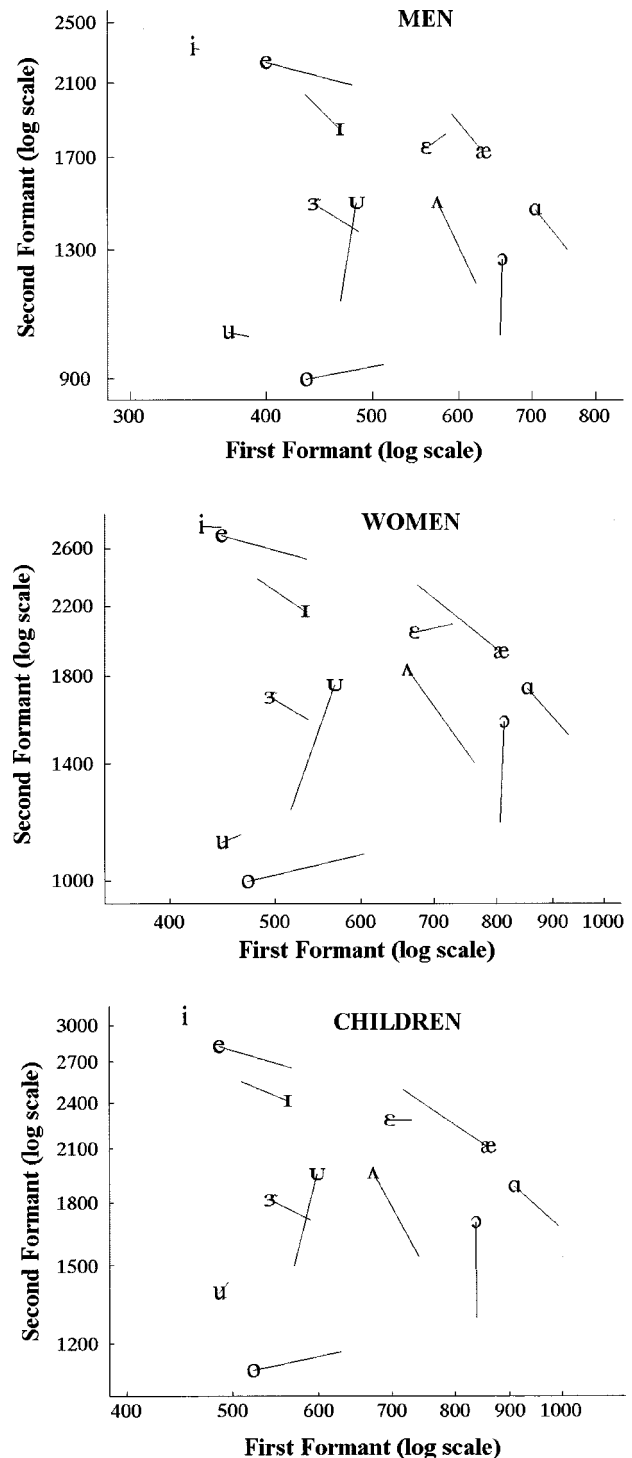


FIG. 1. Spectral change patterns for /hVd/ utterances produced by men, women, and children. The symbol identifying each vowel is plotted at the  $F_1$ – $F_2$  value for the second sample of the formant pattern (80% of vowel duration), and a line connects this point to the first sample (20% of vowel duration). The measurements are from Hillenbrand *et al.* (1995).

original measured formant contours and a second set of signals was synthesized with flat formants, fixed at the values measured at the steadiest portion of the vowel.

## I. EXPERIMENT 1. METHODS

### A. Test signals

The test signals consisted of a 300-stimulus subset of the 1668 /hVd/ utterances recorded by Hillenbrand *et al.* (1995). The talkers in that study consisted of 45 men, 48 women, and 46 10- to 12-year-old children. Recordings were made of subjects producing the vowels /i, I, e, ε, æ, a, ɔ, o, u, u, ʌ, ɜ/ in /hVd/ syllables. A computer program was written to select a 300-stimulus subset from the full set of 1668 signals. The 300 signals were selected at random, but with the following constraints: (a) signals showing formant mergers involving any of the three lowest formants were omitted; (b) signals with identification error rates (measured in the original 1995 study) of 15% or greater were omitted; and (c) all 12 vowels were equally represented. The 300-stimulus set that was selected by this method included tokens from 123 of the 139 talkers, with 30% of the tokens from men, 36% from women, and 34% from children.

### B. Acoustic measurements

Acoustic measurement techniques are described in detail in Hillenbrand *et al.* (1995). Briefly, peaks were extracted from LPC spectra every 8 ms and formant contours for  $F_1$ – $F_4$  were edited by hand during the vowel using a custom interactive editing tool. Measurements were also made of  $F_0$  contour (also edited by hand using the same editing tool) and three temporal quantities: (a) the onset of the vowel; (b) the offset of the vowel; and (c) steady-state time; i.e., the single frame at which the formant pattern was judged by visual inspection to be maximally steady. Vowel onsets and offsets were also judged by visual inspection, using standard measurement criteria (Peterson and Lehiste, 1960). Average spectral change patterns for the full data set are shown in Fig. 1. The figure was created by connecting a line between the formant frequencies sampled at 20% and 80% of vowel duration; the symbol for each vowel category is plotted at the location of the second measurement. The measurement results are described in some detail in Hillenbrand *et al.*, but there are two points about the formant-change patterns in Fig. 1 that are particularly relevant to the present study. First, with the exception of /i/ and /u/, the formant frequency values show a good deal of change throughout the course of the vowel. For example, note that the vowels /e/ and /o/, which are known to be diphthongized, do not show spectral change magnitudes that are unusually large relative to the other vowels. Second, the formants change in such a way as to enhance the contrast between vowels with similar formant patterns. For example, the pairs /æ/–/ε/ and /u/–/ʊ/, which show a good deal of overlap when the vowels of individual talkers are plotted in static  $F_1$ – $F_2$  space (see Fig. 4 of Hillenbrand *et al.*), show very different patterns of formant-frequency change.

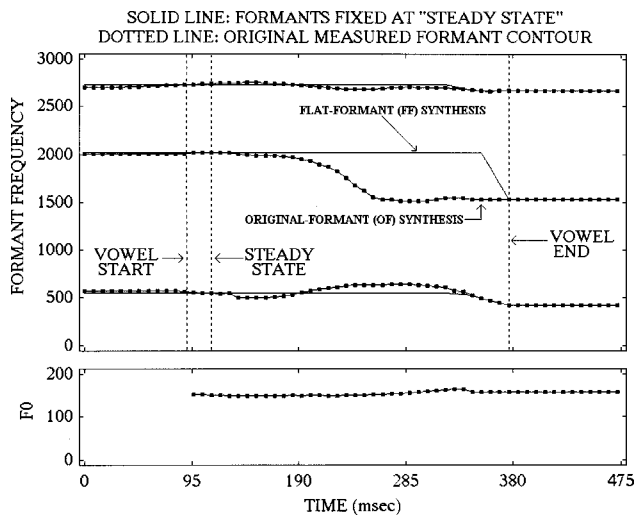


FIG. 2. Schematic spectrograms illustrating the method that was used to synthesize the original-formant (OF) and flat-formant (FF) signals.

### C. Synthesis method

Test signals consisted of the 300 original, 16-kHz digitized utterances and two synthesized versions, for a total of 900 signals. The Klatt and Klatt (1990) formant synthesizer, running at a 16-kHz sample rate, was used to generate two sets of synthetic signals. The “original formant” (OF) and “flat format” (FF) synthesis methods are illustrated in Fig. 2. The OF signals were synthesized using the original measured formant contours, shown by the dashed curves. The vowel in this example is /æ/, and the measured formant contour shows a pronounced offglide which is quite common in our data for this vowel. The FF signals were synthesized with flat formants, fixed at the values measured at steady state, shown by the solid curves in Fig. 2.<sup>1</sup> For these signals a 40-ms linear transition connected the steady-state values to the  $F_1$ – $F_3$  values that were measured at the end of the vowel. Both sets of synthetic signals were generated with the measured  $F_0$  contours and at their measured durations. During the /h/ interval (i.e., between the beginning of the stimulus and the start of the vowel): (a) the voicing amplitude (AV) parameter was set to zero and the aspiration amplitude (AH) parameter was controlled by the measured rms intensity of the signal being synthesized; (b) the  $F_1$  bandwidth was set to 300 Hz; and (c) formant values for  $F_1$ – $F_3$  were set to the values that were measured at the start of the vowel. At all other times formant bandwidths remained at their default values ( $B_1=90$ ,  $B_2=110$ ,  $B_3=170$ ,  $B_4=400$ ,  $B_5=500$ ,  $B_6=800$ ). During the vowel the AH parameter was set to zero and the AV parameter was driven by the measured rms energy of the signal. Values of  $F_4$  were set separately for each vowel and talker group based on data from Hillenbrand *et al.* (1995). Values of  $F_5$  and  $F_6$  were set separately for each talker group based on data from Rabiner (1968). Formant amplitudes were set automatically during the /h/ and vowel by running the synthesizer in cascade mode. A final /d/ was simulated by: (a) ramping  $F_1$  100 Hz below its measured value at the end of the vowel in four steps of 25 Hz; and (b) switching from the cascade to the parallel branch of the synthesizer and setting the resonator

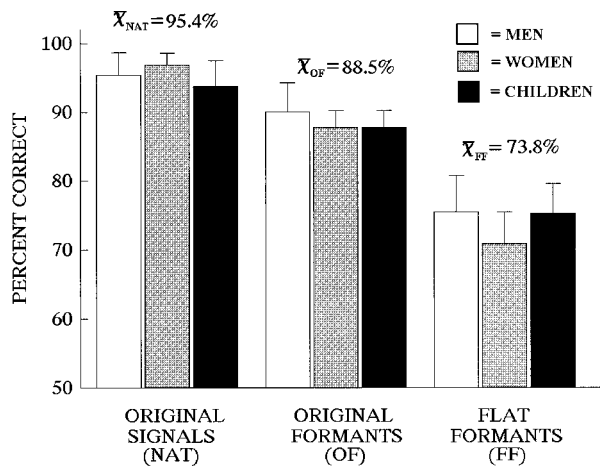


FIG. 3. Percent correct identification of the naturally produced signals, the original-formant synthetic signals, and the flat-formant synthetic signals. Error bars indicate one standard deviation. The percentages at the top indicate the mean percent correct for each condition pooled across the three talker groups.

gains of  $F_2$ – $F_6$  30 dB below the  $F_1$  resonator gain; i.e., producing a “voice bar” with energy primarily at  $F_1$ . Since we were not entirely satisfied with our efforts to generate natural sounding final release bursts with the synthesizer, the signals were generated unreleased, and release bursts that had been excised from naturally produced signals spoken by one man, one woman, and one child were appended to the end of the stimuli.<sup>2</sup>

#### D. Listening Test

Twenty subjects who had taken or were currently enrolled in an undergraduate course in phonetics served as listeners. The choice of listeners with training in phonetic transcription was motivated by the findings of Assmann *et al.* (1982) indicating that many of the identification errors made by untrained subjects are due to listeners’ uncertainty about how to map perceived vowel quality onto orthographic symbols. Subjects were tested one at a time in a quiet room in a single session lasting about 1 h. Listeners identified each of the 900 test signals (300 original signals, 300 OF signals, and 300 FF signals) presented in random order. The presentation order was reshuffled prior to each listening session. Stimuli were low-pass filtered at 6.9 kHz, amplified, and delivered at approximately 75 dBA over a single loudspeaker (Boston Acoustics A60) positioned approximately 1 m from the subject’s head. Subjects entered their responses on a computer keyboard labeled with both phonetic symbols and key words for the 12 vowels. Subjects were allowed to repeat stimuli as many times as they wished before entering a response.

## II. EXPERIMENT 1. RESULTS AND DISCUSSION

Figure 3 shows overall percent correct for each stimulus type and talker group averaged across all vowels. It can be seen that the naturally produced signals (NAT) were identified with the greatest accuracy, followed by the OF and FF synthesized signals. A two-way repeated-measures ANOVA using arcsine-transformed percent correct values showed a

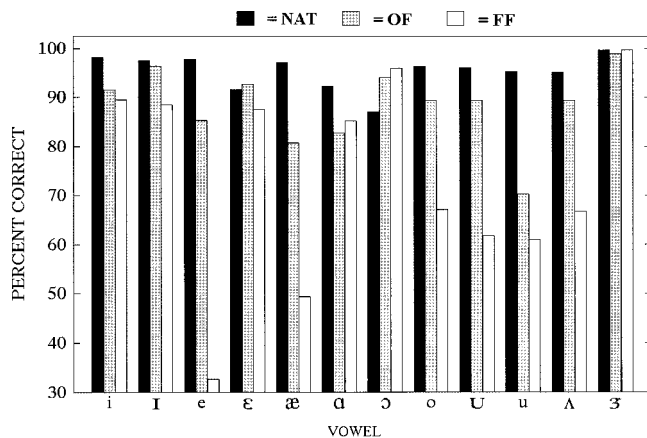


FIG. 4. Percent correct identification of the naturally produced signals (NAT), the original-formant (OF) synthetic signals, and the flat-formant synthetic signals separated by vowel category.

significant effect for stimulus type [ $F(2,38) = 860.7$ ,  $p < 0.0001$ ] and talker group [ $F(2,38) = 3.5$ ,  $p < 0.05$ ] and a significant interaction [ $F(4,76) = 10.2$ ,  $p < 0.0001$ ].<sup>3</sup> Newman–Keuls *post hoc* tests showed significant differences among all three stimulus types. Although statistically reliable, the effects for talker group are relatively small and nonuniform across stimulus type, as revealed by the significant interaction. For the naturally produced signals, *post hoc* analyses showed that the men’s and women’s tokens were identified with greater accuracy than the children’s tokens. The pattern was different for the OF synthetic signals, which showed greater intelligibility for the men’s tokens as compared to those of the women and children. A third pattern was observed for the FF synthetic signals, which showed significantly poorer intelligibility for the women’s tokens as compared to the men and the children. One clear conclusion from this rather mixed set of talker-group results is that there was no evidence for a simple pitch effect; that is, although the spectrum envelope is more poorly defined at higher fundamental frequencies, the talker-group effects show no evidence of a simple inverse relationship between  $F_0$  and vowel intelligibility [see also Carlson *et al.* (1975), and Hillenbrand and Gayvert (1993a), for related findings].

Figure 4 shows percent correct separately for each vowel category. Confusion matrices for the three conditions are shown in Tables I–III. It can be seen in Fig. 4 that the effect of stimulus type varies considerably from one vowel to the next. This was confirmed by a two-way repeated-measures ANOVA (using arcsine-transformed percent correct values) which tested the effects of stimulus type (NAT versus OF synthesis versus FF synthesis) and vowel. The ANOVA showed a significant effect for stimulus type [ $F(2,38) = 799.9$ ,  $p < 0.001$ ] and vowel [ $F(11,209) = 28.0$ ,  $p < 0.001$ ] as well as a significant interaction [ $F(22,418) = 46.7$ ,  $p < 0.001$ ]. The nature of the interaction will be discussed and further analyzed below.

Note that the pattern for /ɔ/ appears to differ markedly from the other vowels since this vowel showed the highest recognition accuracy in the FF condition and the lowest accuracy for the naturally produced signals. Examination of the confusion matrices in Tables I–III suggests that the overall

TABLE I. Confusion matrix for the naturally produced signals. Values on the main diagonal, indicating the percentage of trials in which the listeners' responses matched the vowel intended by the talker, are shown in boldface. The response means given in the last row indicate the percentage of trials in which a given vowel was used as a listener response. Each vowel, as classified by the speaker's intention, was presented on 8.3% of the trials.

		Vowel identified by listener											
		/i/	/ɪ/	/e/	/ɛ/	/æ/	/ɑ/	/ɔ/	/o/	/ʊ/	/u/	/ʌ/	/ɜ/
Vowel intended by talker	/i/	<b>98.2</b>	0.8	0.4	0.6	...	...	...	...	...	...	...	...
	/ɪ/	...	<b>97.6</b>	0.6	1.4	...	...	...	...	0.4	...	...	...
	/e/	0.2	0.2	<b>97.8</b>	1.6	...	...	...	...	0.2	...	...	...
	/ɛ/	...	...	0.2	<b>91.6</b>	8.0	...	...	...	...	...	0.2	...
	/æ/	...	...	...	2.0	<b>97.2</b>	0.4	0.4	...	...	...	...	...
	/ɑ/	...	...	...	...	3.2	<b>92.2</b>	4.0	...	...	...	0.6	...
	/ɔ/	...	...	...	0.2	0.8	9.6	<b>87.2</b>	0.4	0.6	...	1.6	...
	/o/	...	...	...	...	...	...	...	<b>96.4</b>	1.4	0.6	1.6	...
	/ʊ/	...	...	...	...	...	...	...	0.2	<b>96.2</b>	0.4	3.2	...
	/u/	...	...	...	...	...	...	...	1.6	3.0	<b>95.4</b>	...	...
	/ʌ/	...	...	...	...	...	2.0	1.0	...	1.4	0.4	<b>95.2</b>	...
	/ɜ/	...	...	...	0.2	...	...	...	...	...	...	...	<b>99.8</b>
Response means:		8.2	8.2	8.3	8.1	9.1	8.7	7.7	8.2	8.6	8.1	8.5	8.3

percent correct figures may be misleading in some respects. The last row in each of these tables gives the response means; that is, the percentage of trials in which a given vowel was used as a response. Since each of the 12 vowels was presented equally often, an ideal listener whose responses always agreed with the speaker's intention would use each symbol on 8.3% of the trials. Note that the percentage of /ɔ/ responses increases from 7.7% for the naturally produced signals, to 8.6% for the OF synthetic signals, to 11.2% for the FF synthetic signals. In other words, for reasons that are not clear, there is an increasing probability of a listener hearing /ɔ/ across these three conditions. Although the overall percent correct for /ɔ/ improves from natural speech to OF synthesis to FF synthesis, the probability of a correct response on trials in which /ɔ/ was used as a response declines from 94.1% for the natural signals to 91.1% for the OF synthetic signals to 83.7% for the FF synthetic signals.

In the analyses that follow, we will first consider the effect of flattening the formants, and then consider the differences in intelligibility between the natural signals and the OF synthetic signals.

### A. Effects of formant flattening

The drop in intelligibility that occurs as a result of flattening the formants is in general quite large. However, as Fig. 4 shows, the effect varies considerably from one vowel to the next. This was confirmed by a two-way repeated-measures ANOVA comparing just the OF and FF conditions, which showed significant effects for stimulus type [ $F(1,19) = 366.1, p < 0.001$ ] and vowel [ $F(11,209) = 48.4, p < 0.001$ ], and a significant interaction [ $F(11,209) = 44.0, p < 0.001$ ]. Vowels showing the largest changes in intelligibility as a result of formant flattening were /e/ (52.8%), /æ/ (31.4%), /ʊ/ (27.6%), /ʌ/ (22.6%), and /o/ (22.2%).

TABLE II. Confusion matrix for the original-formant (OF) synthetic signals. Values on the main diagonal, indicating the percentage of trials in which the listeners' responses matched the vowel intended by the talker, are shown in boldface. The response means given in the last row indicate the percentage of trials in which a given vowel was used as a listener response. Each vowel, as classified by the speaker's intention, was presented on 8.3% of the trials.

		Vowel identified by listener											
		/i/	/ɪ/	/e/	/ɛ/	/æ/	/ɑ/	/ɔ/	/o/	/ʊ/	/u/	/ʌ/	/ɜ/
Vowel intended by talker	/i/	<b>91.6</b>	4.6	3.6	0.2	...	...	...	...	...	...	...	...
	/ɪ/	0.6	<b>96.4</b>	0.2	2.4	0.4	...	...	...	...	...	...	...
	/e/	7.6	4.8	<b>85.4</b>	2.0	...	0.2	...	...	...	...	...	...
	/ɛ/	...	...	...	<b>92.8</b>	5.6	...	...	...	...	0.2	1.4	...
	/æ/	0.4	...	0.2	18.6	<b>80.8</b>	...	...	...	...	...	...	...
	/ɑ/	...	...	...	0.6	6.4	<b>82.8</b>	6.6	...	...	...	3.0	0.6
	/ɔ/	...	...	...	...	...	4.2	<b>94.2</b>	0.4	...	...	1.2	...
	/o/	...	...	...	...	...	0.4	...	<b>89.4</b>	5.2	4.2	0.8	...
	/ʊ/	...	...	...	...	...	0.2	...	1.0	<b>89.4</b>	2.0	7.0	0.4
	/u/	...	...	...	...	...	0.2	...	11.4	17.6	<b>70.4</b>	0.4	...
	/ʌ/	...	...	...	...	0.6	2.0	2.6	0.8	4.6	...	<b>89.4</b>	...
	/ɜ/	...	...	...	1.0	...	...	...	...	...	...	...	<b>99.0</b>
Response means:		8.4	8.8	7.5	9.8	7.8	7.5	8.6	8.6	9.7	6.4	8.6	8.3



TABLE III. Confusion matrix for the flat-formant (FF) synthetic signals. Values on the main diagonal, indicating the percentage of trials in which the listeners' responses matched the vowel intended by the talker, are shown in boldface. The response means given in the last row indicate the percentage of trials in which a given vowel was used as a listener response. Each vowel, as classified by the speaker's intention, was presented on 8.3% of the trials.

		Vowel identified by listener											
		/i/	/ɪ/	/e/	/ɛ/	/æ/	/ɑ/	/ɔ/	/o/	/u/	/ʊ/	/ʌ/	/ɜ/
Vowel intended by talker	/i/	<b>89.6</b>	6.6	2.8	0.8	...	0.2	...	...	...	...	...	...
	/ɪ/	4.4	<b>88.6</b>	2.8	3.6	0.6	...	...	...	...	...	...	...
	/e/	12.0	29.4	<b>32.6</b>	21.6	4.4	...	...	...	...	...	...	...
	/ɛ/	0.2	0.2	0.8	<b>87.6</b>	9.6	...	...	...	1.0	...	0.4	0.2
	/æ/	...	1.2	7.4	42.0	<b>49.4</b>	...	...	...	...	...	...	...
	/ɑ/	...	...	...	0.4	5.2	<b>85.2</b>	7.4	0.8	0.4	...	0.4	0.2
	/ɔ/	...	...	...	...	...	2.2	<b>96.0</b>	1.8	...	...	...	...
	/o/	...	...	...	0.2	...	3.2	12.2	<b>67.2</b>	9.0	7.0	1.2	...
	/u/	...	...	...	...	...	...	...	13.2	<b>61.8</b>	10.2	14.6	0.2
	/ʊ/	...	...	...	0.2	...	...	0.2	21.2	16.4	<b>61.0</b>	0.8	0.2
	/ʌ/	...	...	...	...	0.4	11.2	18.6	2.2	0.8	...	<b>66.8</b>	...
	/ɜ/	...	...	...	...	...	...	...	...	...	0.2	...	<b>99.8</b>
	Response means:		8.5	10.5	3.9	13.3	5.8	8.5	11.2	8.9	7.5	6.5	7.0

Table IV shows what kinds of changes in identification occurred. The analysis focused on instances in which a given listener identified the OF version of a signal correctly (i.e., as the vowel intended by the talker) but the FF version incorrectly. The symbols going down the rows indicate the vowel as classified both by the speaker's intention and the listener's labeling of the OF version; the columns show how the FF versions of these signals were identified. The last column shows the total number of changes in identification in which the OF version was heard as the intended vowel but the FF version was heard as some other vowel. The most frequently occurring changes in identification (shown in boldface in Table IV) involved /e/ shifting to /ɪ/ or /ɛ/, /æ/ shifting to /ɛ/, /u/ shifting to /ʌ/, /o/, or /ʊ/, /ʌ/ shifting to /ɑ/ or /ɔ/, and /o/ shifting to /ɔ/. Labeling changes involving these five vowels accounted for approximately three-quarters of all correct-to-incorrect vowel shifts. A similar analysis that included all shifts in vowel color between OF and FF signals, whether from correct to incorrect or otherwise, yielded a pattern of results that was quite similar to that shown in Table IV.

As might be expected, vowels that typically show relatively large amounts of spectral change tended to be more strongly affected by formant flattening. Figure 5 shows the relationship between the average magnitude of formant frequency change for each vowel and the total number of correct-to-incorrect changes in identification (i.e., the last column of Table IV). The magnitude of spectral change for each vowel category was represented as the average length of a vector connecting formant measurements sampled at 20% of vowel duration and 80% of vowel duration. The vector was drawn in a three-dimensional space consisting of log-transformed values of  $F_1$ ,  $F_2$ , and  $F_3$ . As Fig. 5 shows, the vowels tend to cluster into one group in the lower left showing relatively little spectral change and few changes in labeling, and a second group in the upper right showing a good deal of spectral change and many shifts in labeling. The relationship is far from perfect, however. For example, the perceptual effect of formant flattening for /e/ is quite large, even though the magnitude of formant frequency change is relatively modest in relation to the other vowels. We experi-

TABLE IV. Changes in phonetic labeling for signals whose original-formant versions were identified correctly but whose flat-formant versions were identified as a vowel other than that intended by the talker. The most frequently occurring vowel shifts are shown in boldface.

		FF synthetic vowel identified as												Total
		/i/	/ɪ/	/e/	/ɛ/	/æ/	/ɑ/	/ɔ/	/o/	/u/	/ʊ/	/ʌ/	/ɜ/	
OF vowel identified as	/i/	...	23	9	2	...	1	...	...	...	...	...	...	35
	/ɪ/	20	...	10	11	...	...	...	...	...	...	...	...	41
	/e/	37	<b>116</b>	...	<b>105</b>	22	...	...	...	...	...	...	...	<b>280</b>
	/ɛ/	1	1	4	...	38	...	...	...	5	...	2	1	52
	/æ/	...	2	28	<b>157</b>	...	...	...	...	...	...	...	...	<b>187</b>
	/ɑ/	...	...	...	...	6	...	26	3	1	...	...	...	36
	/ɔ/	...	...	...	...	...	11	...	7	...	...	...	...	18
	/o/	...	...	...	1	...	16	<b>55</b>	...	27	26	5	...	<b>130</b>
	/u/	...	...	...	...	...	...	...	<b>50</b>	...	<b>45</b>	<b>55</b>	1	<b>151</b>
	/ʊ/	...	...	...	1	...	...	1	<b>57</b>	<b>41</b>	...	3	1	104
	/ʌ/	...	...	...	...	2	<b>51</b>	<b>79</b>	8	2	...	...	...	<b>142</b>
	/ɜ/	...	...	...	...	...	...	...	...	...	1	...	...	1

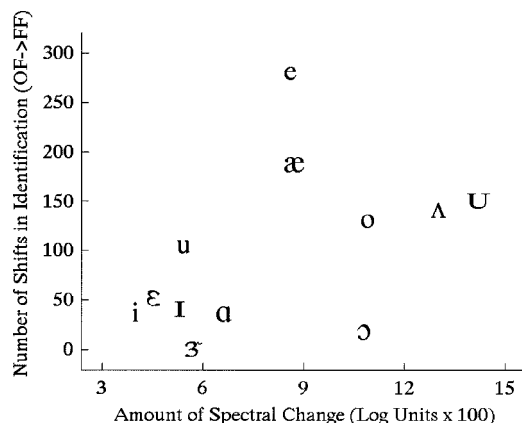


FIG. 5. Total number of correct-to-incorrect changes in identification as a function of the average magnitude of formant frequency change for each vowel. The magnitude of formant change for each vowel category was represented as the average length of a vector connecting formant measurements sampled at 20% of vowel duration and 80% of vowel duration. The vector was drawn in a three-dimensional space consisting of log-transformed values of  $F_1$ ,  $F_2$ , and  $F_3$ .

mented with sample points other than 20% and 80%, and with a number of alternate methods of representing spectral change magnitude, including calculations based on normalization schemes proposed by Syrdal (1985); Syrdal and Gopal (1986); Miller (1984, 1989) and Peterson (1951). Results of these additional analyses were very similar to the general pattern shown in Fig. 5.

### B. Natural signals versus OF synthesis

Although the main purpose of this study was to examine the effects of formant flattening, the difference in intelligibility that was observed between the naturally produced signals and the OF synthetic signals raises some important questions about the cues underlying the perception of vowel color. A two-way repeated-measures ANOVA comparing just the NAT and OF synthesis conditions showed significant effects for stimulus type [ $F(1,19)=371.8$ ,  $p<0.001$ ] and vowel [ $F(11,209)=10.6$ ,  $p<0.001$ ], and a significant interaction [ $F(11,209)=21.0$ ,  $p<0.001$ ]. Vowels showing the

largest changes in intelligibility as a result of formant coding were /u/ (25.0%), /æ/ (16.4%), /e/ (12.4%), and /a/ (9.4%).

Table V shows the distribution of responses for all instances in which a given listener identified the original signal correctly but identified the OF synthetic version as a vowel other than that intended by the talker. The most frequently occurring changes in identification involved /u/ shifting to /ʊ/ or /o/, /æ/ shifting to /ε/, and /e/ shifting to /i/. Labeling changes involving /u/ and /æ/ alone accounted for 37.8% of all correct-to-incorrect vowel shifts. These comparisons raise an obvious question: Why is there any difference in intelligibility between the natural signals and OF synthetic signals? In other words, what phonetically relevant information is not preserved by the formant frequency representation that drives the synthesizer during the vowel?

One possibility that cannot be ruled out is that the drop in intelligibility may occur at least in part as a result of errors in the estimation of formant frequencies. Measurement-remeasurement reliability for the LPC-derived formant frequency estimates is on the order of 1.0%–2.0% of formant frequency for  $F_1$  and 1.0%–1.5% of formant frequency for  $F_2$  and  $F_3$  (Hillenbrand *et al.*, 1995).<sup>4</sup> However, these reliability estimates do not address the validity question, and the possibility exists that LPC produces systematic errors in formant frequency measurement. For example, in a relatively small-scale study, Di Benedetto (1989) reported LPC-derived estimates of  $F_2$  and  $F_3$  that were very similar to those derived from smoothed wide-band Fourier spectra, but estimates of  $F_1$  that were systematically lower when measured with LPC. A larger and more formal comparison of 180 /hVd/ utterances by Hillenbrand *et al.* (1995) also found estimates of  $F_2$  and  $F_3$  that were similar between LPC and smoothed Fourier spectra; however, estimates of  $F_1$  were found to be approximately 40 Hz *higher* for LPC.

The main question that is raised by these measurement issues is whether systematic errors in formant frequency estimation might account for the shifts in vowel quality that were observed between the natural and the OF synthetic signals. For example, the many labeling shifts that occurred from /u/ to /ʊ/ and /æ/ to /ε/ might be explained by positing that estimates of  $F_1$  are systematically high. Figure 6 was

TABLE V. Changes in phonetic labeling for naturally produced signals that were identified correctly but whose original-formant synthetic versions were identified as a vowel other than that intended by the talker. The most frequently occurring vowel shifts are shown in boldface.

		OF synthetic vowel identified as												Total
		/i/	/ɪ/	/e/	/ε/	/æ/	/a/	/ɔ/	/o/	/ʊ/	/u/	/ʌ/	/ɜ/	
NAT vowel identified as	/i/	...	21	16	1	...	...	...	...	...	...	...	...	38
	/ɪ/	3	...	...	12	2	...	...	...	...	...	...	...	17
	/e/	<b>38</b>	21	...	7	...	1	...	...	...	...	...	...	<b>67</b>
	/ε/	...	...	...	...	16	...	...	...	...	1	7	...	24
	/æ/	2	...	1	<b>89</b>	...	...	...	...	...	...	...	...	<b>92</b>
	/a/	...	...	...	3	23	...	26	...	...	...	15	...	<b>67</b>
	/ɔ/	...	...	...	...	...	16	...	2	...	...	5	...	23
	/o/	...	...	...	...	...	2	...	...	22	17	2	...	43
	/ʊ/	...	...	...	...	...	1	...	5	...	10	27	...	45
	/u/	...	...	...	...	...	1	...	<b>54</b>	<b>80</b>	...	2	2	<b>137</b>
	/ʌ/	...	...	...	...	2	9	13	4	19	...	...	...	47
	/ɜ/	...	...	...	5	...	...	...	...	...	...	...	...	5

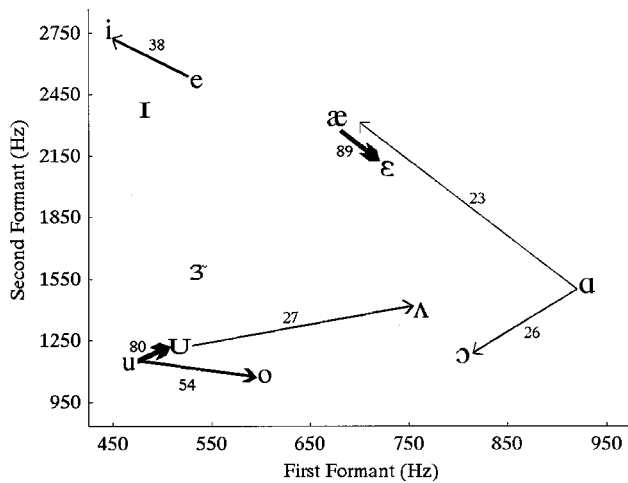


FIG. 6. This figure shows the most frequently occurring shifts in vowel color from NAT to OF; that is, the number of instances in which the naturally spoken version of an utterance was correctly identified but the original-formant synthetic version was identified as some other vowel. To provide a reference frame for formant space, phonetic symbols are plotted at the average values of  $F_1$  and  $F_2$  for the women. The tails of the arrows are plotted at the formant values for the correctly identified natural signal, and the arrow heads point at the phonetic label assigned to the OF synthetic version of the signal. The line thickness is roughly proportional to the number of labeling shifts.

designed to address this question. Plotted on this figure are the most frequently occurring shifts in vowel color from NAT to OF, based on the data in Table V. To provide a reference frame for formant space, phonetic symbols are plotted at the average steady-state values of  $F_1$  and  $F_2$  for the women. The tails of the arrows are plotted at the average formant values for the correctly identified natural signal, and the arrow heads point at the corresponding values for phonetic label assigned to the OF synthetic version of the signal. The line thickness (but *not* line length) is roughly proportional to the number of labeling shifts. The main point to be made about this figure is that no simple, systematic measurement error can account for the most common shifts in vowel quality. For example, while the shifts away from /u/ and /æ/ could conceivably be explained on the basis of systematically high estimates of  $F_1$  (i.e., the arrows point in the direction of vowels with higher first formants), those away from /a/ and /e/ are not consistent with this idea. This is not to suggest that formant measurement error does not play a role in accounting for the differences in intelligibility between the NAT and OF signals, but rather, that no simple, systematic difference in formant estimation seems capable of accounting for these differences.

One other possibility worth considering has to do with differences between the natural and synthetic signals during the /h/ and final /d/ intervals. The synthesizer was driven by acoustic measurements during the vowel only, with the /h/ and /d/ segments being generated by some simple rules. As a result, the initial /h/ and, in particular, the final /d/ segments did not always show a very close match between the original and synthetic utterances. It is possible that there is some limited information in the naturally produced consonants that influenced vowel quality. Alternatively, it may be that there was some information in the synthetic consonants that was

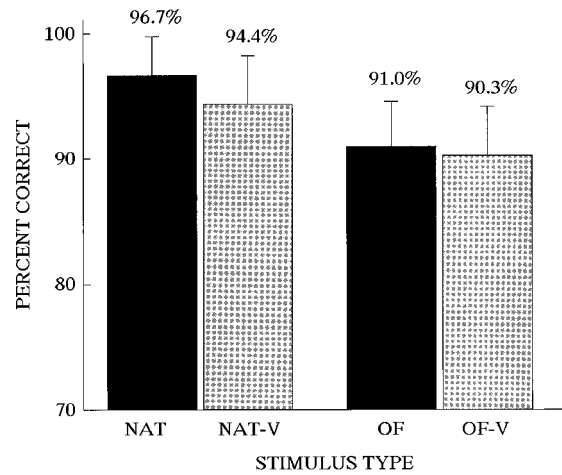


FIG. 7. Percent correct identification for four types of utterances: (a) naturally spoken /hVd/ signals (NAT), (b) the NAT signals with the initial and final consonants edited out (NAT-V), (c) original-formant synthetic /hVd/ signals (OF), and (d) the vowels only from the OF signals (OF-V). Error bars indicate 1 s.d. The percentages above each bar indicate the mean percent correct for each condition.

misleading to listeners in some way. Experiment 2 was designed to test this possibility.

### III. EXPERIMENT 2. METHODS

Experiment 2 presented listeners with four kinds of signals: (a) the 300 natural /hVd/ utterances (NAT); (b) the vowel only from the 300 NAT utterances (NAT-V); (c) the 300 original-formant synthetic /hVd/ utterances (OF); and (d) the vowel only from the 300 OF synthetic utterances (OF-V). The signals were edited from the NAT and OF utterances described above using a simple computer program that was controlled by the hand-measured values of vowel start and vowel end from Hillenbrand *et al.* (1995). After clipping the vowels from the /hVd/ utterances, the NAT-V and OF-V signals were ramped on and off with a 10-ms half-cosine function to prevent onset and offset transients. Listeners consisted of 24 undergraduate students who had taken an introductory phonetics course and had received basic instruction in the use of phonetic symbols for vowels. None of these listeners had participated in experiment 1. Listeners identified each of the 1200 test signals (300 NAT, 300 NAT-V, 300 OF, and 300 OF-V) presented in random order using the same instrumentation and procedures that were described for experiment 1.

### IV. EXPERIMENT 2. RESULTS AND DISCUSSION

Overall percent correct values for the four stimulus conditions of experiment 2 are shown in Fig. 7. The main point to be made about Fig. 7 is that both the natural and OF /hVd/ syllables were identified at a slightly higher rate than the corresponding vowel-only utterances. A repeated-measures two-way ANOVA showed significant effects for both factors (natural versus synthetic:  $F[1,24]=268.4, p<0.01$ ; syllable versus vowel:  $F[1,24]=64.8, p<0.01$ ). As can be seen in Fig. 7, the difference in intelligibility between the /hVd/ and vowel-only conditions is not large overall, and is very small for the OF synthetic stimuli. Newman-Keuls *post hoc* tests

showed that this difference was significant only for the naturally spoken stimuli. These results would seem to be consistent with the idea that there is some limited information in the naturally produced consonants that influences vowel quality. If the synthetic consonants were providing misleading information about vowel quality, the results should have shown an improvement in intelligibility for the OF synthetic signals with the removal of the consonants, and little or no change for the natural signals. The small drop in intelligibility that was observed suggests that some very limited information about vowel identity was lost when the natural consonants were clipped off. However, the absolute magnitude of the effect was quite small. The natural vowel-only stimuli remain highly intelligible, and the drop in intelligibility that results from excising the consonants amounts to an average of just 7 additional misidentified vowels out of the 300 that were presented. The primary conclusion from experiment 2, therefore, is that the failure to faithfully model the initial and final consonants can at best explain a very small portion of the difference in intelligibility between the natural and OF synthetic signals.

## V. PATTERN RECOGNITION MODELS

Two general conclusions seem likely from the foregoing. First, some changes in listeners' perception (most notably those between the natural and OF stimuli) cannot be readily accounted for by any of the acoustic properties controlled in the experiments. Second, despite this, variations in responses across stimuli are at least partly related to differences in spectral change that were manipulated. The modeling work presented below strives to provide a more detailed assessment of just how far we can go in relating response patterns to selected acoustic properties.

This modeling can be viewed as a way to extend the insights that we were seeking in Fig. 5. There, we measured how average identification rates for each vowel category improved from the FF to the OF condition and we attempted to relate that improvement to average spectral change. The Pearson correlation coefficient ( $r = 0.46$ ) between the  $x$  and  $y$  coordinates of Fig. 5 gives us a simple index of association between the two quantities. The directness of such an approach is very appealing, and it seems to provide some evidence for the hypothesis being tested. However, it is deficient in several respects. First, it fails to take into account variation among tokens of the same vowel category. Second, a very high correlation should result only under a very limiting assumption: namely, a unit increase in the magnitude of spectral change will result in a uniform change in identification rate, regardless of the direction of the change and of the overall location in formant space of the tokens involved. However, it is easy to imagine cases where this is most unlikely. For example, a token of a vowel whose overall position in formant space is relatively distant from those of its competitors in neighboring categories is likely to be less sensitive to differences in spectral change than a token that is closer to competing tokens.

Nearey and his colleagues have developed pattern recognition methods that can overcome these difficulties (Assmann *et al.*, 1982; Nearey and Assmann, 1986; Andruski and

Nearey, 1992; Nearey, 1997). These methods can simultaneously accommodate many aspects of the patterns, including magnitude and direction of spectral change and relative distance from competitors, on a token-by-token basis. We will adopt a variation of the general methodology developed by Nearey and Assmann (1986).

### A. Testing and training sets

Much of our modeling uses a two-stage procedure involving the distinction between disjoint training and testing data that is now prevalent in the speech recognition literature. First, a pattern recognition algorithm is constructed based on the statistical regularities in the training data. Second, the parameters of the recognition model are held fixed at the training values, while the algorithm is fed new acoustic data from a testing set that is independent of the training data. This generates predicted response patterns, which are then compared to listeners' performance on the test stimuli.

The training set consisted of 1297 tokens from the larger data of Hillenbrand *et al.* (1995) and will be referred to as the H95 data. The selected tokens included all those with no missing values for any of the measurements required, but excluded the 300 tokens that were used in experiments 1 and 2 above. There was also a corresponding set of responses for each token of the H95 training data. These responses were used for training the logistic regression coefficients in Model B, discussed below. The testing sets consist of the stimuli and responses reported in experiments 1 and 2 above.

### B. Stimulus properties and discriminant analysis

We have chosen a representation similar to that of the Canadian studies, using vowel duration, steady-state  $F_0$ , and  $F_1$ ,  $F_2$  and  $F_3$  (all frequencies were log-transformed) at the 20% and 80% time points. (For the flat formant stimuli, the 20% and 80% formant frequencies were equal to the measured steady-state frequencies.) Linear discriminant function analysis of the H95 training data showed that 92.0% of the tokens could be correctly reclassified using these measurements. When the coefficients estimated from the training data were applied to measurements from the natural or OF testing data (with distinct 20% and 80% formant measures), the results were actually higher, 94.0% correct. Thus the chosen measurements are capable of separating the vowel categories relatively well. But to what degree can a recognition algorithm characterize variation in listeners' response patterns?

### C. Predicting listener responses on the testing data

#### 1. Model A: A posteriori probabilities from discriminant function analysis of the H95 data

The linear classification functions described above are not only suitable for classification, but they can be used to generate *a posteriori* probabilities (APP scores) of group membership for any given measurement point. APP scores can be viewed as estimates of relative strength of group membership (Nearey and Assmann, 1986). For example, a token whose measurements are near the mean of /i/ and remote from the means of the other categories will have an

APP score for /i/ near unity while scores for the other vowels will be near zero. Similarly, an ambiguous token that is roughly equidistant from /i/, /e/, and /ɪ/ will have a score near 0.33 for each of these vowels, but near zero for the rest.

Such graded membership scores can be compared with confusion matrices, including confusion matrices constructed on a token-by-token basis. Such token-by-token confusion matrices will be referred to below as probability matrices. For each 300 token set, there is both a predicted and an observed probability matrix, each with 300 rows and 12 columns. Predictions based on the linear classification functions from the discriminant analysis of the H95 data will be referred to as Model A predictions.

## 2. Model B: Predicted probabilities from logistic regression of the H95 perceptual data

Logistic regression provides another technique for generating predicted probability matrices for the testing data (see Nearey, 1989, 1997, for applications of logistic regression to perceptual data). A 12-category polytomous logistic regression was performed using the training measurements as independent variables and the response matrices from the Hillenbrand *et al.* (1995) study as the dependent measures. This method may result in better correspondence to listeners' behavior because, in effect, it can model ambiguity as well as identity. It does this by optimally matching the gradient, probabilistic identification profiles of a group of listeners to each token of the training set, rather than simply predicting nominal correct categories.

## D. Measures of association between probability matrices

We will compare predicted and observed probability matrices using methods similar to those of Nearey and Assmann (1986).<sup>5</sup> Three measures of association will be reported. The first is percentage of modal agreement ( $P_{ma}$ ), defined as the percentage of tokens for which the predicted and observed probabilities show the same modal category, where the modal category is the response with the highest probability for that token. (See the Appendix for a formal definition of  $P_{ma}$  and other measures of association used in the pattern recognition work.) Note that this measure does not depend on the nominally correct category of the original recordings. For example, both the listeners and the prediction model might agree that a flat formant token from an original /e/ more closely resembles the /ɪ/ category. These results are shown in Table VI.

The second measure of association, called *correct response correlation* ( $r_c$ ), is defined as the correlation between predicted and observed probabilities of nominally correct responses to each stimulus (ignoring all incorrect responses). The value of  $r_c$  will approach a maximum of 1.0 if and only if *variation* in the relative probabilities of correct identification by listeners is matched by *covariation* in predicted probabilities on a token-by-token basis. These correlations are shown in Table VII. Although these are conventional correlation coefficients from a computational standpoint, it is not clear whether the usual statistical assumptions apply. Therefore, following Nearey and Assmann (1986), we will use

TABLE VI. Percent modal agreement values ( $P_{ma}$ ) for each of the stimulus sets used in experiment 1. Columns represent different prediction models. The last column represents an empirical split-sample cross-validation estimate, predicting one-half of the subjects' responses from the other half.

Experiment	Stimulus type	Model A	Model B	Model C	Model C-s	Split-Sample
1	NAT	94.0	97.0	96.3	96.1	99.5
1	FF	71.3	75.3	88.3	87.9	92.5
1	OF	91.7	92.7	94.3	94.2	96.4
2	NAT	94.0	97.0	96.3	96.3	100.0
2	OF	93.7	94.7	96.7	95.6	97.6
2	NAT-V	94.3	96.7	96.0	96.2	100.0
2	OF-V	95.0	95.7	97.0	96.4	97.7

nonparametric randomization tests (Edgington, 1980) to assess significance levels of the correlation coefficients.

## E. Difference correlations

The third measure of association focuses on the ability of the models to predict changes in listener behavior across stimulus conditions. Corresponding stimuli in all the changing formant (i.e., natural and original formant synthetic) stimuli must have exactly the same predicted probabilities because the measurement vectors supplied to the prediction algorithm are identical. However, the measurements for the flat formant synthesis tokens are different, since the formant frequencies are from the steady-state portion. If a pattern recognition algorithm approximates the behavior of our listeners, we would expect *changes* in predicted probabilities of a given token across conditions to be correlated with changes in listeners' responses. Following Nearey and Assmann (1986), we have calculated correlations between changes in predicted probabilities and corresponding changes in observed probabilities. This is done by producing six difference matrices, one for each of the six changing formant conditions (the NAT and OF conditions from experiment 1 plus all four conditions from experiment 2). Each is calculated as the element-by-element difference between the probability matrix of the given changing formant condition and that of the FF condition.

Our analysis here will concentrate on the *correct response difference correlation* ( $r_{cd}$ ). The calculation is analo-

TABLE VII. Correct response correlations ( $r_c$ ) for the same analyses as Table VI. Significance levels (by randomization test) shown for completely cross-validated predictions of models A and B only.

Experiment	Stimulus type	Model A	Model B	Model C	Model C-s	Split-Sample
1	NAT	0.207 <sup>b</sup>	0.370 <sup>c</sup>	0.400	0.345	0.585
1	FF	0.484 <sup>c</sup>	0.547 <sup>c</sup>	0.850	0.814	0.890
1	OF	0.399 <sup>c</sup>	0.387 <sup>c</sup>	0.710	0.660	0.814
2	NAT	0.094	0.305 <sup>c</sup>	0.249	0.197	0.434
2	OF	0.408 <sup>c</sup>	0.426 <sup>c</sup>	0.659	0.626	0.812
2	NAT-V	0.220 <sup>c</sup>	0.372 <sup>c</sup>	0.371	0.326	0.561
2	OF-V	0.480 <sup>c</sup>	0.504 <sup>c</sup>	0.664	0.605	0.751

<sup>a</sup> $p < 0.01$ .

<sup>b</sup> $p < 0.005$ .

<sup>c</sup> $p < 0.001$ .

TABLE VIII. Difference correct response correlations ( $r_{cd}$ ) of predicted with observed correct response difference scores when correct response for the flat formant data is subtracted from each of the corresponding tables. Significance levels (by randomization test) shown for completely cross-validated predictions of models A and B only.

Experiment	Stimulus type	Model A	Model B	Model C	Model C-s	Split-Sample
1	NAT	0.364 <sup>a</sup>	0.494 <sup>a</sup>	0.703	0.669	0.857
1	FF	...	...	...	...	...
1	OF	0.463 <sup>a</sup>	0.541 <sup>a</sup>	0.753	0.699	0.788
2	NAT	0.382 <sup>a</sup>	0.486 <sup>a</sup>	0.694	0.661	0.873
2	OF	0.450 <sup>a</sup>	0.554 <sup>a</sup>	0.772	0.734	0.835
2	NAT-V	0.356 <sup>a</sup>	0.472 <sup>a</sup>	0.701	0.667	0.853
2	OF-V	0.395 <sup>a</sup>	0.497 <sup>a</sup>	0.738	0.698	0.821

<sup>a</sup> $p < 0.001$ .

gous to the correct response correlation, except that difference matrices are substituted for the original probability matrices. This correlation will be large only when predicted and observed correct identification rates change in similar ways across conditions. Correct response difference correlations are shown in Table VIII.

## F. Benchmark split-sample predictions

As in Andruski and Nearey (1992), we have included benchmark measures of association based on the degree of agreement between subgroups of listeners. This was done by half-sample cross validation. For each experimental condition, subjects were split randomly into two groups. An observed probability matrix was calculated from the responses of approximately half the listeners (10 of 20 for experiment 1, 13 of 25 for experiment 2). This was used to provide nonparametric predictions for the entries of a similar matrix compiled from the remaining data. Measures of association from 200 different random splittings were averaged. These results are presented in the last column of Tables VI–VIII. These figures give a rough estimate of the degree of similarity of empirical response tables when the experiment is repeated with different listeners.

## VI. DISCUSSION

### A. Changing formant conditions

Consider first the results for the changing formants conditions, i.e., all cases but FF. In Table VI, we see that model A shows modal agreement ranging from 91% to 95% in these conditions. Model B, which had access to gradient aspects of listeners' categorization of the training stimuli, shows even higher agreement (about 93%–97%). A comparison with split-sample benchmark in the last column of Table VI shows that there is still room for improvement: Listeners are somewhat more consistent with each other (modal agreements range from 96% to 100%) than they are with our models.

Much of the similarity of models A and B for all of the non-FF conditions can be attributed to the simple fact that, for the both the listeners and the models, the modal category is the nominally correct category for most of the stimuli. The nominally correct identification rate is 94% for model A and

97% for model B for each of the changing formant conditions. The correct identification rate for listeners varies from about 95% to 100% across conditions (where the “winning” category is the one with the plurality of listener votes). However, correct response correlations  $r_c$  in Table VII show that more than this overall correspondence of correct responses is involved. Recall that  $r_c$  is positive only to the extent that variations in the probability of nominally correct responses covary in predicted and observed tables. Therefore, simply having high average probabilities of correct responses in both observed and predicted matrices will not result in positive correlations. The correlations for model A and model B are all positive and significant for all of the changing-formant conditions. Although the magnitudes of such correlations are modest, we should bear in mind that a ceiling on this correlation is imposed by listener-to-listener variability. An estimate of this ceiling is given in the split-sample column. For the changing-formant cases, the variance accounted for by model A (calculated as the ratio of the squares of the correlation coefficients) is roughly one-third and that by model B is roughly one-half that accounted for in the corresponding split-sample benchmarks.

### B. Flat formant condition

In the case of the FF stimuli, neither the *a priori* models (A and B) nor listeners' identifications show nominally correct identification rates nearly as high as in the changing-formant cases. The nominally correct category was chosen by the plurality of listeners in only about 80% of the tokens. Perhaps not surprisingly, the corresponding nominally correct identification rate for model A is considerably lower, only about 60%. Nonetheless, the modal agreement between the two is about 71%. Modal agreement with listeners is higher than the algorithm's correct identification rate because model A predictions showed the same “modal error” as listeners in 32 of the 59 tokens nominally misidentified by the plurality of listeners. Model B shows a nominally correct classification rate of about 73%, which is still somewhat lower than listeners. Again, the modal agreement between listeners and model B is higher (about 75%) because model B has also predicted the listeners' “modal errors” correctly in 29 of 59 cases. (This agreement on modal errors is slightly less than with model A. The improvement of model B over model A occurs because model B predicts 197 of the 241 correct responses by listeners, while model A correctly predicts only 182 of them.)

While the above results clearly suggest a reasonable degree of correspondence, we also see that listeners are much more consistent with each other than they are with the models. Although the rates of split-sample modal agreement are lower than they were for any of the changing formant conditions, at about 93% they are still more than 20 percentage points higher than the model A results for the FF stimuli.

The general pattern of the modal agreement results is also supported by the correlations between observed and predicted identification rates for nominally correct tokens  $r_c$ , given in Table VII. Both model A and B show highly significant correlations. However, the magnitudes of the corre-

lations ( $r_c = 0.48$  and  $0.54$ ) account for only about one-third of the variance accounted for in the split-sample predictions ( $r_c = 0.89$ ).

While the above models give a reasonable estimation of the overall fit of the predictions, they give only a very indirect view of the relative success of predicting *changes* in categorization *across* conditions. For this, we turn to the correct response difference correlations,  $r_{cd}$ , in Table VIII. Recall that this measure involves correlations of the changes (from the FF condition to each of the changing formant conditions) in observed probabilities of correct response with corresponding changes in predicted probabilities. Models A and B both show highly significant correlations for changes in response patterns from the FF condition to each of the other (changing formant) conditions. Model A accounts for only about 17%–35% of the variance accounted for by the split-sample benchmark. Model B fares somewhat better, accounting for about 30%–47% as much as the benchmark. The analysis underlying  $r_{cd}$  values is similar in spirit to that of Fig. 5. There are two main differences. First, rather than looking at difference in magnitude of formant change,  $r_{cd}$  involves changes in *a posteriori* probabilities (which, for linear discriminant analysis, are closely related to changes in “relative statistical distance” to category prototypes) between the two conditions. Second,  $r_{cd}$  values are calculated on a token-by-token basis, while Fig. 5 involved averaging over vowel categories.<sup>6</sup> If a similar averaging is done over changes in *a posteriori* probabilities, correlations across vowels are considerably higher for both model A ( $r = 0.60$ ) and model B ( $r = 0.71$ ) than for the spectral distance measure of Fig. 5 ( $r = 0.46$ ).

### C. Model C. Predicted probabilities from the experiment 1 perceptual data

Despite the significance of the results reported above, the modest size of the goodness of fit measures for the FF data relative to the split-sample benchmark must give us some pause. However, it should be kept in mind that models A and B were trained only on citation form tokens that must certainly show less variability than the overall population of tokens (produced, e.g., at various speaking rates and stress conditions) to which listeners are exposed. Thus even model B, which was trained on the rather limited degree of gradient behavior in listeners’ categorization of the H95 training data, might easily have “wrapped itself around” a solution that was dominated by listeners’ behavior to relatively prototypical stimuli. It is perhaps not surprising that such predictions might be rather fragile and that they breakdown somewhat when applied to the FF stimuli, which can present rather different stimulus patterns than those in the H95 data for many vowel categories.

We therefore constructed a third, optimized model, model C. Unlike the other two models which are based on the distinct H95 data set, the predictions here are based on observed response probabilities for the 900 stimuli of experiment 1. Model C prediction results are also shown in Tables VI–VIII. This model is included to see how well a model of the same “size” as model B might do with the stimuli in experiments 1 and 2. Since it is not a fully cross-validated

model, it is likely to present an overly optimistic picture (Efron and Tibshirani, 1993) of prediction errors for new data. However, to the extent that performance of even this model falls below our split-sample benchmarks, we will know that the shortfall is not due simply to the restricted nature of the training data and we will have a useful estimate of the lower bound on how much remains to be explained.

In addition, limited cross validation to distinct listeners and distinct tokens is possible with the available data if we reverse the roles of the training and testing data from those of models A and B. That is, we use logistic regression coefficients “frozen” at the values estimated by model C to predict listeners’ behavior on the much larger set of natural tokens of the H95 data. This analysis yields a modal agreement of 90.5% with H95 listeners, an  $r_c$  correlation of 0.453. If model C is used to classify the H95 training tokens, we find cross-validated classification rate of 90.5%. (This is the same as modal agreement with listeners, because the plurality of listeners’ responses in the H95 actually selects the nominally correct category for all stimuli.) This is rather remarkable, given that self-trained linear discriminant analysis on the same data yielded 92.0%. Recall that in model C we are training on measurements based on only 300 different vowels. Those measurements, when optimally mapped to listeners’ responses in the three presentation conditions (OF, FF, and NAT), are capable of classifying a completely distinct set of nearly 1300 vowels almost as well as linear discriminant analysis trained on the larger data set itself.

Although we are not able here to provide cross validation to entirely new stimulus tokens in all the conditions of experiments 1 and 2, we can provide true cross validation across different listeners for all conditions of experiment 2 (the last four rows in the model C column of Tables VI–VIII) and we can also provide split-sample cross validation even in the case of experiment 1, by training on the data from one-half of the listeners and testing predictions against the other half (model C-s). In the remaining discussion, we will use the predictions of model C-s, since the measures presented should provide unbiased estimates of prediction success for the same set of stimuli across new groups of listeners. (This is actually the only generalization that the empirical split-sample benchmarks also address. There is no statistical basis for generalizing those results to new stimuli.)

We find that modal agreement numbers for model C-s (Table VI) are uniformly very high, although they are smaller than the split-sample benchmarks by about 1–4 percentage points. Correct response correlations,  $r_c$ , in Table VII, are generally within about 0.1 of the corresponding model A and B values, but are higher by about 0.2 to about 0.3 for the FF and the two OF conditions. However, the variance accounted for by model C-s still averages only about half that of the split-sample benchmark, ranging from about 0.21 to 0.83, with the highest value for the FF condition. Correct response difference correlations  $r_{cd}$ , in Table VIII, show somewhat more improvement. Model C-s shows values about 0.16–0.31 higher than corresponding entries for models A and B, accounting for about 57%–79% of the variance accounted for by the split-sample benchmark.

## VII. GENERAL DISCUSSION

The main purpose of this study was to measure the relative importance of formant frequency movements in the recognition of vowel quality. The substantial difference in overall intelligibility between the OF and FF signals strongly confirms several previous findings indicating that spectral change patterns play a secondary but quite important role in the recognition of vowel quality. The vowels of American and Canadian English (and almost certainly many other vowel systems as well—see Watson and Harrington, 1999, for recent data on Australian English) are more properly modeled not as points in formant space but as trajectories through formant space.<sup>7</sup> However, a simple observation that should not be lost in this discussion of spectral change is that the single-slice spectral measurements reported in studies such as Peterson and Barney (1952) capture most of the information that is needed to represent vowel quality. In the present study,  $F_0$ , duration, and steady-state formant measurements were sufficient to signal the intended vowel for roughly three-fourths of the utterances, with nearly all of the misidentifications involving adjacent vowel categories. Strikingly similar identification rates for vowels with static formant patterns were reported by Fairbanks and Grubb (1961), Assmann and Nearey (1986), and Hillenbrand and Gayvert (1993a) in studies using methods that are quite different from those employed here.

The relative importance of formant frequency change varies considerably from one vowel to the next. It was generally the case that the effect of formant flattening was small for vowels that tend to show relatively little formant frequency movement and larger for vowels that tend to show large changes in formant frequencies. The relationship is not quite that simple, however, as demonstrated by the very different effects of formant flattening for /e/, /æ/, and /ɔ/, which showed roughly similar average magnitudes of formant frequency change.

A significant limitation of this study is the exclusive use of the simple /hVd/ environment for all utterances. The relationships between spectral change patterns and vowel identity are guaranteed to be more complex when the consonant environment preceding and following the vowel is allowed to vary. We are currently studying the acoustics and perception of a new multitalker CVC database with variation in both consonants. Preliminary analysis of this database (Hillenbrand and Clark, 1997) using a statistical pattern classifier shows substantially better classification accuracy for two samples of the formant pattern rather than a single sample, in spite of the complexities introduced by variation in consonant environment. The classification advantage for the two-sample case, however, was smaller than we observed in a similar discriminant analysis of our /hVd/ database (Hillenbrand *et al.*, 1995).

The pattern recognition methods described above use a relatively simple approach to vowel specification based on duration, steady-state  $F_0$ , and formant frequency measurements at two temporal intervals in the vowel. Such a representation appears to go a considerable distance toward accounting for the results of experiments 1 and 2 since significant correlations were found between various aspects

of listeners' behavior and the purely *a priori* predictions of models A and B. However, the agreement between listeners measured by split-sample benchmarks is typically better by a factor of about 2 than even our best predictions. This suggests that we have gone no more than about half the distance to the goal of accounting for listeners' behavior in these experiments.

In our view, the most significant challenge presented by the findings reported here is to explain the difference in intelligibility between the natural signals and the OF synthetic signals. The OF signals were, of course, highly intelligible, indicating that most of the information that is needed to capture vowel identity is preserved by the  $F_0$ , duration, and formant measurements that were used to drive the formant synthesizer. But the drop in intelligibility resulting from formant vocoding makes it equally clear that a certain amount of phonetically relevant information was lost.<sup>8</sup> Two possible explanations for this finding were pursued here. First, an analysis of the shifts in vowel identity between the NAT and OF signals suggested that it was unlikely that any simple, systematic error in formant measurement could account for the most common shifts in vowel identity. It was also shown that the failure to faithfully copy the initial and final consonants can at best explain a very small share of this effect. One plausible explanation that was not pursued in this study is that the transformation to a formant representation results in the failure to preserve spectral shape details that are relevant to vowel identity. The formant synthesizer is driven entirely by spectral-peak frequencies, meaning that formant amplitudes, bandwidths, spectral tilt, individual harmonic amplitudes, and other spectral details will often not match well between the natural and synthetic utterances. There has been a fair amount of discussion about the relative contributions of formant frequencies and detailed spectral shape (e.g., Klatt, 1982a,b; Bladon and Lindblom, 1981; Bladon, 1982; Zahorian and Jagharghi, 1986, 1987; Zahorian and Zhang, 1992), but the question is far from resolved. Work that is currently underway involves a close examination of the signals from the present study that showed significant shifts in labeling between the natural and OF formant-synthesis conditions in an effort to understand what specific spectral shape details might play a role in judgments of vowel identity.

## ACKNOWLEDGMENTS

This work was supported by a grant from the National Institutes of Health (2-R01-DC01661) to Western Michigan University and by a grant from the Canadian Social Sciences and Humanities Research Council (410-93-0053) to the University of Alberta. We are grateful to Michael Clark, Robert Erickson, and two anonymous reviewers for helpful comments on an earlier draft.

## APPENDIX. FORMAL DEFINITIONS OF QUANTITIES USED IN PATTERN RECOGNITION STUDIES

The observed probability matrix has elements  $O_s(t,v)$ , where the subscript  $s$  represents a given stimulus condition (e.g., NAT, FF, etc.).  $t$  ranges over the 300 tokens and  $v$  over the 12 vowel response categories. Assume the order of the



vowels /i,ɪ,e,ɛ,æ,a,ɔ,o,u,ʌ,ɜ/. Thus,  $O_s(4,3)$  represents the proportion of listeners who responded to stimulus number 4 with vowel category 3. The predicted probability matrix with elements  $P_s(t,v)$  is similar in structure, but contains a *posteriori* probability estimates from the pattern recognition models. These matrices are used for the definitions of all other quantities used in the pattern recognition studies.

Percentage modal agreement for a given stimulus condition can be defined as:

$$100\sum_t \{M[\text{argmax}_v(O_s(t,v), \text{argmax}_v(P_s(t,v))]/T);$$

where  $M(x,y) = 1$ , if  $x=y$  and 0 otherwise,  $\text{argmax}_v$  indicates the column index of the largest element in a row and  $T$  is the total number of tokens (300 in these experiments).

For correct response correlations, define the observed correct probability score for token  $t$  as  $C_s(t) = O_s(t, c_t)$ , where  $c_t$  is the column number of the correct response for token  $t$  of the observed probability matrix. Thus if stimulus 4 corresponded to the vowel /e/ (vowel number 3) in the original recordings ( $s = \text{NAT}$ ), then  $C_s(4)$  is  $O_s(4,3)$ . Predicted correct probabilities are similarly defined  $D_s(t) = P_s(t, c_t)$ .

Correct response correlations are then defined as the Pearson correlation between the elements of  $C_s(t)$  and the corresponding elements of  $D_s(t)$ . Randomization tests are calculated by randomizing the token index for the predicted tokens. Correct response difference correlations,  $r_{cd}$ , are based on changes in correct identification probabilities between all moving formant conditions and the FF conditions. Define correct response difference scores as  $Y_s(t) = C_s(t) - C_{\text{FF}}(t)$  for observed probabilities and as  $X_s(t) = D_s(t) - D_{\text{FF}}(t)$  for predicted, where the subscript FF refers to the fixed formant condition. Correct response difference correlations for each condition  $s$  (except FF, for which all scores are by definition zero) are then the Pearson correlations between  $Y_s(t)$  and  $X_s(t)$ .

<sup>1</sup>Note that the formant contour for the /æ/ in Fig. 2 shows two relatively stationary segments. While this sort of pattern was not very common overall, it did occur with some regularity for this vowel. There was nothing in the formal procedure for judging steady-state times that would have prevented the research assistants who made these visually based judgments from choosing a frame in the offglide as the steadiest point in the vowel. Inspection of these measurements by the first author showed that this never occurred.

<sup>2</sup>Note that the final /d/ segments of the OF and FF signals matched the original signals fairly closely for overall segment duration, but the original and synthetic signals did not necessarily match with respect to the relative durations of the closure and burst intervals.

<sup>3</sup>Speaker-within-group could also be considered a random effect, although this is not typically done in phonetics experiments. Furthermore, the sampling procedure used did not allow for systematic testing effects as such. Strictly speaking, we have no statistical basis for generalizing to stimuli beyond those used in the experiments. Furthermore, statements about subject and vowel effects are to some degree confounded by individual speaker differences. However, our main focus is on differences among conditions, and we are using ANOVA primarily as a screening tool to draw our attention to possible patterns in the data that are correlated with vowel and speaker group.

<sup>4</sup>As a further check on the stimulus generation process, several dozen of the tokens showing the largest changes in identification from the NAT to OF conditions in experiment 1 were reanalyzed at the University of Alberta using software distinct from that used in Hillenbrand *et al.* (1995). This reanalysis showed that the retracted formant frequencies from the resynthesized stimuli agreed very closely with those of the original signals. Thus, although other aspects of the stimuli (such as formant amplitudes and band-

widths) do vary, we are satisfied that the frequencies of  $F_1$  through  $F_3$  match quite well.

<sup>5</sup>The analyses reported here focus on the probabilities of a single response (either the correct original category or the modal response) in each stimulus row of a probability matrix. We have also performed profile correlation and difference profile correlation analyses similar to those reported in Nearey and Assmann (1986) and Andruski and Nearey (1992) that use all entries of the probability matrices. These analyses revealed patterns generally similar to those reported here.

<sup>6</sup>There is also a third, minor difference. Figure 5 counted the number of tokens whose correct identification changed from OF to FF conditions, regardless of direction, while  $r_{cd}$  measures signed changes in probability of correct identification across tokens. A very similar correlation results if the latter difference in response measure is substituted in the analysis of Fig. 5.

<sup>7</sup>An anonymous reviewer suggested that listeners might, "... base their decisions not on the formant frequencies at the 20%, the 80% or the two combined but on the modal formant frequencies of the vowel production; i.e., the  $F_1, F_2, F_3$  combination that occurs the most often ..." for the nominally monophthongal vowels. We believe that this is an unlikely possibility. The clearest evidence, in our view, comes from the gating experiments of Nearey and Assmann (1986), showing excellent identification of silent-center vowels (including the nominal monophthongs /ɪ/, /ɛ/, and /æ/) consisting of brief onsets and offsets, but poor labeling of the same segments played in reverse order.

<sup>8</sup>It is worth noting that the method that we used to track changes in vowel color from the natural signals to the OF synthetic versions relies entirely on changes in absolute identification. This method is rather coarse. There were clearly many utterances in which the vowel color appeared to us to change from NAT to OF, but the change was not sufficient to induce a labeling shift for most of the listeners. Conversely, in listening to the natural and OF versions of signals showing a large number of labeling shifts, we have generally been impressed at the subtlety of the change in vowel color.

Ainsworth, W. A. (1972). "Duration as a cue in the recognition of synthetic vowels," *J. Acoust. Soc. Am.* **51**, 648–651.

Andruski, J., and Nearey, T. M. (1992). "On the sufficiency of compound target specification of isolated vowels and vowels in /bVb/ syllables," *J. Acoust. Soc. Am.* **92**, 390–410.

Assmann, P., Nearey, T., and Hogan, J. (1982). "Vowel identification: orthographic, perceptual, and acoustic aspects," *J. Acoust. Soc. Am.* **71**, 975–989.

Bennett, D. C. (1968). "Spectral form and duration as cues in the recognition of English and German vowels," *Lang. & Speech* **11**, 65–85.

Black, J. W. (1949). "Natural frequency, duration, and intensity of vowels in reading," *J. Speech Hear. Dis.* **14**, 216–221.

Bladon, A. (1982). "Arguments against formants in the auditory representation of speech," in *The Representation of Speech in the Peripheral Auditory System*, edited by R. Carlson and B. Granstrom (Elsevier Biomedical, Amsterdam), pp. 95–102.

Bladon, A., and Lindblom, B. (1981). "Modeling the judgment of vowel quality differences," *J. Acoust. Soc. Am.* **69**, 1414–1422.

Carlson, R., Fant, G., and Granstrom, B. G. (1975). "Two-formant models, pitch, and vowel perception," in *Auditory Analysis and Perception of Speech*, edited by G. Fant and M. A. A. Tatham (Academic, London), pp. 55–82.

Di Benedetto, M-G. (1989). "Vowel representation: Some observations on temporal and spectral properties of the first formant frequency," *J. Acoust. Soc. Am.* **86**, 55–66.

Edgington, E. (1980). *Randomization Tests* (Dekker, New York).

Efron, B., and Tibshirani, R. (1993). *Introduction to the Bootstrap* (Chapman and Hall, London).

Fairbanks, G., and Grubb, P. (1961). "A psychophysical investigation of vowel formants," *J. Speech Hear. Res.* **4**, 203–219.

Hillenbrand, J. M., and Clark, M. J. (1997). "Effects of consonant environment on vowel formant patterns," *J. Acoust. Soc. Am.* **102**, 3093(A).

Hillenbrand, J. M., and Gayvert, R. T. (1993a). "Vowel classification based on fundamental frequency and formant frequencies," *J. Speech Hear. Res.* **36**, 674–700.

Hillenbrand, J. M., and Gayvert, R. T. (1993b). "Identification of steady-state vowels synthesized from the Peterson and Barney measurements," *J. Acoust. Soc. Am.* **94**, 668–674.

Hillenbrand, J. M., Getty, L. A., Clark, M. J., and Wheeler, K. (1995). "Acoustic characteristics of American English vowels," *J. Acoust. Soc. Am.* **97**, 3099–3111.

- Jenkins, J. J., Strange, W., and Edman, T. R. (1983). "Identification of vowels in 'vowelless' syllables," *Percept. Psychophys.* **34**, 441–450.
- Klatt, D. H. (1982a). "Prediction of perceived phonetic distance from critical-band spectra: A first step," *IEEE ICASSP*, 1278–1281.
- Klatt, D. H. (1982b). "Speech processing strategies based on auditory models," in *The Representation of Speech in the Peripheral Auditory System*, edited by R. Carlson and B. Granstrom (Elsevier Biomedical, Amsterdam), pp. 181–196.
- Klatt, D. H., and Klatt, L. C. (1990). "Analysis, synthesis, and perception of voice quality variations among female and male talkers," *J. Acoust. Soc. Am.* **87**, 820–857.
- Miller, J. D. (1984). "Auditory processing of the acoustic patterns of speech," *Arch. Otolaryngol.* **110**, 154–159.
- Miller, J. D. (1989). "Auditory-perceptual interpretation of the vowel," *J. Acoust. Soc. Am.* **85**, 2114–2134.
- Nearey, T. M. (1989). "Static, dynamic, and relational properties in vowel perception," *J. Acoust. Soc. Am.* **85**, 2088–2113.
- Nearey, T. M. (1992). "Applications of generalized linear modeling to vowel data," in *Proceedings of ICSLP 92*, edited by J. Ohala, T. Nearey, B. Derwing, M. Hodge, and G. Wiebe (University of Alberta, Edmonton, AB), pp. 583–586.
- Nearey, T. M. (1997). "Speech perception as pattern recognition," *J. Acoust. Soc. Am.* **101**, 3241–3254.
- Nearey, T. M., and Assmann, P. (1986). "Modeling the role of vowel inherent spectral change in vowel identification," *J. Acoust. Soc. Am.* **80**, 1297–1308.
- Nearey, T. M., Hogan, J., and Rozsypal, A. (1979). "Speech signals, cues and features," in *Perspectives in Experimental Linguistics*, edited by G. Prideaux (Benjamin, Amsterdam).
- Parker, E. M., and Diehl, R. L. (1984). "Identifying vowels in CVC syllables: Effects of inserting silence and noise," *Percept. Psychophys.* **36**, 369–380.
- Peterson, G., and Barney, H. L. (1952). "Control methods used in a study of the vowels," *J. Acoust. Soc. Am.* **24**, 175–184.
- Peterson, G., and Lehiste, I. (1960). "Duration of syllable nuclei in English," *J. Acoust. Soc. Am.* **32**, 693–703.
- Peterson, G. E. (1951). "The phonetic value of vowels," *Language* **27**, 541–553.
- Potter, R. K., and Steinberg, J. C. (1950). "Toward the specification of speech," *J. Acoust. Soc. Am.* **22**, 807–820.
- Rabiner, L. (1968). "Digital formant synthesizer for speech synthesis studies," *J. Acoust. Soc. Am.* **24**, 175–184.
- Shankweiler, D., Strange, W., and Verbrugge, R. (1977). "Speech and the problem of perceptual constancy," in *Perceiving, Acting, and Comprehending: Toward an Ecological Psychology*, edited by R. Shaw and J. Bransford (Lawrence Erlbaum, Hillsdale, NJ).
- Stevens (1959). "The role of duration in vowel identification," *Quarterly Progress Report 52*, Research Laboratory of Electronics, MIT.
- Stevens, K. N., and House, A. S. (1963). "Perturbation of vowel articulations by consonantal context: An acoustical study," *J. Speech Hear. Res.* **6**, 111–128.
- Strange, W., Jenkins, J. J., and Johnson, T. L. (1983). "Dynamic specification of coarticulated vowels," *J. Acoust. Soc. Am.* **74**, 695–705.
- Syrdal, A. K. (1985). "Aspects of a model of the auditory representation of American English vowels," *Speech Commun.* **4**, 121–135.
- Syrdal, A. K., and Gopal, H. S. (1986). "A perceptual model of vowel recognition based on the auditory representation of American English vowels," *J. Acoust. Soc. Am.* **79**, 1086–1100.
- Tiffany, W. (1953). "Vowel recognition as a function of duration, frequency modulation and phonetic context," *J. Speech Hear. Dis.* **18**, 289–301.
- Watson, C., and Harrington, J. (1999). "Acoustic evidence of dynamic formant trajectories in Australian English vowels," *J. Acoust. Soc. Am.* (in press).
- Zahorian, S., and Jagharghi, A. (1986). "Matching of 'physical' and 'perceptual' spaces for vowels," *J. Acoust. Soc. Am. Suppl.* **1 79**, S8.
- Zahorian, S., and Jagharghi, A. (1987). "Speaker-independent vowel recognition based on overall spectral shape versus formants," *J. Acoust. Soc. Am. Suppl.* **1 82**, S37.
- Zahorian, S., and Zhang, Z.-J. (1992). "Perception of vowels synthesized from sinusoids that preserve either formant frequencies or global spectral shape," *J. Acoust. Soc. Am. Suppl.* **1 92**, S2414(A).

# Three-dimensional mechanical admittance: Theory and new measurement method applied to the violin bridge

Xavier Boutillon<sup>a)</sup>

Laboratoire d'Acoustique Musicale, Université Paris 6, CNRS, Ministère de la Culture, 11 rue de Lourmel, 75015 Paris, France

Gabriel Weinreich<sup>b)</sup>

Randall Laboratory of Physics, University of Michigan, Ann Arbor, Michigan 48109-1120

(Received 13 November 1998; revised 19 February 1999; accepted 3 March 1999)

The mechanical admittance (or mobility) measures the generalized velocities of a system under generalized forces, whereas the impedance measures the forces when velocities are imposed. It is shown that generally, and in cases of practical interest, the experimental determination of the impedance and that of the admittance must comply with different requirements. Therefore, one description cannot be derived from the other unless the degrees of freedom of the system which are not measured are properly dealt with. Some of the experimental methods presented in the literature are discussed along these lines. A new method is proposed for measuring locally the mobility matrix: it is based on comparison with a known mechanical impedance and requires no force measurements. A realization is presented in the case of the bridge of a violin and a quarter-size cello. Theoretical requirements are found to be met between  $\sim 450$ – $1500$  Hz for the violin and  $\sim 250$ – $2000$  Hz for the cello. Limitations of the method are found to be the influence of rotational degrees of freedom and the scarcity of modes at some frequencies. Some implications of the findings for the acoustics of the violin are also discussed. © 1999 Acoustical Society of America. [S0001-4966(99)02406-6]

PACS numbers: 43.75.De [WJS]

## INTRODUCTION

The term “admittance” is extremely general and is used in a variety of contexts for expressing the ratio between the complex amplitudes of two conjugate quantities or sets of quantities which describe the state of a linear dynamical system at a given frequency: for instance, current and voltage in electrodynamics, velocity and force in mechanics, flow and pressure in acoustics. The present article is devoted to the investigation of the mechanical admittance (which is also called “mobility,” a more descriptive term, in part of the literature). The extension of what is presented here to other sets of variables found in acoustics, electrodynamics or electroacoustics is straightforward. In what follows, the term “admittance” will be used in order to stress the potential generality of the underlying concepts.

In Sec. I, the notions of impedance and admittance are analyzed in the context of multiple input variables. The two notions are shown to fit specific and different experimental conditions, a fact which is not always recognized. The analysis is illustrated in the cases of a mechanical point and of a rigid body. The second section reviews part of the literature in the light of these theoretical considerations and presents a new method for measuring the mechanical admittance or mobility of a point of a structure. The experimental procedure is detailed in the next section in the particular instance of measuring the mobility of the bridge of a violin. The last section is devoted to results and to discussion. A very coarse presentation of this research was presented in two conferences.<sup>1,2</sup>

## I. THEORETICAL BACKGROUND

### A. Admittance versus impedance

Defining a dynamical system consists in distinguishing its interior from its exterior. In the approach presented here, it is also necessary to define explicitly all the points at which the (supposedly external) experimenter may act on the system. These points will be called ports. Rather than looking at the system from the inside (considering, for example, the normal modes of a mechanical structure, or the state-space of a dynamical system), this article focuses on the ports of linear systems and is restricted to the case of a finite set of such ports. The variables describing the dynamics of the system at these points will be the generalized velocities (such as the mechanical velocity of a point for a mechanical structure, the current in a wire for an electrical circuit, the acoustical flow entering a pipe, etc.); those describing the actions exerted by the experimenter on the system will be called generalized forces (respectively a mechanical force, an electromotive force, an acoustical pressure). All these form conjugate pairs and choosing one definition is merely a question of convention: in some analogies, acoustical pressures may correspond to currents and flows to voltages. Once a convention is chosen, thought-experiments such as those described in the next section may refer to an imposition of velocity rather than force.

The generalized forces and velocities are viewed as scalar, even when they arise as components of vectors. By definition, the *admittance matrix* is the square matrix which, when multiplied by the column matrix of generalized forces, yields the column matrix of generalized velocities. The *impedance matrix* does the opposite.

<sup>a)</sup>Electronic Mail: xb@moka.ccr.jussieu.fr

<sup>b)</sup>Electronic mail: weinreic@umich.edu

Clearly, the *full* admittance and impedance matrices are inverses of each other, making the two descriptions equivalent in the sense that knowledge of a system's behavior under applied forces (admittance) determines its behavior under imposed velocities (impedance) as well. This is only true, however, if the knowledge of an  $N$ -port system is *complete* in the sense that one has determined the velocity of each port under  $N$  linearly independent sets of forces, so that one knows what each of those velocities will be when, say, a unit force is applied to any one port and no force to any of the others.

In practice, however, one often deals with a *partial admittance* (or *impedance*) *matrix* of dimension lower than  $N$ , say  $m$ . Such a matrix gives the velocity (force) in each of the  $m$  ports when a unit force (velocity) is applied to one of them and the force (velocity) is zero for each of the remaining  $N - m$  ports. Since the conditions at the  $N - m$  "excluded" ports are different for the partial admittance matrix and the partial impedance matrix, it is no longer possible to obtain one from the other. We thus see that, physically, the measurement of impedances must, both conceptually and practically, be carefully distinguished from the measurement of admittances, in that fundamentally different conditions are imposed. (Mathematically, the inverse of a submatrix of one matrix is not a submatrix of the inverse of the matrix.)

To be more specific, the condition at an excluded port in case of an admittance measurement means having a null mechanical force (usual and intuitive understanding of a "free point"), imposing a zero-pressure at an acoustical orifice (which is slightly different from opening it, making this requirement difficult to meet in practice), or short-circuiting an electrical port. In the course of the determination of a partial impedance matrix, the corresponding conditions on excluded ports are imposing a zero-velocity in the mechanical case (in effect, blocking the system there), closing an orifice by a rigid boundary in the acoustical case, or leaving open the port of an electrical circuit. These practical requirements show that one is more likely to measure mechanical admittances on the one hand and acoustical impedances on the other hand. They also call for a thorough examination of what are actually the ports, in a given physical situation.

One case in which complete measurements are taken on *all* ports, so that the distinction between impedance and admittance measurement becomes unimportant, is where the system only has one port altogether. For example, if our interest is in the orifice of a pipe—possibly open at the other end, in which case the free-field is part of the "interior of the system"—it does not really matter how the system is forced. Both force and velocity are measured and one ratio or its reciprocal gives valid results: either the impedance or the admittance. At the other extreme, measuring, say, the admittance of a system with  $N$  degrees of freedom necessitates the forcing by  $N$  independent sets of forces and the measurements of  $N$  sets of  $N$  velocities. Either impedance or admittance components can then be found by the solution of  $N$  linear equations, each having  $N$  unknowns. Any excitation by a mixture of forces and velocities and measurement of the dual complete set is acceptable too in this context.

A factor that often makes practical cases different is that

it is not always possible to isolate, physically, all the ports. For instance, a mechanical point has intrinsically three degrees of freedom; but even if one is interested in only one direction of motion, it may be hard in practice to act only on that particular "port" without affecting the others. Thus applying a force in the  $x$  direction and measuring the velocity in the same direction gives no useful information at all if we are left ignorant of what is going on in the  $y$  and  $z$  directions. If we are careful to keep the forces in the other directions zero, then the ratio of the  $x$  velocity to the  $x$  force gives one element of an admittance matrix. On the contrary, if we are careful to allow the point to move only in the  $x$  direction, then the ratio of the  $x$  force to the  $x$  velocity provides one element of an impedance matrix. Obviously, it would be very dangerous to refer to those two numbers as "the admittance" and "the impedance," since *the two are by no means reciprocals of each other*.

## B. Admittance of a point and admittance of a rigid body

Having three degrees of freedom, a point has also a  $3 \times 3$  mechanical admittance. Under the most natural choice of coordinates, it expresses the velocity taken by the point per unit force in each orthogonal direction:

$$\begin{pmatrix} V_x \\ V_y \\ V_z \end{pmatrix} = \begin{pmatrix} Y_{xx} & Y_{xy} & Y_{xz} \\ Y_{yx} & Y_{yy} & Y_{yz} \\ Y_{zx} & Y_{zy} & Y_{zz} \end{pmatrix} \begin{pmatrix} F_x \\ F_y \\ F_z \end{pmatrix} \quad (1)$$

or, in a condensed form:

$$\mathbf{V} = \mathbf{YF}. \quad (2)$$

In fact, it is impossible to exert a force and measure velocity components at a mathematical point; the measurement locus always has a certain extent and may have to be considered as a rigid body. Thus one may have to deal with six degrees of freedom and a  $6 \times 6$  matrix. It is natural to choose as degrees of freedom the three translations of the point under consideration and three independent rotations  $\Omega_\alpha$ ,  $\Omega_\beta$ , and  $\Omega_\gamma$  of the rigid body around it accompanied by the corresponding torques. The  $x$ ,  $y$ , and  $z$  axes may coincide, respectively, with  $\alpha$ ,  $\beta$ , and  $\gamma$ , but this is not a requirement. We then have

$$\begin{pmatrix} V_x \\ V_y \\ V_z \\ \Omega_\alpha \\ \Omega_\beta \\ \Omega_\gamma \end{pmatrix} = \begin{pmatrix} Y_{xx} & Y_{xy} & Y_{xz} & Y_{x\alpha} & Y_{x\beta} & Y_{x\gamma} \\ Y_{yx} & Y_{yy} & Y_{yz} & Y_{y\alpha} & Y_{y\beta} & Y_{y\gamma} \\ Y_{zx} & Y_{zy} & Y_{zz} & Y_{z\alpha} & Y_{z\beta} & Y_{z\gamma} \\ Y_{\alpha x} & Y_{\alpha y} & Y_{\alpha z} & Y_{\alpha\alpha} & Y_{\alpha\beta} & Y_{\alpha\gamma} \\ Y_{\beta x} & Y_{\beta y} & Y_{\beta z} & Y_{\beta\alpha} & Y_{\beta\beta} & Y_{\beta\gamma} \\ Y_{\gamma x} & Y_{\gamma y} & Y_{\gamma z} & Y_{\gamma\alpha} & Y_{\gamma\beta} & Y_{\gamma\gamma} \end{pmatrix} \times \begin{pmatrix} F_x \\ F_y \\ F_z \\ T_\alpha \\ T_\beta \\ T_\gamma \end{pmatrix}. \quad (3)$$

The  $3 \times 3$  admittance matrix of Eq. (1) is simply the upper left partial matrix of the one in Eq. (3). If, as usual, one measures only the forces  $F_x$ ,  $F_y$ ,  $F_z$  and the velocities  $V_x$ ,  $V_y$ ,  $V_z$ , it is possible to determine this partial matrix provided that no torque is exerted on the rigid body when forces are applied to the point. Otherwise, the measurement has the physical dimensions of an admittance but not the correct physical significance. Another practical requirement is that translations must be measured without being biased by actual rotation of the whole rigid body.

When all the elements of a (possibly partial) impedance or admittance matrix are determined independently, several authors including ourselves have “tested” the validity of the measurements by checking the positiveness of the real part of the diagonal elements (when the structure is dissipative) and the equality of the off-diagonal elements (consequence of the principle of reciprocity): these are necessary conditions that the coefficients must fulfill. However, it should be remembered that the reduction of this test to the case of a measurement in one single direction is not indicative at all of the physical validity of the result: the fact that phases of the force and the velocity along the same direction are less than 90 degrees apart when the structure is dissipative only proves that *some* energy has been put in that direction. It does not prove that energy has been put in *only* that direction, as it should be for a partial impedance or admittance measurement. So this “test” can be very well met even if the ratio of force and velocity does not have the physical meaning of an impedance or an admittance.

After revisiting part of the literature in Sec. II A along these theoretical lines, a new method will be presented in Sec. II B. It will be explained in Sec. III how these practical requirements are addressed.

## II. MEASUREMENT METHODS

### A. Review of some methods or experiments of measuring mechanical admittances

Following the previous discussion, this review focuses on the techniques for imposing the force at the point(s) of measurement. The way of measuring the motion will not be discussed here since it is relatively easy to do it without disturbing the structure. Three kinds of methods can be identified where the force is exerted (a) by an actuator permanently attached to the structure, (b) impulsively by the shock of a hammer on the structure, and (c) by an electromagnetic force, without direct contact.

(a) Measurements with a so-called “impedance head” are generally aimed at measuring the “mobility in one direction,” that is, a diagonal coefficient of an admittance matrix. It applies usually to structures much stiffer in the plane normal to the direction of interest (typically, when the structure undergoes flexural vibrations). The impedance head is attached to the structure, possibly with an intermediate wire, and the ensemble is driven by a mechanical actuator. The following problem must be considered: the mobility matrix generally has off-diagonal elements; hence, the applied force in the direction of interest tends to induce a motion in the normal plane also. If the internal transverse stiffness of the

driver is not negligible compared to the stiffness in the normal plane of the structure, an unknown reaction force is induced and exerted on the structure. More than one port is now excited; in turn, this force is multiplied by off-diagonal terms and the resulting velocity contaminates the measurement of the desired diagonal coefficient. The quantitative analysis has been recently presented by Faure.<sup>3</sup> For this reason, a wire is usually inserted between the impedance head and the structure. It must have a transverse stiffness negligible at all frequencies of interest in comparison with that of the structure. In this case, and despite the name “impedance head,” one diagonal element of an admittance matrix is measured. If no wire is used and the driver has a transverse stiffness much *greater* than that of the structure in the normal plane, one element of the *impedance* matrix is measured. In the intermediate case, the measured quantity has no clear physical meaning and is a linear combination of admittance coefficients, influenced by the internal impedance of the driver.

In the context of violin acoustics, Eggers<sup>4</sup> has used a contact method with a rod linking the shaker to the cello bridge. More recently, Trott<sup>5</sup> used a similar type of mounting between the shaker and the violin bridge. Reportedly the measurements are influenced by a mode whose motion is normal to the bridge plane; only under critical alignment of the driving ensemble could this mode become invisible. This is not surprising: this alignment is simply a situation where a linear combination of the admittance coefficients minimizes that particular mode. Developed for stringed instruments, the design proposed by Firth<sup>6</sup> joins the impedance head and the shaker. This method is also subject to the problems described above.

(b) In the context of the acoustics of stringed instruments, several authors have excited the structure by means of an impact hammer. In the two cases discussed here, the impact was on one side of the bridge whereas the motion was measured on the other side. Jansson<sup>7</sup> deals with only one direction of motion and one direction of force whereas Lambourg and Chaigne<sup>8</sup> present measurements relative to two directions, both for force and motion, and a  $2 \times 2$  admittance matrix. The admittance is derived from a Fourier analysis of the force and acceleration signals. The method is theoretically subject to the same problem as above: a force may be exerted by the hammer during the impact in the direction transverse to the one under consideration (that is, the one oriented along the velocity of the hammer just before the impact) due to a transverse motion of the structure and the reaction of the hammer to that motion. A practical solution is that the hammer have a negligible impedance compared to that of the structure (very light, or very slippery). This in turn would ensure that the reaction force caused during the impact by the coupling terms of the local admittance (as explained for the previous method) are negligible.

According to Woodhouse (personal communication), who practiced the method in Jansson’s laboratory, the absence of rotation of the hammer after its bouncing is easy to check visually and a good sign of the absence of transverse force (at least in one direction) during the shock. Altogether, this method appears to meet the theoretical requirements,

provided that appropriate guidelines are followed during the experiment. Lambourg and Chaigne's experiments also comply almost certainly with these requirements since their measurements of off-diagonal conjugate terms are so very similar to each other (strictly speaking, these "checks" are necessary but not sufficient).

Notice that the dual requirement (zero velocity instead of zero force) for measuring an impedance would be that the hammer be very heavy or constrained to move in only one direction *for the whole duration of data taking*. As the hammer is supposed to bounce on the structure and to lose contact, the boundary conditions change and the requirement cannot be met; this means that, paradoxically, no impedance measurement can be performed with an "impedance hammer."

(c) A third kind of excitation consists in fixing a light magnet on the structure and imposing a current through a nearby coil.<sup>9</sup> In this case, the only concern is the linearity of the excitation process: the motion transverse to the force direction should be sufficiently small as to leave the magnet in a region where the magnetic field can be considered as uniform. Yoder<sup>10</sup> has suggested a way of calibrating the procedure based on a one-dimensional version of the comparison method that we present here.

Besides a discussion of the experimental methods, several other articles or reports are worth reviewing here, mainly because they address the question of multiple ports. As analyzed by Cremer<sup>11</sup> in Chap. 9.2, Steinkopf was the first to discuss the violin bridge as a multiple-port system in an unpublished work of 1963. Three ports of the isolated violin are examined by him: the top of the bridge in the bowing direction and the two feet of the bridge in the direction perpendicular to the top-plate. The nine coefficients of the admittance matrix are not measured independently; instead, hypotheses on the system and reciprocity considerations reduce the number of expected different coefficients to four (therefore, no check could be done on the validity of the results). Three of them are adequately measured: velocity measurements are made under imposition of a force at one port and no constraint at any other port. The fourth coefficient is not reported on. Besides, a measurement of an input-impedance coefficient is described for the same three-port system, at the top of the bridge, with, reportedly, the method of excitation given by Eggers. The two other ports are blocked but, at the point of measurement, the bridge is left free to move in the other directions, instead of being constrained as it should be for a partial-impedance determination. The results are discussed in terms of one single resonance (minimum of the impedance): it could well be that the coupling with other directions of motion shifts the position of a sharp minimum. The same method has been used by Reinicke.<sup>12</sup>

Another part of the work by Reinicke<sup>12</sup> presented by Cremer<sup>11</sup> in Chap. 9.6 deals with the admittance of the violin body at the locations of the violin feet, again perpendicularly to the surface. It is not specified by Cremer how the excitation is produced, but the results comply remarkably well with the theoretical requirements on the admittance matrix over a large frequency range.

## B. The comparison method

The purpose of this method<sup>13</sup> is to avoid measuring forces and to deal only with measurements of motion. The basis of the comparison method relies on the linearity of the system under consideration. Therefore, Eq. (1) or (2) between forces and velocities holds also for differences in forces and differences in velocities between two tests performed on the same system:

$$\Delta \mathbf{V} = \mathbf{Y} \Delta \mathbf{F}. \quad (4)$$

In our case, the two tests are

- (i) one in which the system is shaken by an external source,
- (ii) one where the system has been loaded at the point of interest by a known mass or any other known impedance, and is shaken again by the same external source operating in the very same way. We shall discuss the case where the load is a mass.

The two tests will be referred to respectively as the "free" and the "loaded" tests, and corresponding quantities will be indexed accordingly.  $\Delta \mathbf{V}$  is the velocity difference  $\mathbf{V}_f - \mathbf{V}_l$  between the two tests, measured at the point of interest.  $\Delta \mathbf{F}$  is the difference in force exerted on the point of interest between the two tests. This is simply the opposite of the inertial force on the mass during the second test:

$$\Delta \mathbf{F} = \mathbf{F}_f - \mathbf{F}_l = 0 - (-\mathbf{Z}\mathbf{V}_l), \quad (5)$$

where  $\mathbf{Z}$  is the impedance of the mass. This becomes

$$\mathbf{V}_f - \mathbf{V}_l = \mathbf{Y}(\mathbf{Z}\mathbf{V}_l). \quad (6)$$

In this situation, all degrees of freedom of the system are simultaneously excited by the external source, but only three of them are altered by loading a mass at a specific point, considered as a triple input-port. Provided that the action imposed by the exterior does not change from one test to the other one, this complies with the criteria given in Sec. I A. With only velocity measurements in the three directions of space, one pair of tests gives three linear equations on the nine coefficients of the admittance matrix. This doubling of measurements is the price for getting rid of force measurements. Three pairs of velocity measurements have to be done in dynamically independent situations in order to obtain the nine desired equations. The three dynamical cases are created by changing the external source of excitation, in effect "polarizing" the system in three different states. The linearity of the system guarantees that its dynamics is not changed. Equation (6) between vectors is then written three times, becoming an equation between  $3 \times 3$  matrices, having as columns the three vectors of velocity-differences  $\mathbf{V}_f - \mathbf{V}_l$  and the three vectors  $\mathbf{Z}\mathbf{V}_l$ . By hypothesis, the tests are independent, so that  $\mathbf{Z}\mathbf{V}_l$ , which is now a matrix, can be inverted, giving explicitly the admittance as

$$\mathbf{Y} = (\mathbf{V}_f - \mathbf{V}_l)(\mathbf{Z}\mathbf{V}_l)^{-1}. \quad (7)$$

The name "comparison method" is now justified: it consists in comparing the velocity taken by the loaded system to the one taken by the free system. The physical scale for the admittance is the reciprocal of the impedance of the loading mass. In practice, they must have the same order of

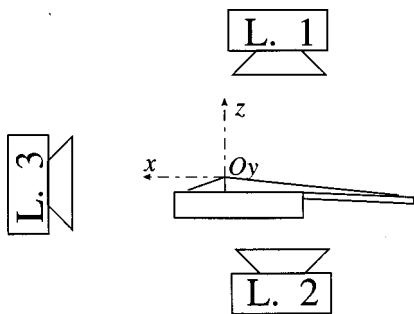


FIG. 1. General setup.

magnitude. If the load is too light, the difference in velocity will be too small, and therefore sensitive to noise; if it is too heavy, the velocity of the loaded system will be too small. Any known impedance can indeed play the role of the mass, but this particular kind is certainly the easiest sort to calibrate.

### III. EXPERIMENTAL PROCEDURES

The experiments have been carried out in a pseudo-anechoic room (~2 m high, 3.5 m long, and 1.8 m wide, lined on all six faces with 4-in. Sonex). The requirement on the room is that the *incoming* acoustical field on the violin is the same for the two corresponding “free” and “loaded” tests. This field can be seen as the superposition of two; one is due to the vibration of the loudspeakers (unchanged from one test to the other) and the other one is due to the waves generated by the vibrations of the violin (different between the two tests) and reflected by the walls of the room. The *outgoing* field generated by the vibrations of the violin is obviously weak compared to the incoming field of the loudspeaker which induces them; therefore, our pseudo-anechoic room is sufficient to make the corresponding reflected field weaker to the point where it can be neglected.

The violin rests on transverse bars across an aluminum frame, hung at mid-distance between the floor and the ceiling. The strings are damped with pieces of foam inserted between them and the fingerboard. The violin is surrounded with three loudspeakers, as shown in Fig. 1: one on the floor underneath the violin, one hanging from the ceiling above the violin and the third one put on a stool at the height of the violin.

The loudspeakers are driven by a power amplifier Crown D150A. A custom-made piece of aluminum of total weight ~0.9 g (a rectangular parallelepiped 9×9×7 mm, with a cylindrical cavity of diameter 7 mm and depth 6 mm) receives the mass and the accelerometers. It is glued on the flat side of the bridge facing the neck, between the A- and the D-string, with cyanoacrylate glue. When a valuable instrument is used, a thin intermediate layer of wood is added to protect the bridge. During the “loaded” test the cavity of the aluminum piece receives a mass (lead cylinder of 3.8 g) surrounded with a coat of wax to keep it firmly in place. The accelerometers are Knowles BU1771 piezoelectric sensors (0.3 g) with a built-in field effect transistor (FET), operated with custom-made amplifiers.

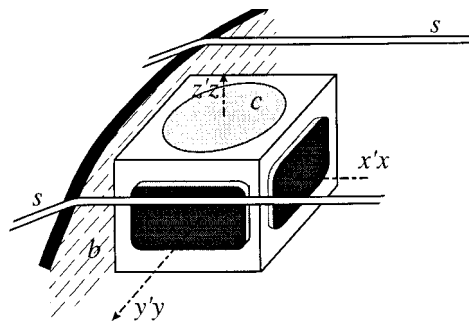


FIG. 2. Measuring device. a: accelerometers (the one measuring along the  $z$  axis is hidden, b: bridge, c: cavity for reference mass, and s: D and A strings.

Directions are chosen for consistency with Cremer so that the  $x$  axis is parallel to the strings, the  $y$  axis is parallel to the normal bow-direction, and the  $z$  axis points from the bottom to the top plate (to be precise, since the plane is in the  $yz$  plane, the  $x$  axis makes a small angle with the string direction). The center of reference is at the center of the added mass, inside the hollow piece of aluminum.

As shown in Fig. 2, the three accelerometers are glued on the aluminum support described above, one on the face opposing the opening of the cavity for measuring the  $z$  component of the acceleration (invisible in picture), one on the face opposing the bridge for measuring the  $x$  component, and the third on one of the two free-remaining parallel faces for the  $y$  component.

Signals are converted at a 8192 Hz sampling rate through a 12-bit DAC/ADC board (Data Translation DT 2128) operated by a 286-PC with custom-written drivers. The excitation signal ( $\frac{1}{8}$  of a second, 1024 points long) was generated according to the algorithm suggested by Schroeder:<sup>14</sup> it has a flat amplitude spectrum on the frequency band 100–3,000 Hz, with a quadratic phase that ensures a very low peak-factor. Its frequency resolution is 8 Hz. At first, a frequency resolution of 1 Hz was used but it has been checked in the course of the experiments that operating with an 8 Hz frequency resolution did not change the graphs of the results presented below in any visible way.

The power amplifier is adjusted to produce an output level of a few volts when a full-scale sinusoid is produced by the DAC. The form of Eq. (7) giving the admittance at the point of reference shows that the incoming acoustical signals have no influence on the result, as long as they induce a vibration of sufficient level to overcome noise. In particular, their spectrum need not be particularly flat. Therefore, the frequency response of the loudspeakers was not considered.

Since the second member of Eq. (7) involves a *ratio* of velocities, one can replace them by the corresponding accelerations. The admittance can also be written, and has been calculated, as

$$\mathbf{Y} = (\mathbf{A}_f - \mathbf{A}_l)(\mathbf{Z}\mathbf{A}_l)^{-1}. \quad (8)$$

Further, a considerable advantage of the method is that no absolute calibration of the accelerometers needs to be made. It is not even necessary that they have a flat response curve. However, the matrix inversion involved in the calculation of the admittance makes it necessary that the three

accelerometers all have the *same* response curve and response level. The accelerometers have a very sharp mechanical resonance at around 12 kHz and a very flat response below 3 kHz. A difference of 1 kHz in the resonance-frequency of the accelerometer (which would correspond to a rather poor manufacturing uniformity) would induce a difference in frequency response of less than 1% in amplitude at 3 kHz. Therefore, the intercalibration of the three accelerometers reduces to the datum of one constant factor, valid at all frequencies. This correction has been done by adjusting the gains of their respective amplifiers so that, when the accelerometers are subjected to the same mechanical vibration, the three accelerometer-amplifier ensembles all give the same response level.

The linearity of the whole experimental chain—DAC, power amplifier, loudspeaker, violin, accelerometers, amplifiers, and ADC—has been checked. In fact, it is a test very sensitive to any defect in the attachment of the accelerometers or of the mass to their support, or of the whole device to the violin bridge.

Each response is obtained by running the excitation signal 20 times without interruption. No data are taken during the first four runs which allow the system to reach its steady regime of vibration. The data are then averaged on the 16 next runs in order to increase the signal-to-noise ratio.

#### IV. RESULTS AND DISCUSSION

The results that are presented here have been obtained on a low-price violin (Fig. 3) and a quarter-size cello of decent quality (Fig. 4). As the main objective of this article is to present a method which measures independently the nine complex coefficients of the admittance matrix, these results will be mainly discussed in perspective of their intrinsic validity. The theoretical requirements that must satisfy the elements of an admittance matrix are that the real part of the diagonal elements be positive (the violin is a dissipative structure) and the off-diagonal symmetric elements be equal (it obeys principle of reciprocity). As can be seen on the figures, the requirements are met within a frequency-range of  $\sim 400$ – $1500$  Hz for the violin and  $\sim 250$ – $2000$  Hz for the small cello. Within these ranges, there are isolated phase-mismatches occurring in correspondence with a very low or very high amplitude level, compared to the amplitude of the reference admittance (a 3.8-g mass): these are of no significance. Other sources of discrepancies are discussed in Secs. IV B and IV C. For allowing comparison with other published measurements of the input-mobility in the bowing direction, the magnitude of  $Y_{yy}$  is redrawn in Fig. 5 in dB, within the frequency range of validity of our results.

##### A. Implications for violin acoustics

The transverse motion of a bowed string exerts a force in the  $yz$  plane (with  $F_y$  usually much larger than  $F_z$  due to the polarization of the motion). Within the frequency ranges mentioned above, the nine mobility coefficients are of comparable order of magnitude in each instrument, the input-mobility  $Y_{yy}$  in the bowing direction being only about three times larger than the other input or transfer mobilities. Con-

sequently, the out-of-plane motion of the bridge (in the  $x$  direction) is as important as the in-plane motion (in the  $yz$  plane) when a string is bowed. This in turn may have at least two effects: coupling transverse waves (mainly in the  $y$  direction) and longitudinal waves (in the  $x$  direction) on the string, and sound radiation of the bridge itself (due to its out-of-plane motion).

##### B. Admittance and normal modes

It has been shown in Sec. II B that three independent excitations are necessary to produce an invertible matrix  $\mathbf{Z}\mathbf{A}_l$  at a given frequency. The concept of independence must be appreciated in a quantitative sense, as indicated by the extent to which the matrix  $\mathbf{Z}\mathbf{A}_l$  is well or ill conditioned. Since any structural vibration of the violin can be projected on the basis of its normal modes, this requirement means that at least three normal modes must exhibit a significant level at that frequency and must be excited differently by the three loudspeakers. This involves modes that (a) are well coupled to the acoustical field, (b) have an eigenfrequency not too far from the frequency of interest, and (c) do not have a node near the point of interest.

Criterion (b) is best discussed in terms of the modal overlap, defined as the ratio between the half-power bandwidth of a single mode and the average modal spacing. When the modal overlap is large, the number of possible active modes is of the order of the overlap; this is favorable to the method. This situation is encountered at high enough frequencies. When the modal overlap is small (that is for high Q-factors and large modal spacing), difficulties will arise at least in the resonance band of each individual mode, where it largely dominates the whole vibration. In between the resonance frequencies, damping has almost no influence on the level of vibration of each mode so the modal overlap is irrelevant; the number of active modes depends only on the details of the modal structure, and on the modal projection of the applied force.

If, at a given frequency, the vibration of a point is largely dominated by the contribution of one single mode (because all the other modes have a large modal mass or a remote eigenfrequency or a nodal line at that point), there is effectively only one single mechanical port at that point. A reference frame can be found in which the velocity has only one spatial component and in which the admittance matrix reduces to one scalar. Under this circumstance, or when only two modes are present, neither a mechanical nor an acoustical excitation can produce three independent velocity-vectors and the algorithm presented here for finding the admittance matrix is not applicable. This situation may be encountered at any frequency but there is no theoretical limit (related to the lowest eigenfrequency of the violin, for instance) below which this situation would be necessarily met: omitting the spatial factor, the mobility of a mode is asymptotically equal to  $\omega/M\omega_0^2$ , where  $M$  is the modal mass of the mode and  $\omega_0/2\pi$  its eigenfrequency. It is theoretically possible that three modes have comparable contributions at an arbitrary low-frequency. Therefore, one has to evaluate at each frequency whether the vibration is one-, two- or three-dimensional. If the velocity has a linear or planar polariza-



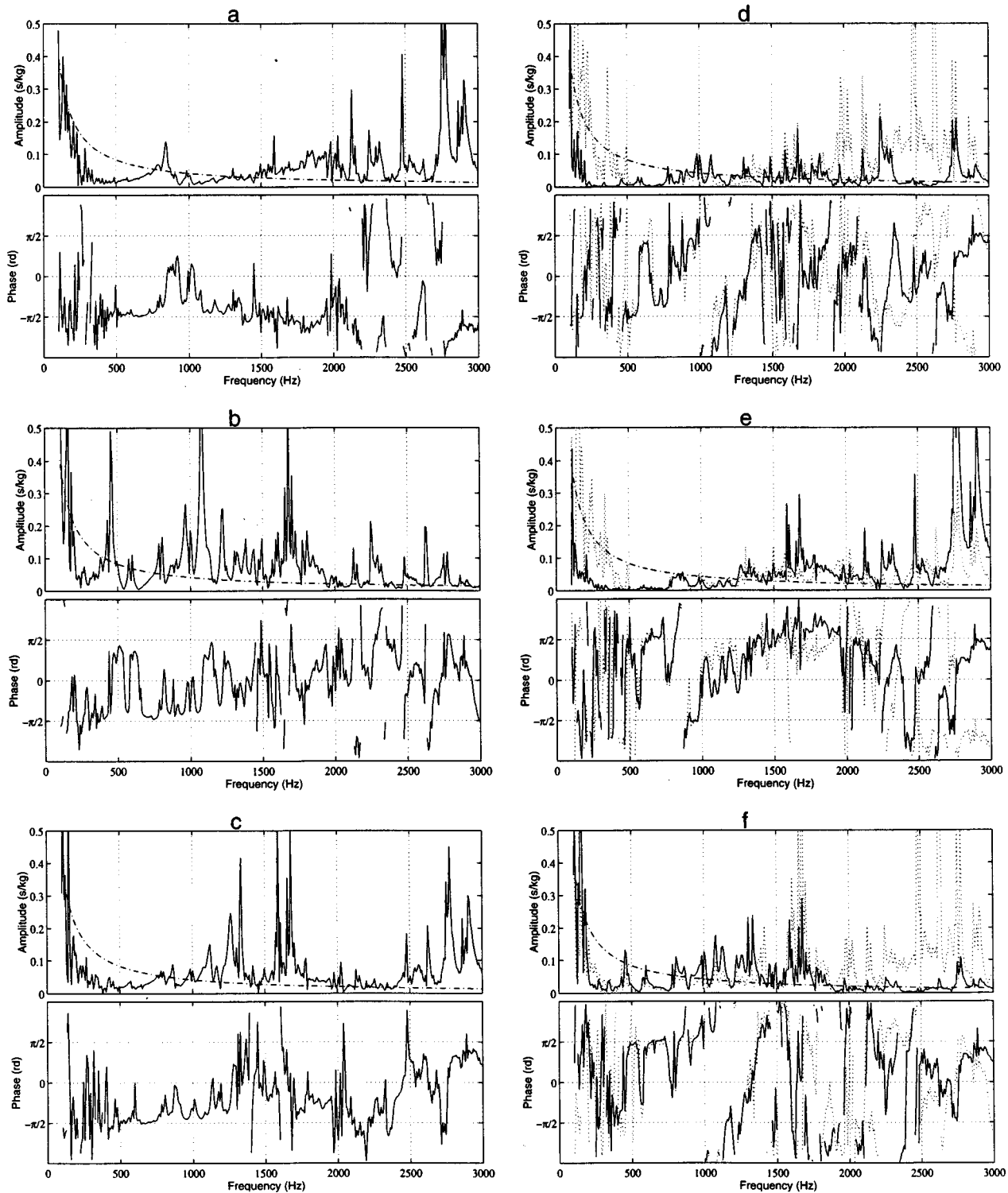


FIG. 3. Mobility at the top of the bridge of a violin. (a)  $Y_{xx}$ , (b)  $Y_{yy}$ , (c)  $Y_{zz}$ , (d)  $Y_{yx}$  (solid),  $Y_{xy}$  (dashed), (e)  $Y_{zx}$  (solid),  $Y_{xz}$  (dashed), and (f)  $Y_{zy}$  (solid),  $Y_{yz}$  (dashed). Dash-dot: Mobility of the reference mass (3.8 g).

tion, the admittance matrix could still be written in a  $3 \times 3$  presentation but it would not be invertible. Given these conditions, the corresponding  $3 \times 3$  impedance matrix would exhibit one or two extremely large eigenvalues, corresponding to the fact that the system cannot physically be put into more than two or one patterns of motions.

In order to estimate at which frequency the matrix  $\mathbf{Z}\mathbf{A}_l$  cannot be inverted safely (i.e., the violin presents only one or

two degrees of freedom at that point or the signal becomes very weak), its eigenvalues have been evaluated and their ratio examined. The sorted eigenvalues of the matrix of acceleration  $\mathbf{A}_l$ , called  $A_1, A_2, A_3$  in ascending order of their modulus, are presented for the violin in Fig. 6. The figure shows that the vibration level becomes very low below 180 Hz: this is because the monopole moment of the violin vanishes at low frequencies.<sup>15</sup> A weak level of  $A_1$  compared to

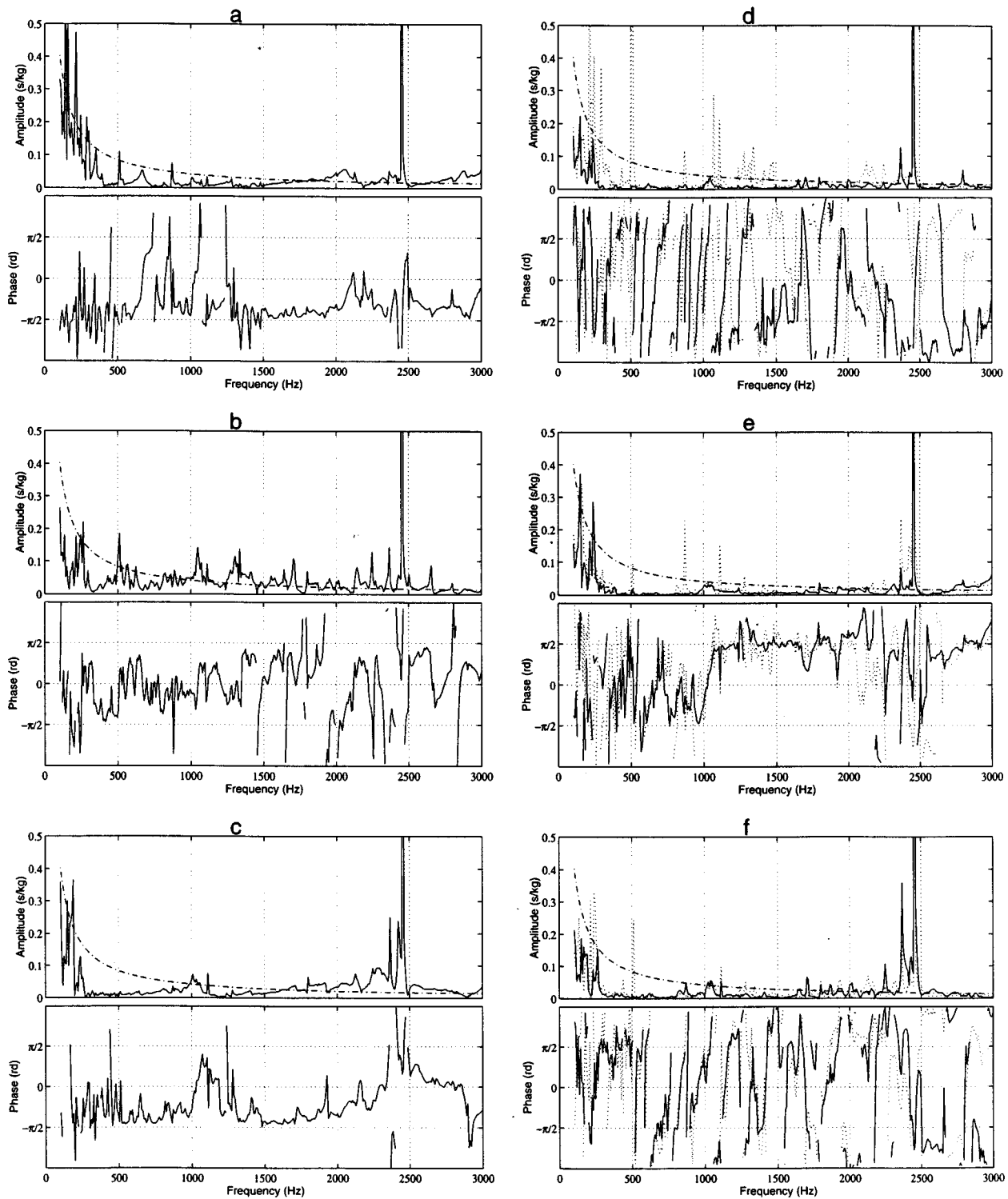


FIG. 4. Mobility at the top of the bridge of a quarter-size cello. (a)  $Y_{xx}$ , (b)  $Y_{yy}$ , (c)  $Y_{zz}$ , (d)  $Y_{yx}$  (solid),  $Y_{xy}$  (dashed), (e)  $Y_{zx}$  (solid),  $Y_{xz}$  (dashed), and (f)  $Y_{zy}$  (solid),  $Y_{yz}$  (dashed). Dash-dot: Mobility of the reference mass (3.8 g).

$A_3$  indicates that the motion has a planar polarization whereas a weak level of  $A_2$  compared to  $A_3$  indicates a linear polarization. This is clearly the case below  $\sim 450$  Hz for the violin, which explains why the  $3 \times 3$  admittance matrix could not be determined successfully in this low-frequency range. Above the frequency range given in the figure, there was no problem of that sort for the violin, except for isolated fre-

quencies. The same phenomenon could be observed below  $\sim 250$  Hz for the small cello.

### C. Rotations and torques

Kinematical and dynamical effects may couple rotations or torques to the measurements:

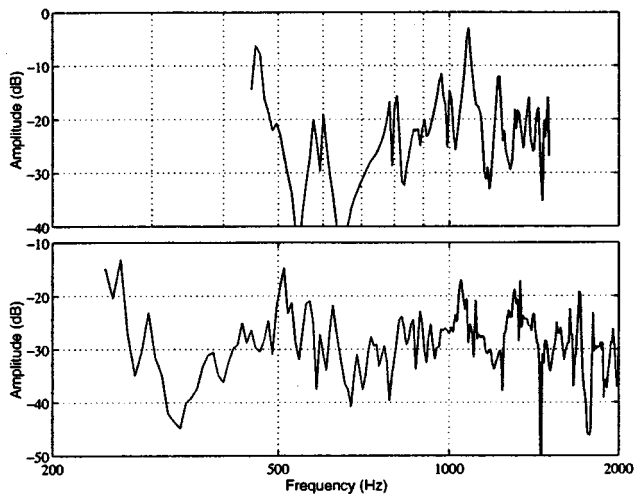


FIG. 5. Amplitude of the input mobility in the bowing direction  $Y_{yy}$  for the violin (upper) and the small cello (lower). Data are limited to the valid frequency range. 0 dB corresponds to 1 s/kg (i.e.,  $1 \text{ m} \cdot \text{s}^{-1} \cdot \text{N}^{-1}$ ).

(a) the accelerometers may be sensitive to rotational motions,

(b) some translational motion may result from applied torques due to off-diagonal elements of the  $6 \times 6$  admittance matrix written in Eq. (3).

Theoretically, the method measures the admittance of a point. In fact, the dimensions of the measuring device make it a solid body unless some characteristic dimension of the device is negligible compared to some physical length characteristic of the experiment. One considers the top plate a thin and almost flat shell. Hence, at the surface of the top plate where each foot of the bridge lies, the vibration is largely dominated by translational motion except at nodal lines (if any) where the motion is purely rotational. Below the first resonance of the bridge (known to occur near 3 kHz) these vibrations are transmitted rigidly to the measuring device. Therefore, nodal lines on the top plate are the physical axes of rotation of the measuring device (“nodal line” does

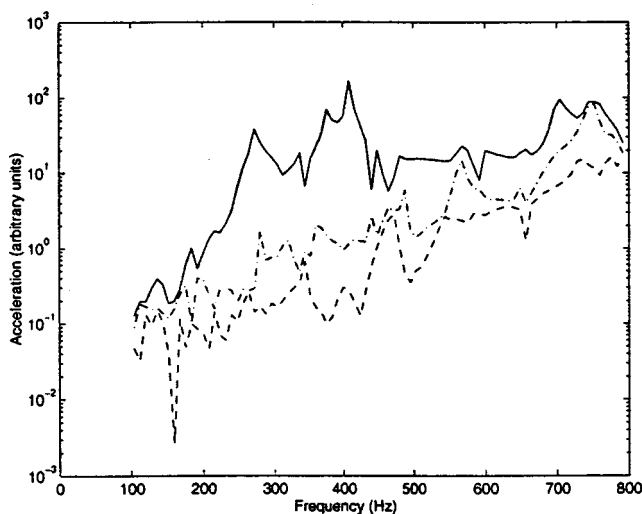


FIG. 6. Eigenvalues of the  $A_j$  matrix of acceleration measurements. Dash-dot: smallest eigenvalue  $A_1$ . Dashed: middle eigenvalue  $A_2$ . Solid: largest eigenvalue  $A_3$ .

not necessarily pertain to one particular mode since several may be contributing to the vibration at one frequency; it simply means the locus at which there is no vibration for that frequency and that excitation).

(a) Kinematical effects relate to the separation of translations from rotations, for a given choice of the reference point (in our case, the center of the reference mass, or the center of the cavity during the “free” tests) and of the three axes of measurement. The accelerometers have a very small sensitive volume: a piezoelectric thin beam ( $\sim 2.4 \times 0.25$  mm), attached to the inside of the box of the accelerometer. Considering them like point accelerometers means that they have a center of sensitivity and are insensitive to rotations around any axis going through that center of sensitivity. One must then put the center of sensitivity of the accelerometer measuring  $a_z$  (respectively  $a_x$  and  $a_y$ ) on the  $z$  (resp.  $x$  and  $y$ ) axis; in addition, the effect of  $\Omega_x$  and  $\Omega_y$  (resp.  $\Omega_y$  and  $\Omega_z$ ,  $\Omega_z$  and  $\Omega_x$ ) becomes minimal. Experimentally, the validity of the results (as indicated by the phase of the diagonal elements and the equality of off-diagonal elements of the matrix) at high frequency proved to be sensitive to small changes in the positioning of the accelerometers. Here the relevant ratio is formed by the distance between the axes of sensitivity of the accelerometers and the distance to the nearest nodal line. It can be reduced significantly by careful positioning.

(b) Dynamical effects are only inertial, since there are no external forces applied to the bridge or the measuring device. Therefore, the relevant parameter is the square of the ratio between the radius of gyration of the lead mass (typically 2 mm) and the distance to the nearest nodal line (effective axis of rotation). The latter is at least equal to the height of the bridge ( $\sim 3$  cm), as the measuring device is glued between the two central strings. The rotational effects of inertia are therefore less than  $\frac{1}{100}$  of translational effects unless the bridge itself deforms. This occurs at its first resonance ( $\sim 3$  kHz) and the closest nodal line is then considerably closer to the measuring device: this is a very plausible source for the 1.5 kHz limit in the validity of the results for the violin.

Whereas the loss of validity of the results above  $\sim 2$  kHz for the small cello and  $\sim 1.5$  kHz for the violin seems to be rooted in the finite size of the device used for the measurements, reducing the overall dimensions of the measuring device is not advisable: it has a dramatic effect (going with the third power) on the mass that can be used as the comparison reference. Also, the overall dimensions of the accelerometers are no less than several millimeters.

## V. CONCLUSIONS

The theoretical discussion together with the analysis of some experimental methods in mechanics showed that admittance rather than impedance measurements are performed in most situations. In practice, it is difficult to ensure the conditions necessary to determine true impedance matrices or coefficients; some techniques such as the impact hammer are fundamentally inappropriate for that purpose but can be excellent for admittance measurements. The comparison method with an external load presented here has several ad-

vantages: neither force measurements nor the absolute calibrations of motion sensors are necessary. Drawbacks include the necessity of dynamically independent sets of excitations which may be difficult to obtain when the modal density and the modal damping are low. Sensitivity to rotational effects depends highly on experimental circumstances: the measured structure, the measuring device, and the size of the load.

One could extend the method to a 1-D measurement by means of a load constituted of a mass attached to the end of a short stiff wire, whose other end is attached to the structure at the point of interest. Between the resonance of this mass-spring system and the first resonance of the clamped-clamped beam formed by the wire between the mass and the structure, the impedance of the load in the direction of the wire is that of the mass whereas it is negligible in the transverse plane<sup>16</sup>. This system could in turn be replicated in three dimensions by three converging loaded wires, with the potential advantage of concentrating the load on a smaller space than that needed by the solid mass used here.

Our measurements show that the bridge of a violin or a cello seems to have comparable mobilities in all directions of space and that coupling terms between different directions are also roughly of the same magnitude. The dynamics of motion and coupling of the bowed string to the instrument in the direction of the string may have received less attention than they deserve.

#### ACKNOWLEDGMENTS

We are grateful to Dr. James Woodhouse for a number of stimulating discussions and useful suggestions, not the least of which was the suggestion to write up this paper. This work was performed at the University of Michigan where research on musical acoustics has been supported by the National Science Foundation. Main support for the first author

is due to the French National Center for Scientific Research (CNRS) and the French Ministry of Culture.

- <sup>1</sup>X. Boutillon, G. Weinreich, and N. R. Michael, "Experimental developments for the measurement of violin bridge admittance," *J. Acoust. Soc. Am.* **84**, 179 (1988).
- <sup>2</sup>X. Boutillon and G. Weinreich, "Measurement of the three-dimensional admittance at a violin bridge," *ICA-13: Proc. International Congress on Acoustics, Belgrade* **3**, 87–90 (1989).
- <sup>3</sup>C. A. Faure, "Mécanique des vibrations tridimensionnelles - Méthodes de mesure et calibration," Ph.D. thesis, University Pierre et Marie Curie, Paris, 1997.
- <sup>4</sup>F. Eggers, "Untersuchung von Corpu-Schwingungen am Violoncello," *Acustica* **9**, 453–465 (1959).
- <sup>5</sup>W. Trott, "The violin and its bridge," *J. Acoust. Soc. Am.* **81**, 1948–1954 (1987).
- <sup>6</sup>I. Firth, "Small mechanical impedance head for use with musical instruments," *Acustica* **35**, 348 (1976).
- <sup>7</sup>E. Jansson, "Admittance Measurements of 25 High Quality Violins," *Acustica* **83**, 337–341 (1997).
- <sup>8</sup>C. Lambourg and A. Chaigne, "Measurements and modeling of the admittance matrix at the bridge in guitars," *SMAC-93: Proc. Stockholm Music Acoust. Conference 1993, Royal Swedish Academy of Music, Stockholm*, pp. 448–453 (1994).
- <sup>9</sup>J. Moral and E. Jansson, "Eigenmodes, input admittance, and the function of the violin," *Acustica* **50**, 329–337 (1982).
- <sup>10</sup>L. Yoder, "Normalizing a coil-magnet driver for stringed instrument measurements," *Ninth ISMA, C.A.S. Proc., IIA-4* (1991).
- <sup>11</sup>L. Cremer, *The Physics of the Violin* (MIT Press, Cambridge, MA, 1984).
- <sup>12</sup>W. Reinicke, "Übertragungseigenschaften des Streichinstrumentensteges," Ph.D. thesis, Technical University of Berlin, 1973. See also "Übertragungseigenschaften des Streichinstrumentensteges," *Catgut Acoust. Soc. Newsletter* **22**, 26 (1974).
- <sup>13</sup>G. Weinreich and L. Yoder, "New method of measuring input admittance of stringed instruments," *J. Acoust. Soc. Am.* **81**, 83 (1987).
- <sup>14</sup>M. Schroeder, "Synthesis of low-peak-factor signals and binary sequences with low autocorrelation," *IEEE Trans. Inf. Theory* **16**, 85 (1970).
- <sup>15</sup>G. Weinreich, "Sound hole sum rule and the dipole moment of the violin," *J. Acoust. Soc. Am.* **77**, 710–718 (1985).
- <sup>16</sup>L. Yoder, "Violin admittance measurements using a one-dimensional mass-loading technique," *J. Acoust. Soc. Am.* **89**, 1878 (1991).

# Behavior of piano-action in a grand piano. I. Analysis of the motion of the hammer prior to string contact

Eiji Hayashi

*Department of Mechanical System Engineering, Faculty of Computer Science and System Engineering, Kyushuu Institute of Technology, 680-4, Kawatu, Iizuka, Fukuoka Pref., 820-8502, Japan*

Masami Yamane

*Department of Mechanical Engineering, School of Science and Engineering, Waseda University, 3-4-1, Okubo, Shinjyuku-ku, Tokyo, 169, Japan*

Hajime Mori

*Apex Co. Ltd., Yasukuni Kudan Minami Building, 2-3-14, Kudan minami, Chiyoda-ku, Tokyo, 102, Japan*

(Received 12 January 1998; revised 19 October 1998; accepted 19 February 1999)

Producing a stable soft tone, difficult even for pianists, is one of the critical challenges in the development of an automatic piano. The motion of the hammer prior to string contact was studied to provide a basis for creating a stable, soft tone. A simple dynamic mechanical model was used to analyze the motion. Using the actuator developed for the touch system, the motion of the hammer under various input waveforms was observed and recorded. Based on this analysis, it can be said that the calculated behavior correlates with the measured behavior. These results indicate that the proposed model is useful because it clearly illustrates the behavior of the hammer. The modified constant speed necessary to produce an optimum touch input waveform has now been found, which allows the system to achieve stable soft tones. © 1999 Acoustical Society of America.

[S0001-4966(99)00706-7]

PACS numbers: 43.75.Mr [WJS]

## INTRODUCTION

An automatic piano has been developed that has superior playback capability with the ability to produce stable, soft tones and to achieve fast repetition of the same key.

Automatic pianos, including roll-operated and reproducing pianos, have been manufactured from the late 19th to the early 20th century. The earlier models played back the notes by simply following the length of the slit on a roll, which represented the actual length of the notes to be played. They could not, however, vary the volume of the note being played. The reproducing pianos that followed in the early 1910s, however, featured volume control and had a greatly improved ability to record and play back music compared to the earlier roll-operated pianos. In recent years, automatic pianos have been introduced into the market by several companies. These new machines are superior in terms of recording the touches and strokes of the original pianists as data, and they achieve a far superior playback compared with previous reproducing pianos. However, in these machines, solenoids, driven in an open loop, are used to control the touch of the key. This results in a narrow dynamic range due to the unstable reproduction of soft tones as well as difficulties playing the same key with fast repetitions. There is still much room for improvement.

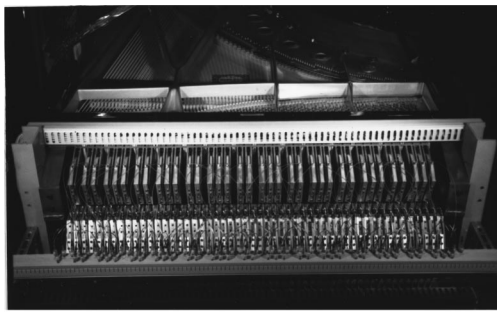
Reliably and repeatedly producing stable tones at a very soft dynamic level is difficult even for pianists, and producing these tones is one of the critical challenges in the development of an automatic piano. Little experimental research has been done on the motions of the grand piano-action,<sup>1-4</sup> which consists of several individual parts, including a key and hammer. As a result, there have not been adequate quan-

titative and theoretical evaluations of these actions. Therefore, a thorough analysis of the motion of the hammer prior to string contact is necessary. Unfortunately, the mechanism of the grand piano-action is complicated, and in order to accurately analyze its motion, it has been necessary to develop dynamic models of a high order, the analysis of which is extremely difficult.

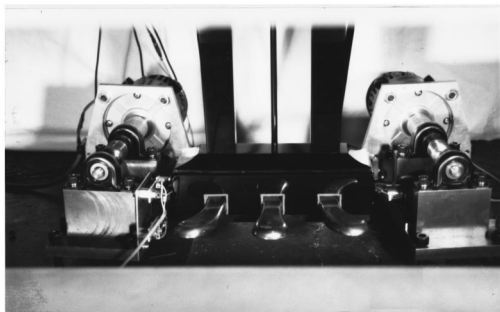
In this article, we analyze the motion of the hammer prior to its striking of the string to provide a basis for creating a stable, soft tone. Although one of the problems with such tones is producing them repeatedly at a very soft dynamic level, we concern ourselves in this article only with a reliable initial production of the tone.

A simple dynamic mechanical model was used to analyze the motion of the hammer. Using the actuator developed for the touch system, which is later described in detail, the motion of the hammer under various input waveforms was observed and recorded. The data obtained here correlate well with the calculated theoretical values, demonstrating that the model used for the analysis is valid and that the behavior of the hammer has been made clear. An optimum touch input waveform for the mechanical input motion of the key has been established that allows the system to create a stable soft tone.

The motions of the key and hammer were observed, and data were recorded for the bass, baritone, and high tones on the keyboard of the piano. The motion was extremely similar for the entire keyboard range. For this reason, this article deals only with the 11th key as an example. The piano used in the demonstration was a C7 grand type from the Yamaha corporation.



(a) The touch actuators



(b) The pedaling actuator

FIG. 1. View of the touch actuators (a) and the pedaling actuators (b) for the automatic piano.

## I. TWO TYPES OF ACTUATORS FOR THE AUTOMATIC PIANO

The two kinds of actuator used in the experiments discussed in this article are described below. These actuators were designed and developed<sup>5,6</sup> for the automatic piano to control touch and pedaling, respectively (see Fig. 1). The touch actuator, as shown in Figs. 1(a) and 2, is no wider than the width of the white key on the piano, and one is allotted to each key on the keyboard, with 88 actuators in all. This actuator is a moving coil type. An epoxy resin laminated plate was chosen as the rotor material for its lightness, and a coil of aluminum ribbon wire was embedded in the rotor. High magnetic flux is obtained in the actuator through the use of a rare earth cobalt magnet. The servo system functions

by means of position feedback, measuring the position of the rotor with an eddy-current-type sensor, and having a follow-up control that allows the fingertip (the touch-point of the rotor) to follow any input waveform. The target plate is fixed in a tilted position to the rear of the rotor. The tilt of the plate is set so that the gap between the sensor and the plate remains constant within  $\pm 0.8$  mm for a rotational angle of  $\pm 0.1$  rad. This actuator produces a torque of 1.33 Nm/A, a thrust of 13.3 N/A at the touch-point of the fingertip, and a frequency response of 60 Hz.

The actuator can function from 10 mm above the key, which is an effective way of producing a very strong tone. However, this presents two problems at the moment the actuator comes into contact with the key: an impact sound is produced, and the key separates from the fingertip due to an impact force. Therefore, the best initial position of the fingertip has been determined to be at the surface of the key when the key is in its unplayed position. To obtain a very strong fortissimo, it is necessary to achieve a velocity of at least a constant 0.8 m/s in the movement of the key from its initial position to the bottom. Though the mechanical impedance of bass keys is greater than that of the middle and treble keys, a bass-key velocity of approximately 2 m/s can be achieved with a force of 50 N at the key front. The servo amplifier for the touch actuator is a pulse width modulated type, with a maximum current of 9 A. The force produced by the actuator at the fingertip can reach a maximum greater than 100 N.

The pedaling actuator was designed as a combination between a dc servo motor and a self-locking mechanism utilizing a worm gear and wheel. The dc servo motor provides high torque with a small moment of inertia. The self-lock mechanism enables a reduction in the heavy load on the dc motor because the pedal remains mechanically locked in the depressed position, allowing the current to the dc motor to be cut off. This servo system is similar to the touch servo system, with position feedback being generated from the pedal.

The system comprises 90 actuators, 88 for the keys and 2 for the pedals, with a servo mechanism for each actuator. In order to control this network of actuators, a three-level, hierarchically structured computer system was introduced.

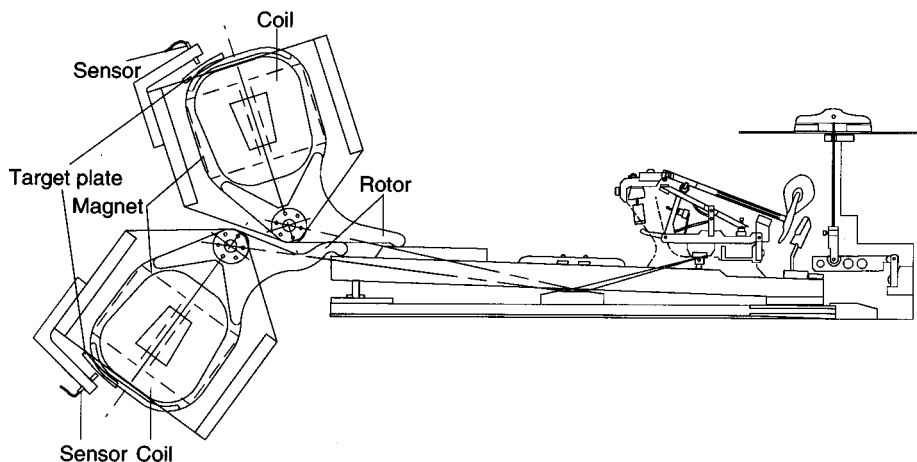


FIG. 2. Schematic view of the touch actuators on a grand piano.

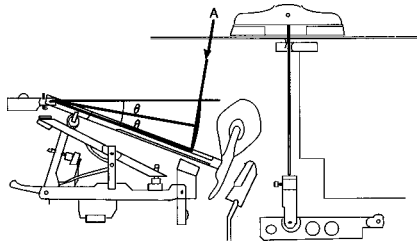


FIG. 3. View of the measurement method of the hammer motion.

At the bottom level, 90 microprocessors are employed to efficiently control the 90 actuators.

## II. MEASUREMENT OF KEY, HAMMER, AND SOUND PRESSURE

Using a laser displacement gauge (Keyence Corporation), the motions of the key and hammer were measured. Although a damper above the hammer head prevented us from directly measuring the motion of the hammer at the hammer head, we were able to align the spot of the laser with the hammer shank by directing the beam between two strings of the grand piano. The laser displacement gauge was secured as shown in Fig. 3.

The rotational motion of the hammer shank could only be approximated with a straight line using the laser displacement gauge. The motion of the hammer had a rotational angle of approximately 18 deg, and as a result there could be significant error in these straight-line measurements. Therefore, the laser was secured so that the beam could be pointed toward the hammer shank from A in Fig. 3, which is equal to the direction of the tangent at the half position  $\theta$  of the total rotational angle in Fig. 3. As a result, the maximum error for the displacement and the velocity of the hammer motion became approximately 1% and 2.5%, respectively. It was therefore possible to observe and accurately record the displacement and velocity. Using this measuring method, the displacement signal is obtained, and the velocity signal for the hammer can be simply determined by differentiation of the displacement signal.

The C-weighting sound pressure level of a source in decibels was measured using the sound level meter. The microphone of the sound level meter was secured in place 0.15 m above the strings near the rear edge of the damper.

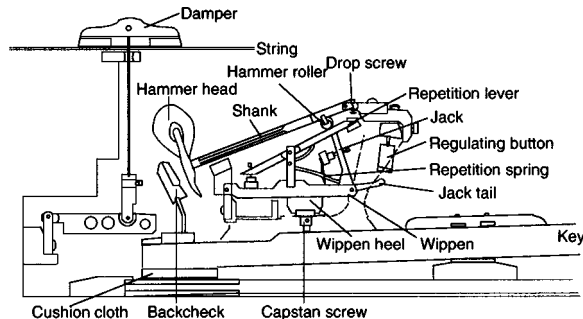


FIG. 4. View of the piano-action on a grand piano.

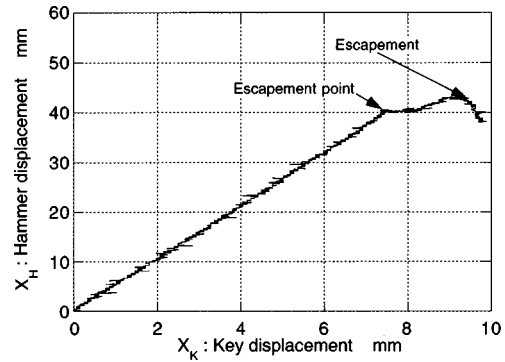


FIG. 5. Sample relationship between the displacement of key and hammer.

## III. DYNAMIC MODEL OF PIANO-ACTION

Concerning the piano-action shown in Fig. 4, when the key is depressed slowly, the wippen moves upward. In the initial movement, the repetition lever pushes up the hammer roller (attached near the rotational center of the shank) and next the jack pushes up the hammer roller, and the hammer head moves toward the string. As the key is depressed further, the jack tail reaches the regulating button (referred to as the “escapement point” in the following) as the hammer head comes within several millimeters of the string. After that, the jack is forced by the regulating button to rotate and disconnect itself from the hammer roller. This is referred to as the “escapement” in the following (see Fig. 5), and becomes the end spot for the rising motion of the hammer. If the key is depressed extremely slowly, the hammer falls on the repetition lever without striking the string. In cases where the key is struck with the usual force, the hammer will have acquired sufficient velocity at the escapement point to continue and strike the string.

The suggested dynamic model of piano-action is shown in Fig. 6. Here  $M$  is the mass of the hammer head, and  $K$  is

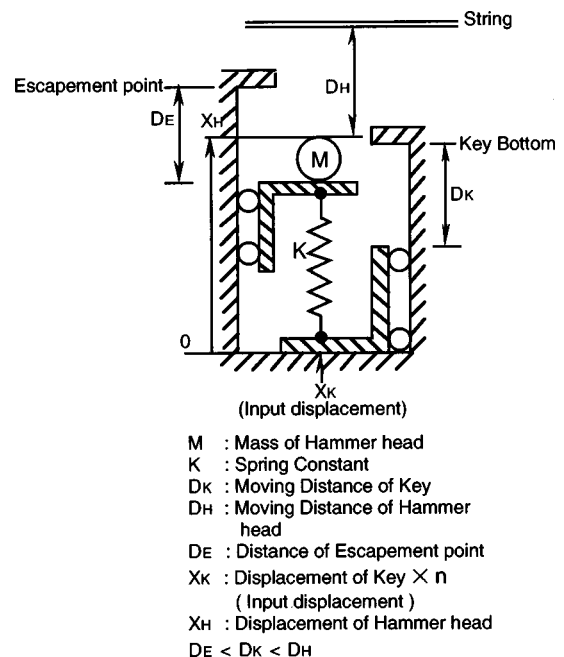


FIG. 6. Dynamic model to be analyzed for the grand piano-action.

TABLE I. Parameters of analysis.

Parameter	Unit	Value	Reference
$D_H$	m	0.048	Moving distance of hammer head
$D_K$	m	0.045	Moving distance of key $\times n$
$D_E$	m	0.042	Distance of escapement point
$\omega_1$	rad/s	81.7	Natural frequency of excitation
$\omega_2$	rad/s	113	Natural frequency of free vibration
$n$	...	5.20	Moving distance of hammer head Moving distance of key

the spring constant, which translates the elasticity of all the piano-action parts, including the elasticities of the key, the wippen, the hammer roller, and the shank, to the position of the hammer head.  $D_H$  is the distance moved from the original resting position of the hammer head to the string, and  $D_E$  is the distance moved from rest to the escapement point. The initial deflection and damping due to the influence of gravity on  $M$  were disregarded.  $D_K$  is the value obtained by multiplying the key travel distance by the lever ratio  $n$ . The maximum movement of the edge under the spring is limited by  $D_K$ .

We also assumed the following. Because the spring is not connected to  $M$ , the length of the spring does not exceed its original length in the static equilibrium position. And because the displacement of the top of the spring cannot exceed  $D_E$ ,  $M$  cannot receive the force from the spring after it reaches  $D_E$ , and will move on because of its inertia.

To begin the analysis, values for  $M$ ,  $K$ , and  $n$  were determined by means of the following survey.

If the key was moved slowly, a good proportional relationship was obtained between the displacement  $X_K$  of the key and the displacement  $X_H$  of the hammer head, so that  $n$  could be obtained from the slope (see Fig. 5).

To measure the spring of stiffness  $K$  with an experimental compression machine, the hammer was constrained, and the key was loaded at the touch-point. The relationship between the deflection and the force of the key showed that the stiffness  $K$  had nonlinear characteristics, and a single value for the stiffness of the spring could not be determined on the basis of the results obtained. Therefore, the deflection range was divided into two regions based on the relationship between the deflection and the force, and it was assumed that the spring constant exhibited piecewise-linear characteristics. The natural frequency,  $\omega_1$ , of the region of the small deformation was determined by observing the movement of the hammer in the piano. The natural frequency,  $\omega_2$ , of the re-

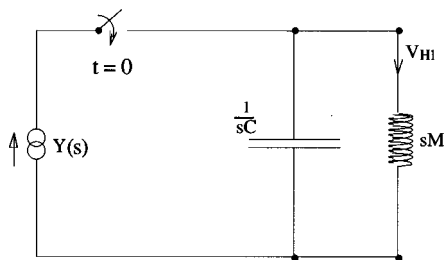
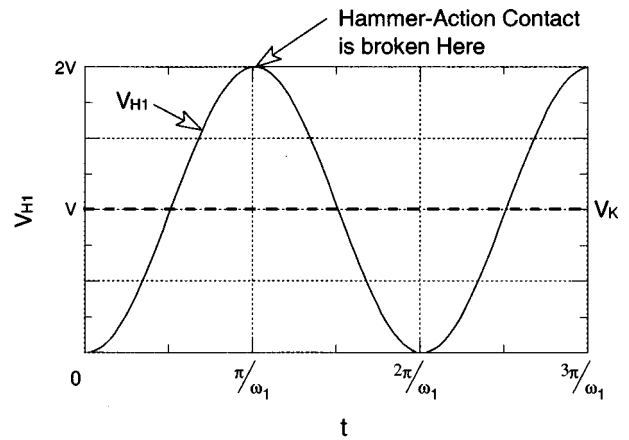
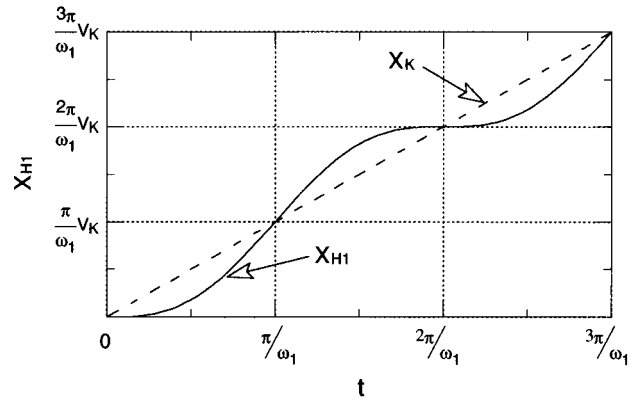


FIG. 7. Equivalent circuit of the model for which the half-period closely corresponds to the length of time that the hammer  $M$  is in contact with the spring in the model shown in Fig. 6.



(a)



(b)

FIG. 8. Velocities (a) and displacements (b) of the hammer and the key at constant speed  $V$  in Fig. 7.

gion of the large deformation was estimated from the ratio of the slopes based on a comparison of the small and large deformations obtained from both calculated and measured data.

The parameters for the calculations are indicated in Table I.

### A. Introduction to an analysis of the dynamic model

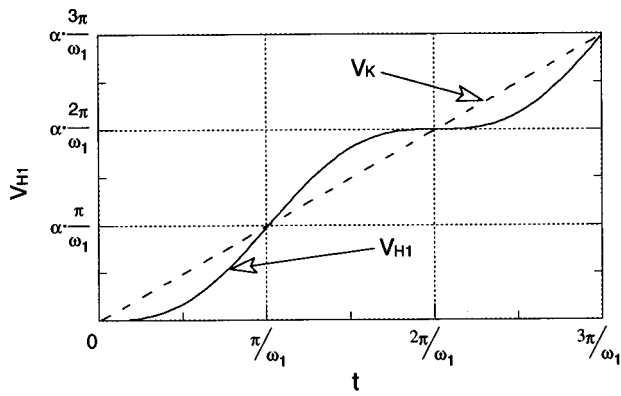
Using the equivalent electrical circuit shown in Fig. 7, for which the half-period closely corresponds to the length of time the hammer  $M$  is in contact with the spring in the model shown in Fig. 6, an application of the Laplace transform with the initial conditions  $X_H(0)=0$ ,  $V_H(0)=0$ , and the velocity function  $Y(s)$  gives the following expression for the hammer velocity  $V_{H1}(s)$ :

$$V_{H1}(s) = Y(s) \frac{1/sC}{1/sC + sM} = \frac{\omega_1^2 \cdot Y(s)}{(s^2 + \omega_1^2)}, \quad \omega_1^2 M C = 1. \quad (1)$$

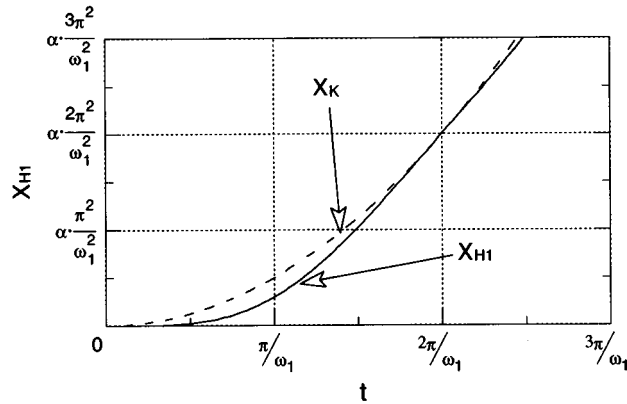
#### 1. Constant speed

In the case of a constant speed  $V$ , by substituting  $V/s$  into  $Y(s)$  of Eq. (1), the following expression for the hammer velocity  $V_{H1}(s)$  is obtained:





(a)



(b)

FIG. 9. Velocities (a) and displacements (b) of the hammer and the key at constant acceleration  $\alpha$  in Fig. 7.

$$V_{H1}(s) = \frac{V}{s} \cdot \frac{1/sC}{1/sC + sM} = \frac{\omega_1^2 \cdot V}{s(s^2 + \omega_1^2)}. \quad (2)$$

From the inverse Laplace transformation of  $V_{H1}(s)$ , the following expressions for the hammer velocity  $V_{H1}(t)$  and displacement  $X_{H1}(t)$  are obtained:

$$V_{H1}(t) = V \cdot (1 - \cos(\omega_1 \cdot t)) \quad (3)$$

and

$$X_{H1}(t) = V \cdot (t - 1/\omega_1 \cdot \sin(\omega_1 \cdot t)). \quad (4)$$

As the key starts moving with constant speed  $V$  at  $t=0$  (see Fig. 8), the compression of the spring introduces a delay into the displacement  $X_H$  of  $M$  compared to the displacement  $X_K$  of the input waveform. After half a period of the natural frequency at  $t=\pi/\omega_1$ , the spring recovers its original length, and the delay momentarily decreases to zero. At this moment,  $M$  has moved the same distance as the input distance  $X_K$  and has reached its maximum velocity,  $2V$ . The spring then extends, exerting a force on the mass, moving  $M$  more than the input distance  $X_K$ . After a period of the natural frequency at  $t=2\pi/\omega_1$ , the spring again recovers its original length and the delay momentarily decreases to zero.

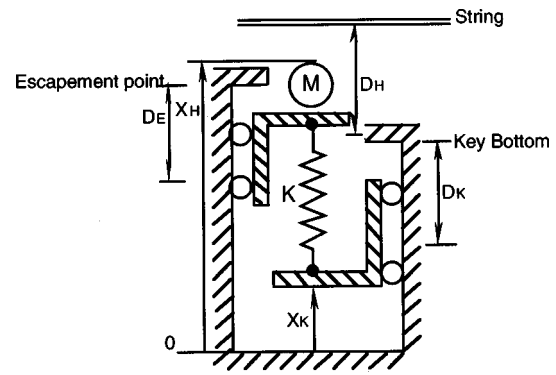


FIG. 10. Behavior of model in Fig. 6 for group I at constant speed.  $M$  breaks contact with the top of spring after a time of half the natural period has elapsed.

## 2. Constant acceleration

In the case of constant acceleration  $\alpha$ , by substituting  $\alpha/s^2$  into  $Y(s)$  of Eq. (1), the following expression for the hammer velocity  $V_{H1}(s)$  is obtained:

$$V_{H1}(s) = \frac{\alpha}{s^2} \cdot \frac{1/sC}{1/sC + sM} = \frac{\omega_1^2 \cdot \alpha}{s^2(s^2 + \omega_1^2)}. \quad (5)$$

From the inverse Laplace transformation of  $V_{H1}(s)$ , the following expressions for the hammer velocity  $V_{H1}(t)$  and displacement  $X_{H1}(t)$  are obtained:

$$V_{H1}(t) = \alpha \cdot t - \frac{\alpha}{\omega_1} \cdot \sin(\omega_1 \cdot t) \quad (6)$$

and

$$X_{H1}(t) = \frac{\alpha \cdot t^2}{2} - \frac{\alpha}{\omega_1^2} \cdot (1 - \cos(\omega_1 \cdot t)). \quad (7)$$

As the key starts moving with constant acceleration  $\alpha$  at  $t=0$  (see Fig. 9), the compression of the spring also introduces a delay into  $X_{H1}$  compared to  $X_K$ . After a period of the natural frequency at  $t=2\pi/\omega_1$ , the spring recovers its original length and the delay momentarily decreases to zero. At this moment  $M$  has moved the same distance and reached the same velocity as the input distance  $X_K$  and the input velocity  $V_K$ . Then, because the velocity  $V_{H1}$  becomes smaller than the input velocity  $V_K$ , the spring is again compressed.

## IV. CONSTANT SPEED

In the case of a constant speed input, to differentiate the behavior of the hammer from the driving speed of the key, the behavior has been broken into three speed groups for analysis.

### A. Constant speed group I

Using a slow constant speed starting at  $t=0$ , the motion of the model was observed and analyzed. The initial motion of  $M$  exhibits a delay from the input displacement of  $X_K$ . At the moment when half the period of the natural oscillation has passed, the delay becomes zero, and the spring returns to its original length. At this time  $M$  breaks contact with the

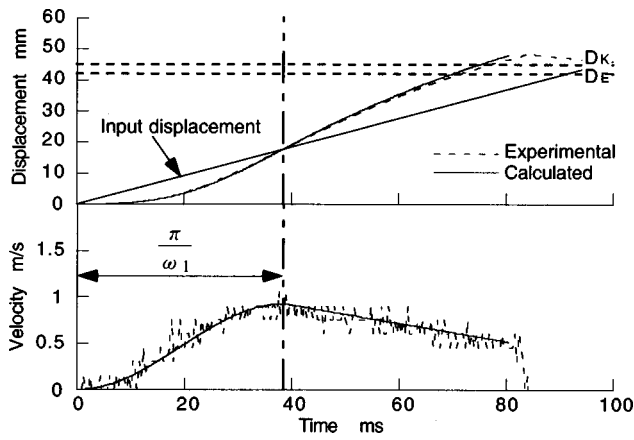


FIG. 11. General histories of displacement and velocity of the hammer measured and calculated for group I.

spring and is thrown up at a velocity twice the input speed (see Fig. 10). After this moment, while the velocity decreases due to the influence of gravity,  $M$  goes up and then strikes the string.

The displacement and velocity of  $M$  until  $t < t_A$ , where  $t_A = \pi/\omega_1$  is half the period of the natural frequency, can be expressed by the equation

$$X_{H1}(t) = V \cdot (t - 1/\omega_1 \cdot \sin(\omega_1 \cdot t)), \quad (8)$$

$$V_{H1}(t) = V \cdot (1 - \cos(\omega_1 \cdot t)), \quad (9)$$

where  $V$  is the input speed of the base excitation and  $\omega_1$  is the natural frequency. Because the motion of the hammer becomes parabolic for  $t > t_A$ , the displacement and velocity of  $M$  from  $t_A$  to the time of the string-striking is given by the equation

$$X_{H2}(t) = V_{H1} \cdot (t - t_A) - g \cdot (t - t_A)^2/2 + X_{H1}, \quad (10)$$

$$V_{H2}(t) = V_{H1} - g \cdot (t - t_A), \quad (11)$$

where  $g$  is the acceleration of gravity. Because the motion after the time  $t_A$  is parabolic motion, when the peak of the parabolic motion is equal to the string position, the input speed can be defined as the minimum input speed. With  $t = t_A$ , Eqs. (8) and (9) are reduced, and the obtained solutions can be substituted into Eqs. (10) and (11), respectively. Therefore, reducing Eqs. (10) and (11) with  $X_{H2} = D_H$  and  $V_{H2} = 0$ , the minimum input speed can be expressed by the equation

$$V_{K\min} = \frac{V_{\min}}{n} = \frac{\sqrt{(g \cdot \pi/\omega_1)^2 + 8 \cdot D_H \cdot g} - g \cdot \pi/\omega_1}{4 \cdot n}, \quad (12)$$

where  $V_{K\min}$  is the minimum input speed for the key, reduced to the position of the touch-point. In this case,  $M$  comes into contact with the string; however, because the string-striking velocity of  $M$  is zero, no sound is produced.

The general motion of the hammer in group I is shown in Fig. 11. The data obtained, i.e., the displacement and velocity of the hammer, is in good agreement with the calculated theoretical values and the demonstrated values.

In the case of group I, because the hammer breaks contact with the other parts before it reaches the escapement point, escapement does not occur.

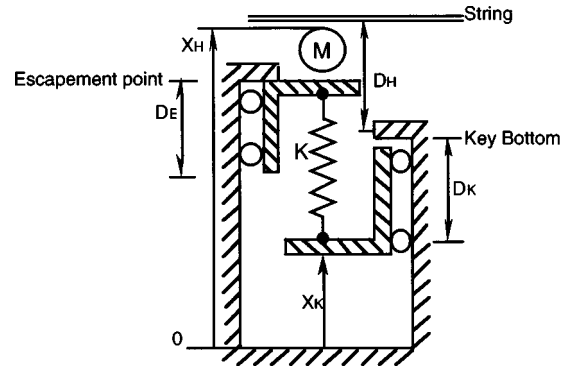


FIG. 12. Behavior of model in Fig. 6 for group II at constant speed.  $M$  breaks contact with the top of the spring at the escapement point.

## B. Constant speed group II

The speed range ( $V_{KE} < V_{KD}$ ) in this group is greater than the speed of group I and can be expressed in terms of the following quantities:

$$V_{KE} = \frac{D_E \cdot \omega_1}{n \cdot \pi}, \quad (13)$$

$$V_{KD} = \frac{D_K \cdot \omega_1}{n \cdot \pi}, \quad (14)$$

where  $V_{KE}$  and  $V_{KD}$  are the speeds of the touch-point of the key that the top and bottom of the spring respectively reach at the escapement point, when a half-period of the natural frequency has elapsed (see Fig. 12).

The motion of the hammer for  $t < t_E$ , where  $t_E$  is the time elapsed until the hammer reaches the escapement point, is similar to that described by Eqs. (8) and (9). As shown in Fig. 11, for  $t > t_E$ , the motion of the hammer becomes parabolic with the initial conditions of a displacement  $D_E$  and  $V_H = V_{H1}(t_E)$ . Hence, the response can be expressed by equations that are similar to Eqs. (10) and (11):

$$X_{H3}(t) = V_H \cdot (t - t_E) - g \cdot (t - t_E)^2/2 + D_E, \quad (15)$$

$$V_{H3}(t) = V_H - g \cdot (t - t_E). \quad (16)$$

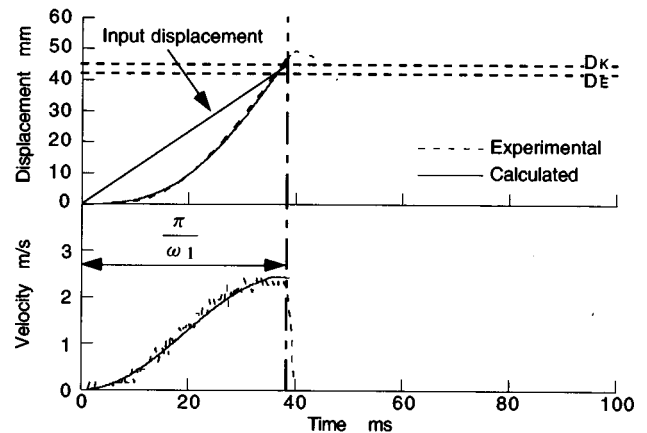


FIG. 13. General histories of displacement and velocity of the hammer measured and calculated for group II.

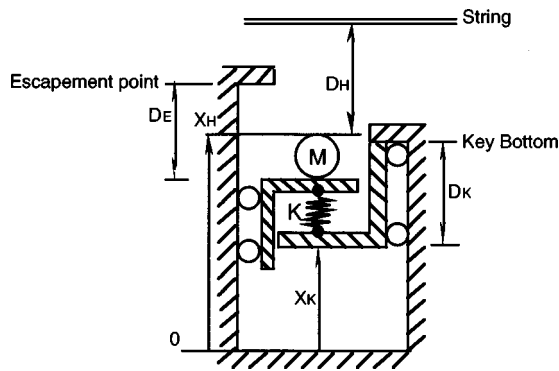


FIG. 14. Behavior of model in Fig. 6 for group III at constant speed. A large applied force for input is excited for the model in Fig. 6. Before  $M$  starts the motion, the spring deforms maximally.

In addition, to make the hammer reach the string in group II, the minimum velocity of the hammer at the escapement point can be expressed in terms of the following quantities:

$$V_{Hmin} = \sqrt{2 \cdot g \cdot (D_H - D_E)}. \quad (17)$$

The general motion of the hammer in group II is shown in Fig. 13. The displacement and velocity of the hammer, too, are in good agreement with the calculated theoretical values and the demonstrated values.

Because the motion of the hammer after it passes the escapement point is extraneous to the driving input waveforms, we will only discuss the motion of the hammer until it reaches the escapement point in the following.

### C. Constant speed group III

The speed in this group is greater than the speed of group II, in that the bottom of the spring reaches  $D_K$  (i.e., the key reaches bottom) before the top of the spring reaches the escapement point (see Fig. 14). At this moment the compression of the spring increases with increases in the speed.  $M$  is then pushed toward the string up to the escapement point by a force that is due only to the compression of the spring.

The motion of the hammer for  $t < t_b$ , where  $t_b$  is the time elapsed until the bottom of the spring reaches  $D_K$ , can be expressed by Eqs. (8) and (9). In the group III case for

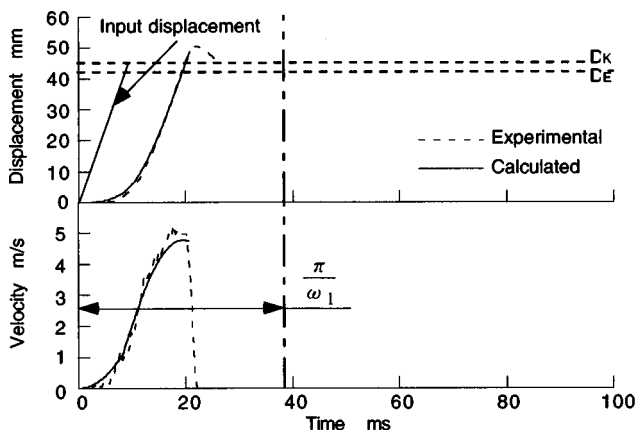


FIG. 15. General histories of displacement and velocity of the hammer measured and calculated for group III.

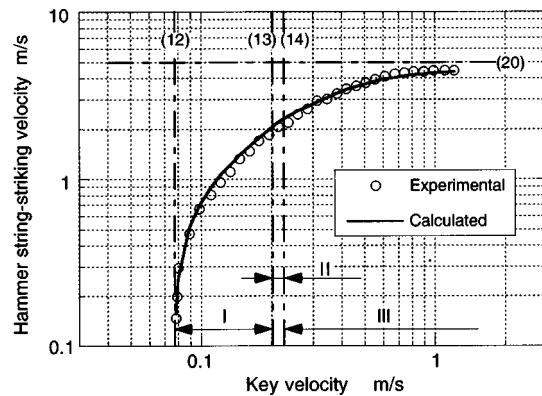


FIG. 16. Relationship between the key speed and the hammer string-striking velocity. The numerical values in parentheses indicate the numbers of equations in the main text.

constant velocity and the group II case for constant acceleration described in Sec. VI B, because the greater deformation results in increases in speed, the natural frequency  $\omega_2$  must be used. A comparison of the motion of  $M$  until the bottom of the spring reaches  $D_K$  was made between calculated and measured data, and the result shows a gross error in the calculated data for  $\omega_2$  compared to the measured data (data not shown). Hence, the natural frequency of the motion until the bottom of the spring reaches  $D_K$  in these group cases was determined to be  $\omega_1$  based on a comparison between calculated and measured data.

After this moment,  $M$  begins to oscillate freely.  $M$  is pushed up toward the string by the spring force due to the initial displacement  $[D_K - X_{H1}(t_b)]$ . The displacement and velocity of  $M$  with the initial conditions of displacement  $[D_K - X_{H1}(t_b)]$  and velocity  $V_{H1}(t_b)$  can be given by the equations

$$X_{H4}(t) = D_K - ((D_K - X_{H1}(t_b)) \cdot \cos(\omega_2 \cdot (t - t_b)) - V_{H1}(t_b) / \omega_2 \cdot \sin(\omega_2 \cdot (t - t_b))), \quad (18)$$

$$V_{H4}(t) = (D_K - X_{H1}(t_b)) \cdot \omega_2 \cdot \sin(\omega_2 \cdot (t - t_b)) + V_{H1}(t_b) \cdot \cos(\omega_2 \cdot (t - t_b)). \quad (19)$$

If the speed becomes infinity, the string-striking velocity of the hammer converging to  $V_{Hammer}$  can be expressed as the equation

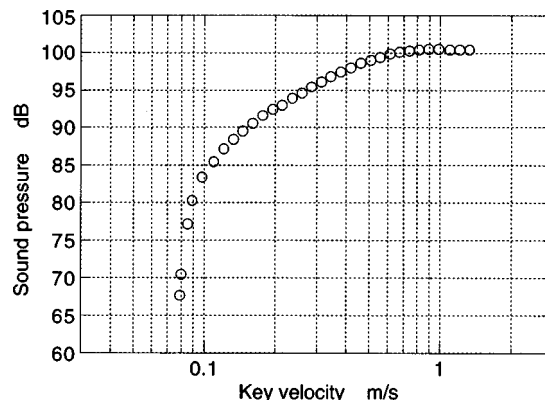


FIG. 17. Relationship between the key speed and the sound pressure.

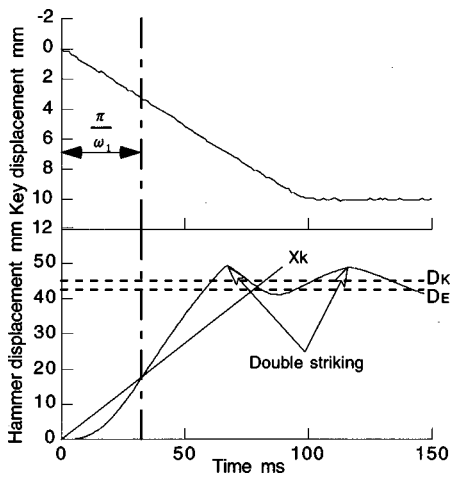


FIG. 18. General histories of the displacement of the hammer for "double-striking."

$$V_{\text{Hammer}} \doteq \sqrt{(D_K \cdot \omega_2)^2 - 2 \cdot g \cdot (D_H - D_E)}. \quad (20)$$

The quantity in this equation corresponds to the hammer velocity when the key reaches  $D_K$  before the hammer goes into motion, as shown in Fig. 14, and the hammer is pushed up only by the spring force due to displacement. This value is independent of the input waveform.

The general motion of the hammer in group III is shown in Fig. 15. The displacement and velocity of the hammer, too, are in good agreement with the calculated theoretical values and the demonstrated values.

#### D. Results and discussion

The relationship between key speed and the string-striking velocity as well as key speed and the sound pressure are shown in Figs. 16 and 17, respectively. The group I case is applicable for constant input speeds  $V$  from very low values up to about 0.2 m/s. At this upper limit, the striking velocity of the hammer is about 2 m/s, corresponding to mezzo forte.

In the case of group I, because the gradient for the driving speed of the key is great, the string-striking velocity and the sound pressure vary greatly with the input speed. Therefore, if the physical properties of the piano-action also vary with environmental changes in temperature and humidity, it appears that producing a soft tone is even more difficult.

In addition, for the group-I-type touch, the hammer breaks contact with the other parts before it reaches the escapement point (see Fig. 18). After it strikes, it is pushed up by the other parts, moving at a constant speed, and strikes the string again. In what follows, this phenomenon is referred to as "double-striking." This double-striking phenomenon occurs only in response to group-I-type touch at constant speed in response to the touch actuator.

#### V. MODIFIED CONSTANT SPEED

The double-striking problem as well as the other problems described above are caused by the hammer breaking contact with the other parts at an early stage, before it reaches the escapement point. To solve this problem, the

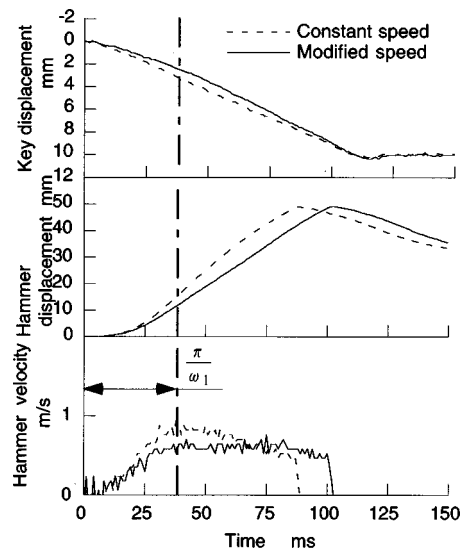


FIG. 19. Comparison of constant speed and the modified speed.

method described below was developed. The speed ranges of groups I, II, and III are similar to the respective groups discussed in Sec. IV.

#### A. Modified constant speed group I

In the model shown in Fig. 6, the input speed is modified so that it is equal to the velocity of  $M$  when a half-period of the natural oscillation has elapsed. As a result,  $M$  maintains contact with the top of the spring and can move at a constant speed to the escapement point. The motion of  $M$  up until half the natural period and after the escapement point is given by Eqs. (8) and (9) and Eqs. (10) and (11), respectively. The relationship between the input speed  $V_S$  and the string-striking velocity  $V_{HC}$  of  $M$  is expressed by the equation

$$V_{HC} = \sqrt{4 \cdot V_S^2 - 2 \cdot g \cdot (D_H - D_E)}. \quad (21)$$

Because the minimum input speed  $V_{KCmin}$  becomes  $\frac{1}{2}$  of  $V_{Hmin}$ , which is given by Eq. (17), we have the relationship

$$V_{KCmin} = \frac{V_{Hmin}}{2 \cdot n}. \quad (22)$$

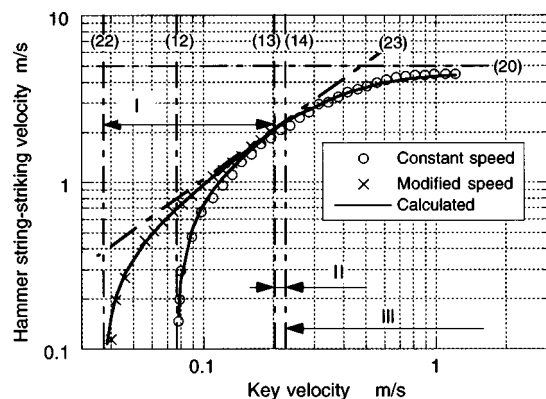


FIG. 20. Relationship between the key speed and the hammer string-striking velocity. The numerical values in parentheses indicate the numbers of equations in the main text.

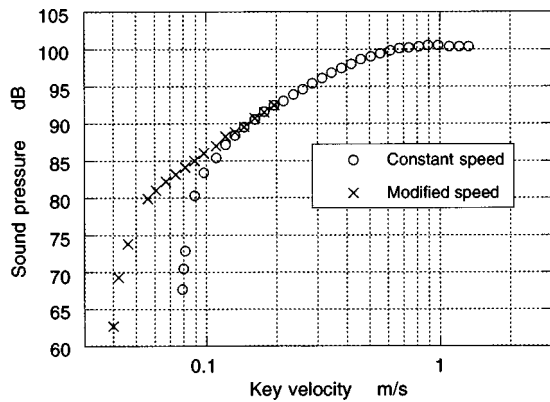


FIG. 21. Relationship between the key speed and the sound pressure.

All of the hammer motions were observed and recorded so that the string-striking velocities for the constant speed and for the modified speed were equal. These results are shown in Fig. 19.

### B. Modified constant speed groups II and III

It is not necessary for the speed to change in these groups because the top of the spring reaches the escapement point when a half-period of the natural oscillation has elapsed, and  $M$  breaks contact with the other parts as the top of the spring reaches the escapement point. As a result, the motion is similar to that described in Secs. IV B and C.

### C. Results and discussion

The effects of changing the input speed on the string-striking velocity and sound pressure are shown in Figs. 20 and 21, respectively. The equation for the dashed line 16 is

$$V_{ZH} = 2 \cdot V_Z, \quad (23)$$

where  $V_Z$  is the input speed, and  $V_{ZH}$  is the string-striking velocity of the hammer. Therefore, because of the distance  $(D_H - D_E)$ , the values for the string-striking velocity cannot be above the dashed line.

The problems described above that occur with a constant speed for group I, i.e., the large gradient for the input speed and the “double-striking,” are solved by changing the input speed, and a stable soft tone can thus be produced.

## VI. CONSTANT ACCELERATION

As with the effects of constant speed described in the previous section, the constant acceleration for the input waveform also changes the behavior of the hammer in response to the driving acceleration of the key. As a result, this behavior can also be broken into two acceleration groups and analyzed.

### A. Constant acceleration group I

Using a constant acceleration starting at  $t=0$ , the motion of the model was observed and analyzed. First, the initial motion of  $M$  exhibits a delay from the input displacement of  $X_K$ . The delay becomes maximum at half the natural oscillation period. The spring returns to its original length after a

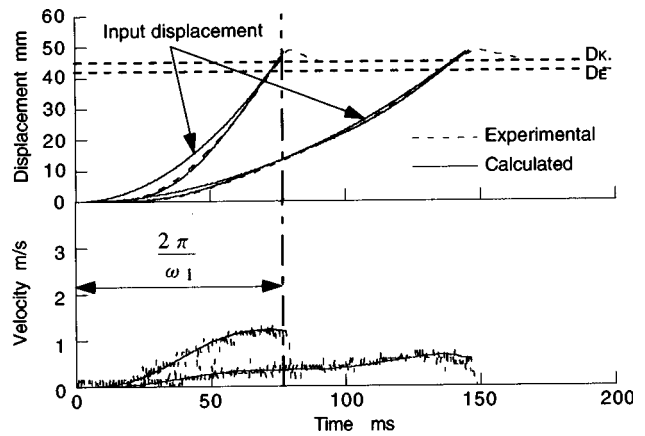


FIG. 22. General histories of the displacement and velocity of the hammer measured and calculated for group I.

full period of the natural oscillation has elapsed. Because the velocity of  $M$  is equal to the velocity of the input,  $M$  does not break contact with the other parts before the escapement point as in the case of a constant input speed. After the first period of the natural oscillation,  $M$  repeats the periodic motion. The motion of  $M$ , when it reaches the escapement point, becomes parabolic with an initial condition  $\dot{x}$ , which is the velocity of  $M$  at the escapement point. The displacement  $X_{H1}$  and the velocity  $V_{H1}$  of  $M$  before it reaches the escapement point can be expressed by the equations

$$X_{H1}(t) = \frac{\alpha \cdot t^2}{2} - \frac{\alpha}{\omega_1^2} \cdot (1 - \cos(\omega_1 \cdot t)), \quad (24)$$

$$V_{H1}(t) = \alpha \cdot t - \frac{\alpha}{\omega_1} \cdot \sin(\omega_1 \cdot t), \quad (25)$$

where  $\alpha$  is the input acceleration.

By setting  $V_{H\min}$  of Eq. (17) equal to the right side of Eq. (25), the minimum input acceleration  $\alpha_{K\min}$  can be expressed by the equation

$$\alpha_{K\min} = \frac{V_{H\min}}{(t - (\sin(\omega_1 \cdot t)/\omega_1)) \cdot n}. \quad (26)$$

Then, by substituting the right side of Eq. (26) into  $\alpha$  of Eq. (24) from  $D_E$ =the right side of Eq. (24),

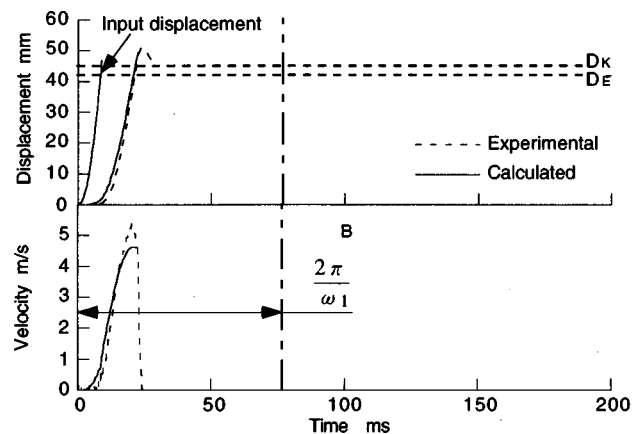


FIG. 23. General histories of the displacement and velocity of the hammer measured and calculated for group II.

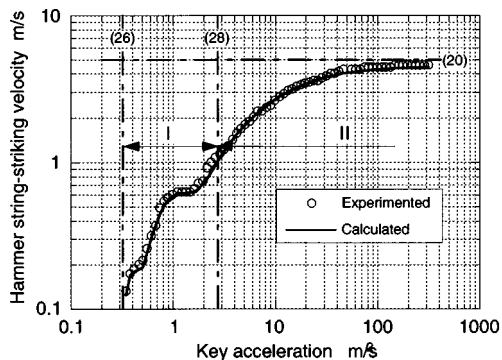


FIG. 24. Relationship between the key acceleration and the hammer string-striking velocity. The numerical values in parentheses indicate the numbers of equations in the main text.

$$D_E = \frac{V_{H\min} \cdot t^2}{2 \cdot (t - (\sin(\omega_1 \cdot t) / \omega_1)) \cdot n} - \frac{V_{H\min}}{(t - (\sin(\omega_1 \cdot t) / \omega_1)) \cdot n \cdot \omega_1^2} (1 - \cos(\omega_1 \cdot t)). \quad (27)$$

Reducing Eq. (27) by repeating the calculation with  $t = 0$ , the minimum acceleration for the input can be obtained by substituting the solution of  $t$  into Eq. (26).

When the top of the spring reaches the escapement point after one period of the natural oscillation has elapsed, the acceleration  $\alpha_{KE}$  reaches a maximum in this group and can be expressed by the equation

$$\alpha_{KE} = \frac{D_E \cdot \omega_1^2}{2 \cdot n \cdot \pi^2}. \quad (28)$$

The motion after the escapement point is similar to that described in previous sections.

The general motion of the hammer in group I is shown in Fig. 22. The calculated theoretical values and the demonstrated values for the displacement and velocity of the hammer are in good agreement.

## B. Constant acceleration group II

The acceleration in this group is greater than the acceleration of group I, and in the motion of this group the bottom

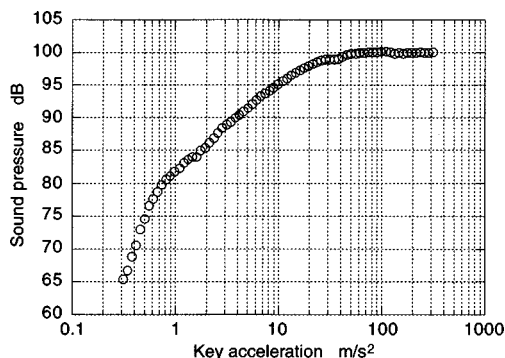


FIG. 25. Relationship between the key acceleration and the sound pressure.

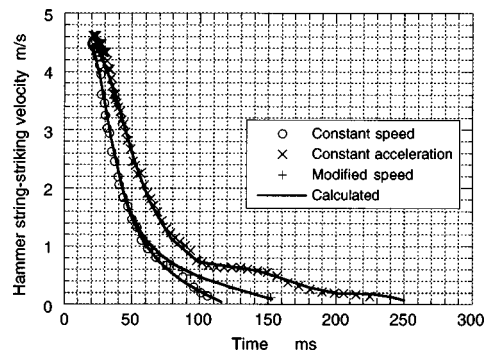


FIG. 26. Measured and calculated relationship between the hammer string-striking velocity and the interval between the initial input and the time at which the string was struck.

of the spring reaches  $Dk$  (i. e., the key reaches the key bottom) before the top of the spring reaches the escapement point.

The motion of the hammer for  $t < t_c$ , where  $t_c$  is the time elapsed until the bottom of the spring reaches  $D_K$ , can be expressed by Eqs. (24) and (25), because the motion is similar to that described in Sec. IV C. After this moment,  $M$  begins to oscillate freely.  $M$  is pushed up toward the string by the spring force due to the initial displacement  $[D_K - X_{H1}(t_c)]$ . The displacement and velocity of  $M$  with the initial conditions of displacement  $[D_K - X_{H1}(t_c)]$  and the velocity  $V_{H1}(t_c)$  can be given by the equations

$$X_{H5}(t) = D_K - ((D_K - X_{H1}) \cos(\omega_2 \cdot (t - t_c)) - V_{H1} / \omega_2 \cdot \sin(\omega_2 \cdot (t - t_c))), \quad (29)$$

$$V_{H5}(t) = (D_K - X_{H1}) \cdot \omega_2 \cdot \sin(\omega_2 \cdot (t - t_c)) + V_{H1} \cdot \cos(\omega_2 \cdot (t - t_c)). \quad (30)$$

The motion after  $M$  reaches the escapement point can be expressed by Eqs. (15) and (16).

The general motion of the hammer for group II is shown in Fig. 23. The calculated theoretical values and the demon-

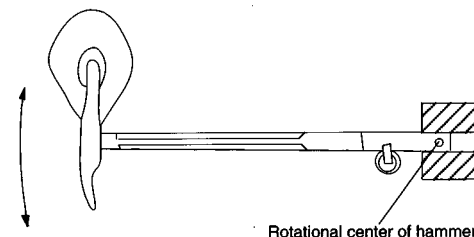


FIG. 27. Free vibration of hammer clamped at center of rotation.

strated values for the displacement and velocity of the hammer are in good agreement.

### C. Results and discussion

The relationship between key acceleration and string-striking velocity as well as key acceleration and sound pressure are shown in Figs. 24 and 25, respectively. In group I, the curves for both the string-striking velocity versus key acceleration and the sound pressure versus key acceleration cause a waviness. This phenomenon is caused by the hammer velocity being slower than the input velocity at the escapement point for the odd-numbered half periods of the natural oscillation. The group I case is applicable for constant input accelerations  $\alpha$  from very low values up to about  $3 \text{ m/s}^2$ . At this upper limit the striking velocity of the hammer is about  $1 \text{ m/s}$ , corresponding to mezzo forte.

Because the hammer does not break contact with the other parts at an early stage in group I, a stable soft tone can be produced. As shown in Fig. 26, however, with a constant acceleration, the interval between the initial input and the time at which the string is struck will be longer for a soft tone than with other waveforms.<sup>4</sup>

### VII. CONCLUSION

A dynamic model of a simulated hammer motion was presented in this article. A motion analysis was carried out, with various input waveforms observed and recorded. The resulting data are in good agreement with the calculated theoretical values and demonstrate that the model used for the

analysis is valid, making the behavior of the hammer clear. The change in speed necessary to produce the optimum touch input waveform has now been established.

In the case of the piano-action described in this article, the  $\omega_1$  and  $\omega_2$  natural frequencies used in this analysis were 81.7 and 113 rad/s, respectively. When the hammer shank was clamped, as shown in Fig. 27, its natural frequency  $f$  was found to be approximately 440 rad/s. Based on this value, it can be predicted that the natural frequencies of the individual wooden parts that make up the action, except for the felt found on the wippen heel and the hammer roller, are greater than the natural frequencies  $\omega_1$  and  $\omega_2$ . Further, the natural frequency of the piano-action (except for the felt), estimated by a series-type synthetic method, will be greater than the natural frequency of  $\omega_1$  and  $\omega_2$ . This indicates that the stiffness of the felt has a significant effect on the physical values of the piano-action.

<sup>1</sup>W. B. White, "The human element in piano tone production," *J. Acoust. Soc. Am.* **1**, 357–367 (1930).

<sup>2</sup>H. C. Hart, M. W. Fuller, and W. S. Lusby, "A precision study of piano touch and tone," *J. Acoust. Soc. Am.* **6**, 80–94 (1934).

<sup>3</sup>A. Askenfelt and E. V. Jansson, "From touch to string vibration. I: Timing in the grand piano action," *J. Acoust. Soc. Am.* **88**, 52–63 (1990).

<sup>4</sup>A. Askenfelt and E. V. Jansson, "From touch to string vibration. II: The motion of the key and hammer," *J. Acoust. Soc. Am.* **90**, 2383–2393 (1991).

<sup>5</sup>E. Hayashi, M. Yamane, and H. Mori, "Development of Moving Coil Actuator for an Automatic Piano," *Int. J. Jpn. Soc. Precis. Eng.* **28**(2), 164–169 (1994).

<sup>6</sup>E. Hayashi, T. Ishikawa, K. Yamamoto, M. Yamane, and H. Mori, "Development of a Piano Player," *International Computer Music Conference* (1993), pp. 426–427.

# Time-domain modeling and numerical simulation of a kettledrum

Leïla Rhaouti

INRIA, Domaine de Voluceau-Rocquencourt, BP 105, 78153 Le Chesnay Cedex, France

Antoine Chaigne<sup>a)</sup>

Ecole Nationale Supérieure des Télécommunications, Département TSI, CNRS URA 820, 46 Rue Barrault, 75634 Paris Cedex 13, France

Patrick Joly

INRIA, Domaine de Voluceau-Rocquencourt, BP 105, 78153 Le Chesnay Cedex, France

(Received 18 September 1998; accepted for publication 26 February 1999)

A kettledrum is made of a circular elastic membrane stretched over an enclosed air cavity. It is set into vibration by the impact of the mallet. The motion of the membrane is coupled with both the external and internal sound field. A time-domain modeling of this instrument is presented which describes the motion of the mallet and its nonlinear interaction with the membrane, the transverse displacement of the membrane, and the sound pressure inside and outside the cavity. Based on a variational formulation of the problem, which uses the pressure jump over the boundaries of the instrument as a new variable, a numerical scheme is derived by means of three-dimensional finite elements. Higher-order absorbing conditions are used to simulate the free space. The validity of the model is illustrated by successive snapshots showing the pressure fields and the displacement of the membrane. In addition, time histories of energetic quantities help in explaining how the energy is balanced between mallet, membrane, and acoustic field in real instruments. Simulated external pressures show particularly good agreement with the sound field radiated by real instruments in both time and frequency domains. © 1999 Acoustical Society of America. [S0001-4966(99)04806-7]

PACS numbers: 43.75.Hi, 43.40.Dx [WJS]

## LIST OF SYMBOLS

### Mallet

$u(t)$  position of the center of gravity of the mallet ( $z$  component)

$v_o$  initial velocity of the mallet (in  $z$  direction)

$\delta$  initial position of the mallet's center of gravity (at rest)

$m$  effective mass of the mallet

$K$  coefficient of mallet stiffness

$\alpha$  stiffness nonlinear exponent

### Interaction force

$F(t)$  interaction force between mallet and membrane

$f(t)$  force density

$(x_0, y_0)$  impact point on the membrane

$g(x, y)$  spatial window

$W(t)$  mean displacement of membrane's area in contact with the mallet

### Membrane

$\Sigma$  surface of the membrane

$a$  membrane radius

$(x, y, t)$  point on  $\Sigma$  at time  $t$

$w(x, y, t)$  transverse displacement of the membrane

$\sigma$  area density of the membrane (in  $\text{kg/m}^2$ )

$T$  membrane tension (in N/m)

$\eta$  viscoelastic damping coefficient

$[p]_{|\Sigma}(x, y, t)$  pressure jump on the membrane

### Air

$\Omega_e$  external space (outside the cavity)

$\Omega_i$  internal space (inside the cavity)

$(x, y, z, t)$  point in  $\mathbb{R}^3$  ( $=\Omega_e \cup \Omega_i$ ) at time  $t$

$p_j(x, y, z, t)$  pressure in  $\Omega_j$  for  $j=e, i$

$\mathbf{V}_j(x, y, z, t)$  acoustic velocity in  $\Omega_j$  for  $j=e, i$

$c_a$  speed of sound in air

$\rho_a$  air density

### Kettledrum

$\mathcal{C}$  surface of the kettle

$H$  height of the kettle

$V$  volume of the kettle

$\Gamma = \Sigma \cup \mathcal{C}$  surface of the kettledrum

$\mathbf{n}$  exterior unit vector normal to  $\Gamma$

## INTRODUCTION

The kettledrums, or timpani, are percussive instruments made of a thin circular membrane stretched over an air cav-

ity. The membrane is set into vibration by the impact of a mallet. The sound of such instruments can be roughly divided into two parts: the attack, or initial transient, where the membrane strongly interacts with the mallet, and the decay, which is made of the relatively slow decrease of the free vibrations of the system.<sup>1-3</sup>

<sup>a)</sup> Author to whom correspondence should be addressed; Electronic mail: chaigne@sig.enst.fr



The vibrations of a circular membrane *in vacuo* have been extensively investigated in the past. The mathematical solutions for the eigenmodes and the eigenfrequencies can be found in many textbooks.<sup>4,5</sup> A model of the acoustics of timpani has been studied in the frequency domain by Christian *et al.* by computing the eigenfrequencies of a baffled membrane coupled to the exterior fluid and to a cylindrical attached cavity.<sup>6</sup> The results obtained show to what extent the presence of both external and internal air loading shifts the eigenfrequencies of the membrane. In the time-domain, a previous work by the second author of this paper was limited to the modeling of the interaction between mallets and a membrane *in vacuo*.<sup>7</sup> Other previous studies related to the kettledrums are more specifically devoted to the kettle-shape dependence of timpani normal modes,<sup>8</sup> to the properties of mallets used in percussion instruments<sup>9</sup> and to the accurate signal analysis of timpani sounds.<sup>10</sup>

However, a number of physical problems related to timpani sound production still remain unsolved. No satisfactory model exists, for example, for the evolution with time of timpani sounds and vibrations, starting from the nonlinear initial impact, and taking the interaction of the unbaffled membrane with external air and cavity into account.

It is a well-known fact that the quality of sound produced by timpani is highly dependent on the various damping mechanisms in the instrument. These mechanisms are essentially due to losses in the mallet's felt and in the membrane's material, to absorption at the membrane's boundaries, and to acoustic radiation. It turns out that these losses are very difficult to measure and to discriminate experimentally, and thus, a simulation program is of major help for separately investigating each damping term. Such a program could also be used for optimizing cavity volume and shape, which should be of interest for instrument makers.

The purpose of this study is to present a time-domain numerical formulation for the vibroacoustics of timpani. The governing equations of the model are presented in Sec. I. This model includes the motion of the mallet, the nonlinear interaction between both mallet and membrane, and the transverse vibrations of the membrane coupled with both the external and internal acoustic field. The problem is complemented by initial and boundary conditions for the mallet, the membrane, the kettle, and the sound field, respectively.

In order to solve these equations numerically, the problem is first rewritten by using an equivalent variational formulation. In this procedure a new variable is introduced for convenience: the pressure jump over the boundaries of the instrument. A numerical scheme is then derived by means of three-dimensional finite elements. For space discretization, the acoustic pressure and the acoustic velocity are discretized on a regular volume mesh. The pressure jump is discretized on a triangular mesh for the surface  $\Gamma = \Sigma \cup \mathcal{C}$  of the kettledrum and the displacement of the membrane is discretized on another triangular mesh for  $\Sigma$ . For time discretization, a classical centered finite differences scheme is used. Higher-order absorbing conditions are implemented to simulate the free space.<sup>11</sup> The numerical formulation of the problem is presented in Sec. II.

In Sec. III, the model is first tested by a comparison

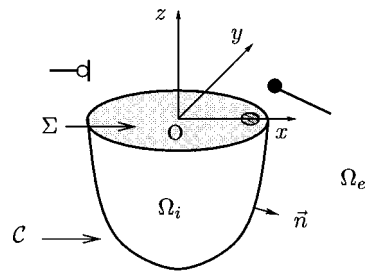


FIG. 1. Geometry of the kettledrum. The origin  $O$  of the coordinate system lies in the center of the membrane.  $\Omega_i$  is the cavity domain bounded by the surface  $\mathcal{C}$  of the bowl and the surface  $\Sigma$  of the membrane.  $\Omega_e$  is the external free space. The mallet strikes the membrane at position  $(x_0, y_0)$ .

between numerical and analytical results, in the simple case of a membrane *in vacuo*. Other simulations, including external air loading and cavity, are used for displaying successive snapshots of the pressure, inside and outside the cavity, together with the displacement of the membrane. It will be shown, also, to what extent an energy approach helps in understanding the energy exchange between the mallet, the membrane, and the acoustic field during the evolution of sound.

Section IV is devoted to the comparison between measured and simulated radiated sound pressure. Comparisons are made in both time and frequency domains by examining similarities and differences in waveforms, eigenfrequencies, and decay times. The effect of improper tensioning of the membrane, in connection with the appearance of characteristic beats, is also examined.

One difficulty of the comparison between simulations and experiments follows from the necessity of inserting accurate values of geometrical and physical parameters into the model. Therefore, a discussion on the method used for extracting the parameters from measured signals is also provided in this section, with particular emphasis on the tension of the membrane, on the damping coefficient, and on the mallet's model.

## I. GOVERNING EQUATIONS

The vibroacoustics of kettledrums can be described by a set of coupled equations that govern the motion of the mallet, the propagation of elastic waves in the membrane, the internal pressure in the cavity, and the external pressure in the surrounding fluid, respectively.

The motion of the mallet is described by the displacement  $u(t)$  of its center of gravity. It starts at time  $t=0$ , when the felt just reaches the membrane with initial velocity  $v_0$ , regardless of how the mallet has been set into motion by the player before the origin of time. The initial conditions for the mallet are, at this time, according to the orientation of the  $z$ -axis (see Fig. 1),

$$u(0) = \delta, \quad \frac{du}{dt}(0) = -v_0. \quad (1)$$

During the contact, the mallet is subjected to the interaction force  $F(t)$  between mallet and membrane, due to the compression of the felt, and the displacement is governed by Newton's second law,

$$m \frac{d^2 u}{dt^2} = F(t). \quad (2)$$

Notice that, in the normal use of the instrument, the mallet is held at one end by the hand of the player. As a consequence, the effective dynamic mass  $m$  in Eq. (2) is not equal to the static mass, obtained, for example, by weighing the head on a precision balance (see Sec. IV).

$F(t)$  can be conveniently described by a nonlinear function of the felt compression. This model of compression can be viewed as a 2D extension of a previous model which has been successfully applied in the past for the modeling of hammer-string interaction in the piano,<sup>12</sup>

$$F(t) = K[(\delta - u(t) + W(t))^+]^\alpha, \quad (3)$$

where  $K$  is a stiffness coefficient and  $\alpha$  a phenomenological exponent. These coefficients are derived from a curve fitting of the force-deformation curve with a power law, in the usual range of the excitation force.<sup>7</sup> The symbol “+” means “positive part of.”  $W(t)$  is the mean displacement of the membrane's area in contact with the mallet defined as

$$W(t) = \int_{\Sigma} w(x, y, t) g(x, y) dx dy, \quad (4)$$

assuming that the impact is distributed over a small area of the membrane surface  $\Sigma$  at the excitation point of coordinates  $(x_0, y_0)$  with spatial window  $g(x, y)$ . The smooth positive function  $g(x, y)$  is normalized so that  $\int_{\Sigma} g(x, y) dx dy = 1 \text{ m}^2$ . The size of the spatial window has been estimated by measuring the size of the spots drawn on the membrane by mallets previously colored with ink. The variation of the contact area with time, during the impact, is neglected. The hysteretic cycle, due to the relaxation of the felt, and the elasticity of the stick, are also neglected in the model.

Assuming that the transverse displacement  $w(t)$  is sufficiently small,<sup>4</sup> then the membrane equation can be written

$$\sigma \frac{\partial^2 w}{\partial t^2} = \text{div} \left( T \nabla \left( w + \eta \frac{\partial w}{\partial t} \right) \right) - f(t) g - [p]_{\Sigma}, \quad (5)$$

which expresses that the inertial force is balanced by the sum of three terms: the restoring force due to the tension  $T$ , the impact force, and the pressure jump  $[p]$  on the membrane. The internal damping in the membrane is modeled by a relaxation (viscoelastic) term with coefficient  $\eta$ . The essential purpose of this coefficient is to introduce a frequency-dependent damping in the membrane. Its relevance will be discussed in Sec. IV. Notice that, for nonuniform membranes, the three coefficients  $\sigma$ ,  $T$ , and  $\eta$  are functions of space. The numerical simulation presented in Sec. II will allow such variations.

The force density term  $f(t)$  is related to the interaction force  $F(t)$  by the relation

$$F(t) = f(t) \int_{\Sigma} g(x, y) dx dy. \quad (6)$$

Due to the properties of  $g$ , one has  $F(t) = f(t)$ . As a consequence, there will be no distinction between these two variables in the following parts of the paper.

The membrane is assumed to be clamped at its periphery  $\partial\Sigma$ , which implies

$$w(x, y, t) = 0 \quad \forall (x, y) \in \partial\Sigma, \quad \forall t > 0. \quad (7)$$

The losses at the edge are neglected. At the origin of time, the membrane is assumed to be at rest, which yields

$$w(x, y, 0) = \frac{\partial w}{\partial t}(x, y, 0) = 0 \quad \forall (x, y) \in \partial\Sigma. \quad (8)$$

The acoustic fields both inside ( $\Omega_j$ ) and outside ( $\Omega_e$ ) the cavity are governed by the equations

$$\frac{\partial p_j}{\partial t} = -c_a^2 \rho_a \text{div} \mathbf{V}_j \quad \text{in } \Omega_j, \quad \text{for } j = e, i, \quad (9)$$

$$\rho_a \frac{\partial \mathbf{V}_j}{\partial t} = -\nabla p_j \quad \text{in } \Omega_j, \quad \text{for } j = e, i, \quad (10)$$

where  $c_a$  is the speed of sound,  $\rho_a$  is the density of air,  $p_j$  is the sound pressure in  $\Omega_j$ , and  $\mathbf{V}_j$  the acoustic velocity in  $\Omega_j$ . These equations are complemented by a condition of continuity for the normal component of the velocity on the surface  $\Sigma$  of the membrane in the plane  $z=0$  (see Fig. 1),

$$\mathbf{V}_j(x, y, 0, t) = \frac{\partial w}{\partial t}(x, y, t) \quad \forall (x, y) \in \Sigma, \quad \forall t > 0, \quad \text{for } j = e, i, \quad (11)$$

and by the initial conditions

$$p_j = 0, \quad \mathbf{V}_j = 0 \quad \text{in } \Omega_j, \quad \text{for } j = e, i. \quad (12)$$

Finally, it is assumed that the kettle is perfectly rigid, which means that

$$\mathbf{V}_j(x, y, z, t) \cdot \mathbf{n} = 0 \quad \forall (x, y, z) \in \mathcal{C}, \quad \forall t > 0, \quad \text{for } j = e, i, \quad (13)$$

where  $\mathbf{n}$  is the exterior unit vector normal to the surface  $\Gamma$  of the kettledrum. In the model, the shape of the kettle is assumed to be ellipsoidal. A comparison between a real bowl and this mathematically defined shape will be made in Sec. IV.

The system of coupled equations described above satisfies a property of energy decay. The energies of the mallet  $E^m$ , of the membrane  $E^w$ , and of the sound field (external field + cavity)  $E^a$ , respectively, are given by

$$E^m(t) = \frac{m}{2} \left| \frac{du}{dt} \right|^2 + \frac{K}{\alpha + 1} [(\delta - u(t) + W(t))^+]^{\alpha+1}, \quad (14)$$

$$E^w(t) = \frac{1}{2} \int_{\Sigma} \sigma \left| \frac{\partial w}{\partial t} \right|^2 ds + \frac{1}{2} \int_{\Sigma} T |\nabla w|^2 ds, \quad (15)$$

$$E^a(t) = \frac{1}{2} \int_{\mathbb{R}^3} \rho_a |\mathbf{V}|^2 d\Omega + \frac{1}{2} \int_{\mathbb{R}^3} \frac{1}{c_a^2 \rho_a} |p|^2 d\Omega, \quad (16)$$

where  $p$  and  $V$  are defined by their restrictions to  $\Omega_j$  ( $j = e, i$ ):  $p|_{\Omega_j} = p_j$  and  $\mathbf{V}|_{\Omega_j} = \mathbf{V}_j$ .

The first term in the right-hand side of Eqs. (14)–(16) is the kinetic energy, whereas the second term is the potential energy. Because of the viscoelastic damping term in the membrane equation, the total energy of the system  $E = E^m + E^w + E^a$  is decreasing with time. It can be shown that

$$\frac{dE}{dt}(t) = -\eta \int_{\Sigma} T \left| \nabla \frac{\partial w}{\partial t} \right|^2 ds \Rightarrow E(t) \leq \frac{m}{2} v_0^2. \quad (17)$$

One goal of the numerical method described in the following section will be to preserve the energy property expressed in Eq. (17) for the discrete system in space in order to ensure the stability of the numerical formulation.

## II. THE NUMERICAL METHOD

One main difficulty in the numerical resolution of the previous set of coupled equations results from the size of the three-dimensional problem. Another difficulty arises from the correct approximation of the bowl's shape. One possibility, in order to save time and memory, is to use a regular mesh composed of small cubes for the sound pressure and to approximate roughly the bowl by means of these cubes. This method, which is very easy to implement, leads unfortunately to numerical spurious diffractions on the artificial boundaries of the bowl due to the cubes and, thus, does not provide an acceptable solution. Another possibility is to use a tetrahedric mesh for the pressure, which fits the geometry of the bowl with a better approximation. However, this method is quite expensive in both computing time and memory. Finally, another alternative is to use the fictitious domain method.<sup>13,14</sup> It consists of rewriting the system of equations via the introduction of a new unknown  $\lambda$  which represents the pressure jump  $[p] = p_e - p_i$  across the boundary  $\Gamma$  of the instrument.

Using this method, the mallet's equations [Eqs. (2) and (3)] remain unchanged while the other four equations of the model can be rewritten in the following variational form (see Appendix A for more details):

$$\begin{aligned} \text{a) } & \frac{d^2}{dt^2} \int_{\Sigma} \sigma w w^* ds + \int_{\Sigma} T \nabla w \nabla w^* ds \\ & + \eta \frac{d}{dt} \int_{\Sigma} T \nabla w \nabla w^* ds + \int_{\Sigma} \lambda|_{\Sigma} w^* ds \\ & = -f(t) \int_{\Sigma} g w^* ds \quad \forall w^*, \\ \text{b) } & \frac{d}{dt} \int_{\mathbb{R}^3} \frac{1}{c_a^2 \rho_a} p p^* d\Omega + \int_{\mathbb{R}^3} \text{div } \mathbf{V} p^* d\Omega = 0 \quad \forall p^*, \\ \text{c) } & \frac{d}{dt} \int_{\mathbb{R}^3} \rho_a \mathbf{V} \cdot \mathbf{V}^* d\Omega - \int_{\mathbb{R}^3} p \text{div } \mathbf{V}^* d\Omega - \int_{\Gamma} \lambda \mathbf{V}^* \cdot \mathbf{n} ds \\ & = 0 \quad \forall \mathbf{V}^*, \\ \text{d) } & \frac{d}{dt} \int_{\Sigma} w \lambda|_{\Sigma}^* ds - \int_{\Gamma} \mathbf{V} \cdot \mathbf{n} \lambda^* ds = 0 \quad \forall \lambda^*, \end{aligned} \quad (18)$$

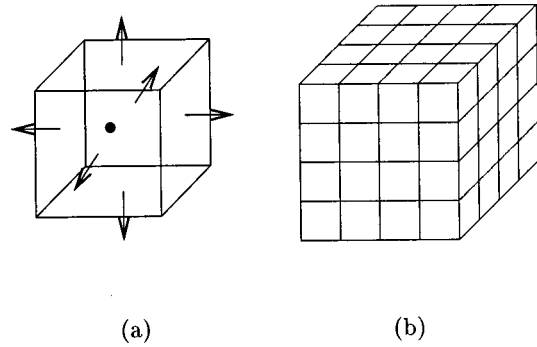


FIG. 2. Mesh for the pair of variables  $(p_h, \mathbf{V}_h)$ . (a) "Degrees of freedom" on a cube:  $\bullet$  for  $p_h$ ,  $\rightarrow$  for the normal components of  $\mathbf{V}_h$ . (b) Mesh for  $\Omega$ .

where the unknowns  $(w, p, \mathbf{V}, \lambda)$  and the test functions  $(w^*, p^*, \mathbf{V}^*, \lambda^*)$  are chosen in the appropriate spaces of functions so that all integrals in Eq. (18) are defined.

Notice that there are only integrals on  $\Sigma$ ,  $\mathbb{R}^3$ , and  $\Gamma$  but not on  $\Omega_e$  and  $\Omega_i$  in Eq. (18). Due to this property, one can ignore the interior and exterior volume when meshing the space for the pressure. One must work with a space of continuous functions for  $w$ , because of the term  $\nabla w$ , and with a space of vector fields whose normal components across the surface are continuous for  $\mathbf{V}$ , because of the term  $\text{div } \mathbf{V}$ . The space for  $p$  is not required to satisfy any condition of continuity.

To simulate the free space, the actual computations are restricted to a box  $\Omega$  of finite size which surrounds the instrument, and higher-order absorbing conditions (of the 12th order) are implemented on the boundary of this domain (see Ref. 11). For the sake of simplicity, the additional equations resulting from these absorbing conditions will be omitted in this section. These equations are presented in Appendix B.

The continuous variables  $(p, \mathbf{V})$  are approximated by discrete variables in space  $(p_h, \mathbf{V}_h)$  on a regular mesh of  $\Omega$  composed of small cubes [see Fig. 2(b)]. The index  $h$  indicates that the variables are discrete in space,  $h$  being related to the selected spatial step.  $p_h$  is constant for each element and  $\mathbf{V}_h$  belongs to a space of mixed finite elements,<sup>15</sup> so that each normal component of  $\mathbf{V}_h$  is linear in one direction and constant in the two other directions, on each cube.  $p_h$  is fully determined by its value in the center of each cube, and  $\mathbf{V}_h$  is given by the values of its normal components across each face of the cube. Figure 2(a) illustrates the degrees of freedom per cube for  $p_h$  and for the six normal components of  $\mathbf{V}_h$ . A major interest of the fictitious domain method is that the cubic mesh used for computing the pressure is uniform since it ignores the shape of the instrument (see Appendix A).

The variables  $w$  and  $\lambda$  are discretized by means of  $P_1$ -continuous finite elements on a triangular mesh, on  $\Sigma$  and  $\Gamma$ , respectively. The approximations  $w_h$  and  $\lambda_h$  are entirely determined by their values at each node of the mesh. Figure 3 shows the three "degrees of freedom" for  $\lambda_h$  on a triangle and the mesh for  $\Gamma$ . Figure 4 shows the mesh for  $w_h$ . In order to facilitate the computations, the mesh for  $\Sigma$  is a refinement on the membrane of the mesh for  $\Gamma$ , so that each triangle of the  $\Gamma$ -mesh contains 16 triangles for the mesh of  $\Sigma$  [see Fig. 4(a)]. This refinement is consistent with the ratio

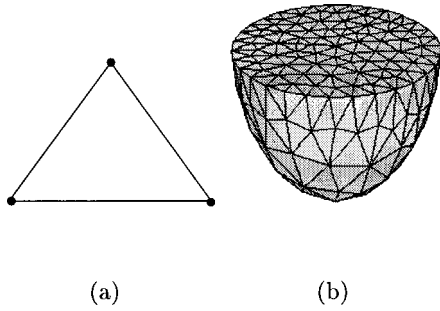


FIG. 3. Mesh for  $\lambda_h$ . (a) “Degrees of freedom” on a triangle. (b) Mesh for  $\Gamma$ .

between the wave velocity on the membrane and the speed of sound in air.

The discretized system in space is obtained by substituting  $(w_h, p_h, \mathbf{V}_h, \lambda_h)$  for  $(w, p, \mathbf{V}, \lambda)$  and  $(w_h^*, p_h^*, \mathbf{V}_h^*, \lambda_h^*)$  for  $(w^*, p^*, \mathbf{V}^*, \lambda^*)$  in the continuous equations [Eq. (18)]. The integrals were computed using quadrature formula so that the mass matrices  $(M_w, M_p, M_V)$  [see Eq. (19)] are diagonal. The resulting differential system, including the mallet’s equations, can be written in the following matrix form:

$$\begin{aligned}
 \text{a) } & f_h = K[(\delta - u_h + G^t w_h)^+]^\alpha, \\
 \text{b) } & m \frac{d^2 u_h}{dt^2} = f_h, \\
 \text{c) } & M_w \frac{d^2 w_h}{dt^2} + R_w w_h^n + \eta R_w \frac{dw_h}{dt} + A_w \lambda_h = -G f_h, \\
 \text{d) } & M_p \frac{dp_h}{dt} + D_V^t \mathbf{V}_h = 0, \\
 \text{e) } & M_V \frac{d\mathbf{V}_h}{dt} - D_V p_h - B_V \lambda_h = 0, \\
 \text{f) } & B_V^t \mathbf{V}_h - A_w^t \frac{dw_h}{dt} = 0,
 \end{aligned} \tag{19}$$

where  $M_w$ ,  $R_w$ ,  $M_p$ ,  $M_V$ ,  $D_V$ ,  $B_V^t$ , and  $A_w^t$  are matrices,  $G$  is a vector representing the spatial window of the mallet, and the superscript  $t$  on matrices stands for transpose matrices.

One property of the fictitious domain method is that the energy property expressed in Eq. (17) remains valid for the space discretized system and is not perturbed by the introduction of the new unknown  $\lambda$ . The discrete energies of the

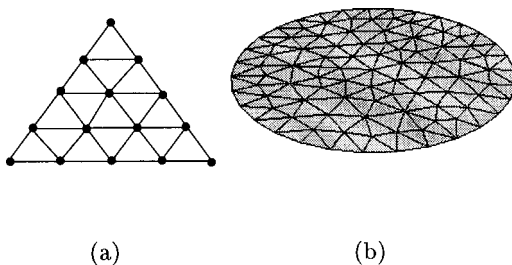
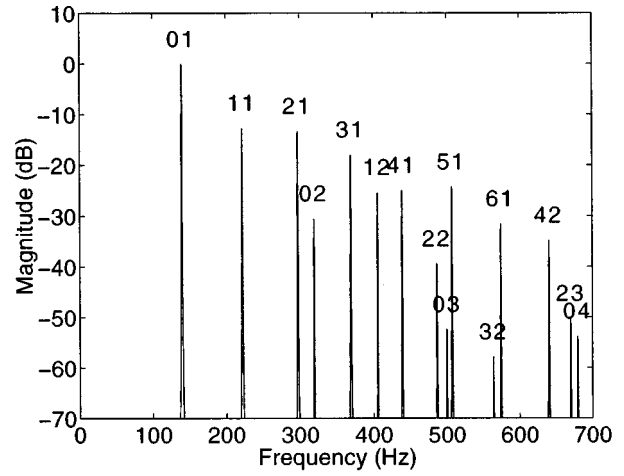
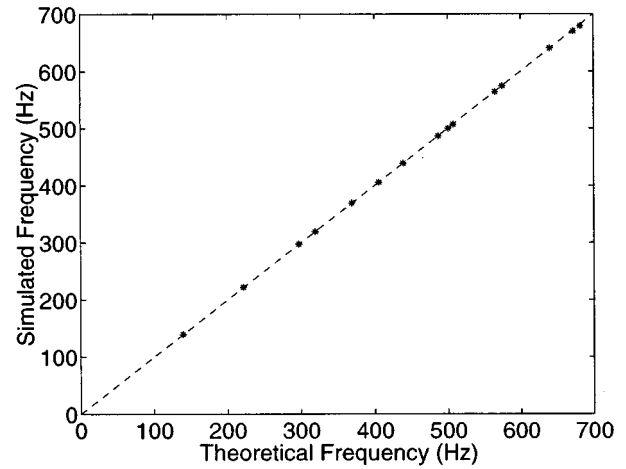


FIG. 4. Mesh for  $w_h$ . (a) Refinement of 1 triangle into 16 triangles, the dots indicate the “degrees of freedom.” (b) Mesh for  $\Sigma$ .



(a)



(b)

FIG. 5. Membrane *in vacuo*: Comparison between theoretical and simulated eigenfrequencies. (a) Spectrum of the membrane displacement. The corresponding modes are indicated above each spectral line. (b) Simulated versus theoretical frequencies.

mallet, of the membrane, and of the acoustic field, which correspond, respectively, to the continuous expressions (14), (15), and (16), are defined as follows:

$$E_h^m(t) = \frac{m}{2} \left| \frac{du_h}{dt} \right|^2 + \frac{K}{\alpha + 1} [(\delta - u_h + G^t w_h)^+]^{\alpha+1}, \tag{20}$$

$$E_h^w(t) = \frac{1}{2} \left( M_w \frac{dw_h}{dt} \right)^t \frac{dw_h}{dt} + \frac{1}{2} (R_w w_h)^t w_h, \tag{21}$$

$$E_h^a(t) = \frac{1}{2} (M_V \mathbf{V}_h)^t \mathbf{V}_h + \frac{1}{2} (M_p p_h)^t p_h. \tag{22}$$

It can be shown that the total discrete energy of the system, defined as the sum  $E_h = E_h^m + E_h^w + E_h^a$ , is decreasing with time. As a consequence, the energy at time  $t$  is lower than the initial discrete energy,

TABLE I. Typical values of the parameters used in the simulation.

Mallet		
$v_o = 1.4 \text{ m s}^{-1}$	$\delta = 0.025 \text{ m}$	$m = 0.028 \text{ Kg}$
$K = 1.6 \times 10^8 \text{ N m}^{-\alpha}$	$\alpha = 2.54$	
Interaction force location		
$(x_0, y_0) = (0.21, 0)$	$g(x, y) = \frac{\exp[-10^7((x-x_0)^4 + (y-y_0)^4)]}{\int_{\Sigma} \exp[-10^7((x-x_0)^4 + (y-y_0)^4)]}$	
Membrane		
$a = 0.31 \text{ m}$	$\sigma = 0.262 \text{ kg m}^{-2}$	
$T = 3325 \text{ N m}^{-1}$	$\eta = 0.6 \times 10^{-6} \text{ s}$	
Air constants		
$c_a = 344 \text{ m s}^{-1}$	$\rho_a = 1.21 \text{ kg m}^{-3}$	
Kettle		
$H = 0.5 \text{ m}$	$V = 0.1 \text{ m}^3$	
Meshes		
membrane:	nodes = 2963; triangles = 5744	
pressure jump:	nodes = 361; triangles = 718	
pressure:	$\Delta x = \Delta y = \Delta z = 0.025 \text{ m}$	
	observation point: $(x_M, y_M, z_M) = (-0.31, 0, +0.10) \text{ m}$	
	number of nodes: $N_x = N_y = 40, N_z = 50$	
Sampling		
Time sampling frequency $f_e = 24 \text{ kHz}$	number of iterations = 72 000	

$$\frac{dE_h}{dt}(t) = -\eta \left( R_w \frac{dw_h}{dt} \right)^t \frac{dw_h}{dt} \Rightarrow E_h(t) \leq \frac{m}{2} v_0^2. \quad (23)$$

For the time discretization of the equations, a classical centered finite difference scheme is used. In order to obtain a centered system, the variables  $p_h$ ,  $w_h$ , and  $\lambda_h$  are computed at instants  $t^n$ , and the variable  $V_h$  is computed at instants  $t^{n+1/2}$ , where  $t_l = l\Delta t$ , and  $\Delta t$  is the time step.

The resulting discrete system is of the following form:

$$\begin{aligned} \text{a) } f_h^n &= K[(\delta - u_h^n + G^t w_h^n)^+ ]^\alpha, \\ \text{b) } m \frac{u_h^{n+1} - 2u_h^n + u_h^{n-1}}{\Delta t^2} &= f_h^n, \\ \text{c) } M_w \frac{w_h^{n+1} - 2w_h^n + w_h^{n-1}}{\Delta t^2} + R_w w_h^n + \eta R_w \frac{w_h^{n+1} - w_h^{n-1}}{2\Delta t} \\ &+ A_w \lambda_h^n = -G f_h^n, \\ \text{d) } M_p \frac{p_h^{n+1} - p_h^n}{\Delta t} + D_V^t \mathbf{V}_h^{n+1/2} &= 0, \\ \text{e) } M_V \frac{\mathbf{V}_h^{n+1/2} - \mathbf{V}_h^{n-1/2}}{\Delta t} - D_V p_h^n - B_V \lambda_h^n &= 0, \\ \text{f) } B_V^t \mathbf{V}_h^{n+1/2} - A_w^t \frac{w_h^{n+1} - w_h^n}{\Delta t} &= 0. \end{aligned} \quad (24)$$

This system is solved by considering Eq. (24) as a linear system in  $(u_h^{n+1}, w_h^{n+1}, p_h^{n+1}, \mathbf{V}_h^{n+1/2}, \lambda_h^n)$ , all other terms being known. In practice, most of the computations are explicit, for  $p_h$  and  $V_h$  in particular, and one simply has to invert a linear system for  $\lambda_h$ . Other details of the method

used for this resolution are beyond the scope of this paper and will not be presented here.

### III. RESULTS OF SIMULATION

The model is first validated by a comparison between theoretical and simulated eigenfrequencies for the ideal case of a circular membrane *in vacuo* clamped at its periphery. In another series of numerical experiments, the complete kettle-drum model presented in Sec. I is used in order to investigate the vibroacoustics of the instrument in both the time and frequency domain. The values of the parameters used in the numerical computations are listed in Table I.

TABLE II. Membrane *in vacuo*: Comparison between theoretical and simulated eigenfrequencies (uniform tension of 3325 N/m).

Modes (mn)	$f_{mn}$ (theoretical) (Hz)	$f_{mn}$ (simulated) (Hz)	Relative error %
01	139.1	139.3	0.18
11	221.6	222.0	0.17
21	297.0	297.3	0.10
02	319.3	319.3	0.02
31	369.0	369.3	0.09
12	405.8	405.7	0.02
41	438.9	439.0	0.03
22	486.8	486.7	0.03
03	500.5	500.0	0.10
51	507.3	507.0	0.06
32	564.6	564.0	0.10
61	574.7	574.3	0.06
42	639.9	640.3	0.06
23	672.1	670.3	0.26
04	682.0	679.7	0.34

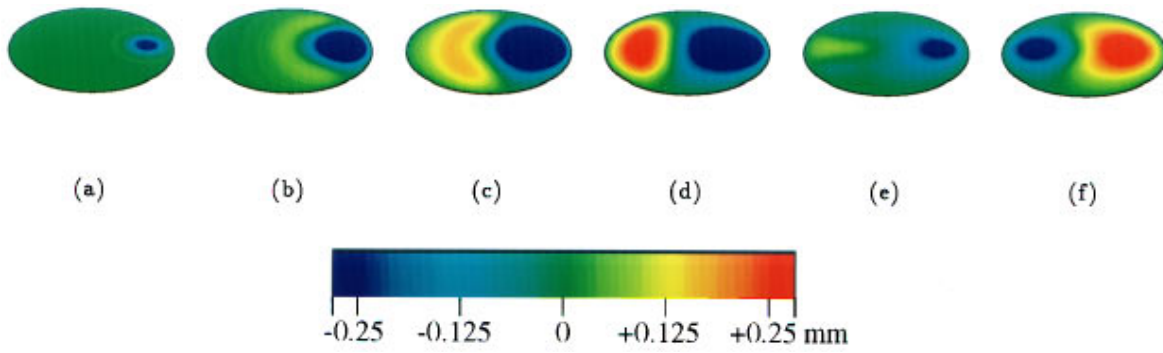


FIG. 6. Snapshots of the membrane displacement during the first 10 ms.

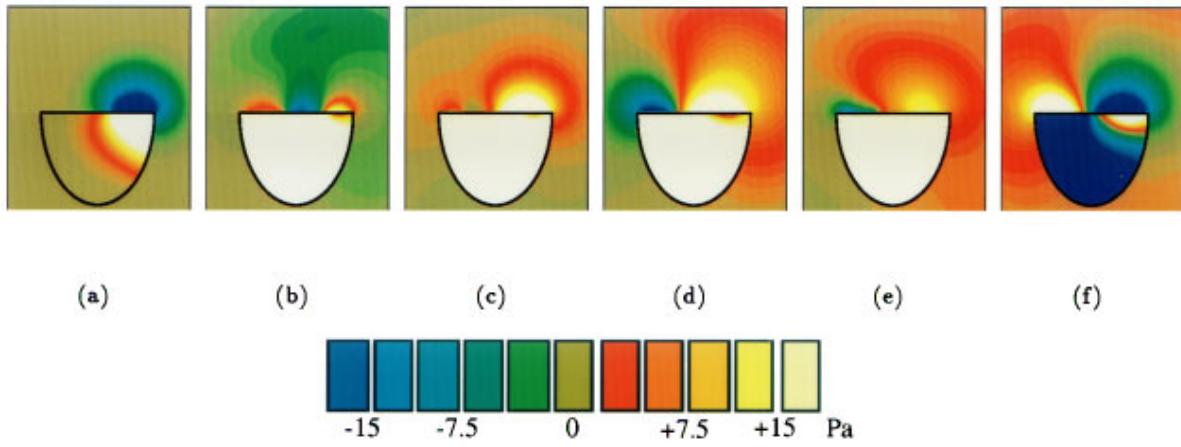


FIG. 7. Snapshots of the acoustic pressure during the first 10 ms.

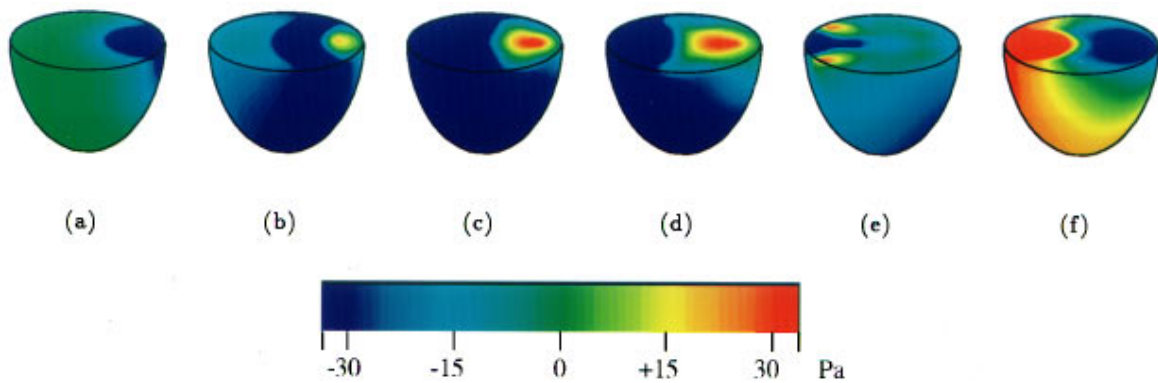


FIG. 8. Snapshots of the pressure jump during the first 10 ms.

## A. Membrane *in vacuo*

For an ideal circular membrane with uniform tension, clamped at its periphery, the normalized mode functions and angular frequencies ( $\xi_{mn}, \omega_{mn}$ ) for  $m=0,1,2,\dots$  and  $n=1,2,3,\dots$  satisfying,

$$-c^2 \Delta_{\Sigma} \xi_{mn} = \omega_{mn}^2 \xi_{mn}, \quad (25)$$

are given by<sup>5</sup>

$$\xi_{mn}(\rho, \phi) = \frac{2}{a J_{m+1}(x_{mn})} J_m \left( x_{mn} \frac{\rho}{a} \right) \frac{e^{im\phi}}{(2\pi)^{1/2}}, \quad (26)$$

$$\omega_{mn} = x_{mn}(c/a) = 2\pi f_{mn},$$

where  $c = \sqrt{T/\sigma}$  is the transverse wave velocity on the membrane,  $J_m$  are the Bessel functions of the first kind of order  $m$ , and  $x_{mn}$  is the  $n$ th zero of  $J_m$ . Each mode is designated by two integers, where  $m$  corresponds to the number of nodal diameters and  $n$  to the number of nodal circles (including the fixed boundary).

Numerical simulations are made by considering only the ideal case where the nonlinear mallet strikes a lossless membrane *in vacuo*. For this simple case, both external and cavity pressure are set to zero in the model. The spectrum of the membrane displacement waveform obtained from the simulation program is analyzed with standard fast Fourier transform (FFT) tools. The sample rate here is 24 kHz and the length is 3 s (72 000 points). A Hanning window has been used for the analysis. The results shown in Table II show a remarkable agreement between the first 16 theoretical eigenfrequencies obtained from Eq. (26) and the values obtained from the Fourier analysis of the simulated membrane motion. Below 700 Hz, the relative error is less than 0.5% between these two sets of frequencies (see Fig. 5).

## B. Complete kettledrum model

The model now couples the membrane with the mallet and with both the cavity and external sound pressure. The complete set of equations presented in Sec. I is used in the simulations.

Figure 6 shows the transverse displacement of the membrane at successive instants of time. The green color corresponds to the equilibrium state of the membrane at its initial position [see Fig. 6(a)]. When the mallet strikes the membrane, it pushes it down. This corresponds to the blue color all around the point of impact [see Fig. 6(b)]. The elastic membrane reacts to this sudden downward force and rises, shown in red, ahead of the impact [see Fig. 6(c) and (d)]. The transverse wave is reflected at the border of the membrane with a change of sign [see Fig. 6(d) and (e)]. Finally, in its phase of decompression, the mallet leaves the membrane which pushes it away [see Fig. 6(f)].

Figure 7 shows the pressure field in the plane of symmetry of the kettledrum perpendicular to the membrane, which contains the impact point, at the same instants of time as in Fig. 6. At the beginning of the impact, the volume of the cavity decreases which, in turn, leads to a positive acoustic pressure inside the cavity [in white in Fig. 7(a)]. In con-

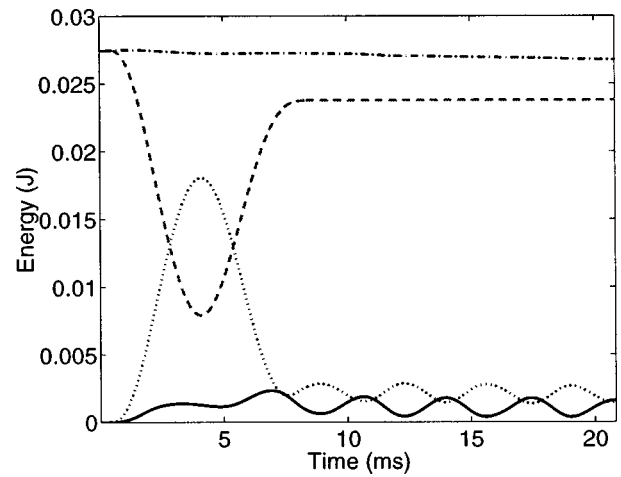


FIG. 9. Exchange of energy between the mallet (dashed line), the membrane (dotted line), and the acoustic field (solid line) during the first 20 ms of the sound. The total energy of the system is represented in dash-dotted line.

trast, outside the cavity, a negative pressure can be observed [in blue in Fig. 7(a)]. The inside wave is naturally restrained by the shape of the bowl [see Fig. 7(b)].

Both internal and external acoustic wavefronts are propagating at sound speed  $c_a$ , i.e., approximately three times faster than the elastic wave on the membrane: comparing Figs. 6(b) and 7(b), for example, shows that, at that time, the internal sound wave has already reached the opposite side of the cavity, compared to the impact point, while the membrane is still at rest in this area.

In the vicinity of the membrane itself, the sound field shows an interesting feature: it seems that an apparent “guided” wave, which immediately precedes the elastic wavefront on the membrane, is propagating [in green in Fig. 7(b) to (d)]. However, this feature, for which we have yet no convincing explanation, is more clearly seen on animated pictures than on fixed illustrations (see Ref. 16).

To see the effects of the discontinuity in the sound field on both sides of the kettledrum surface  $\Gamma$  more clearly, it is interesting to observe the time history of the pressure jump  $[p]_{\Gamma} = p_e - p_i$ , where  $p_i$  is the inside pressure and  $p_e$  the pressure outside the instrument on  $\Gamma$  (see Fig. 8). The blue color here corresponds to a negative pressure jump, which means that the magnitude of the pressure on the internal cavity surface is higher than on the surface in the external field. The red color indicates a positive pressure jump. A comparison between Figs. 6 and 8 shows that the pressure jump in the membrane plane is very similar to the elastic displacement field.

Finally, the time history of the energy terms defined in Sec. II were calculated in order to understand how the total energy of the system is balanced between mallet, membrane, and air. At the origin of time, the unique energy of the system is the kinetic energy of the mallet. During the compression, this energy is partially transformed into potential energy, because of the mallet’s stiffness, and partially transferred to the kettledrum. Conversely, during its phase of decompression, the kinetic energy of the mallet increases whereas its potential energy decreases. At the end of the interaction time, it can be seen in Fig. 9 that the mallet leaves

the membrane with an energy lower than its initial value, as expected. During the impact, the membrane energy  $E^w$  is significantly larger than the acoustic energy. After the mallet has left the membrane, one can see in Fig. 9 an almost periodical exchange of energy between membrane and air. Due to both the internal losses, through the viscoelastic term  $\eta$  in Eq. (5), and radiation in free space, simulated here by means of absorbing conditions at the boundaries, the total energy  $E$  of the system decreases slowly with time, as shown in Fig. 9.

#### IV. COMPARISON WITH EXPERIMENTS

We now examine to what extent the present model accounts for the sounds produced by real timpani.

In order to compare the results of the simulations with measured sound pressures, it is necessary to estimate first the parameters of the model from experiments on a real instrument. The experiments were conducted on a hand-tuned 25" (63.5 cm) Premier kettledrum with fiberglass bowl and clear Mylar head (Premier 8435). The tuning was adjusted so that  $f_{11} = 147$  Hz (note D<sub>3</sub>). The instrument is set into vibration by means of a hard mallet Vibrawell FC7 with carbon fiber handle whose head has a diameter of 46 mm. The membrane is struck at 10 cm from the edge. The pressure is recorded in the near field at a distance of 10 cm above the edge by means of an omni-directional Schoeps MC2 microphone, following the recommendations by Sullivan<sup>10</sup> (see Fig. 1). All measurements were made in a studio recording room with very low reverberation ( $\tau_{60} \leq 0.4$  s at 125 Hz and  $\leq 0.15$  s at 500 Hz) so that measurements of decay times are not altered by the room.

The spectral plots presented in the section are limited to the range 0–700 Hz. This upper limit has been selected by considering that this range contains most of the energy of the sound: the magnitudes of the spectral lines above 700 Hz are approximately 60 dB lower than the main peaks. There are no theoretical limitations here for enlarging this frequency range, but the price to pay would naturally be a refinement of the mesh which would, in turn, lead to an increase in computing time.

##### A. Measurements on real instruments

The effective mass  $m$  of the mallet and its initial velocity  $v_0$  are determined by using the same procedure as the one used by Chaigne and Doutaut for xylophones.<sup>17</sup> In order to measure  $m$ , an impedance head B&K 8001 is struck by the mallet. This impedance head yields force and acceleration signals which are very similar in shape. Thus,  $m$  is obtained from the ratio between these two signals. These experiments show that the effective mass of the mallet to be considered here is about 20% higher than the static mass of the head. The initial velocity  $v_0$  is obtained through integration of the acceleration signal of the mallet, during its contact with the membrane. The stiffness coefficient  $K$  and the nonlinear exponent  $\alpha$  of the mallet were obtained from force-deformation curves, following the same method as for piano hammers.<sup>12,18</sup>

The density of the membrane is simply obtained by weighting small samples of a torn membrane. The estimation

TABLE III. Complete kettledrum: Comparison between measured and simulated most salient eigenfrequencies (uniform tension of 3325 N/m and viscoelastic constant  $\eta = 0.6 \times 10^{-6}$  s).

Mode (mn)	$f_{mn}$ (measured) (Hz)	$f_{mn}$ (simulated) (Hz)
01	136	139
11	147	147
21	221	221
02	248	245
31	288	288
12	315	315
41	357	355
22	395	395
03	403	408
51	424	419
61	480	483
71	552	544

of the tension, however, is more difficult and cannot be obtained directly. In order to measure this parameter, we use the property that the upper eigenmodes of the membrane are only slightly perturbed by the air loading. From the series of eigenfrequencies above 1 kHz, an estimate for the transverse wave velocity on the membrane is then obtained, from which the tension is derived. This method is fast and attractive since it yields a good estimate for the tension directly from the pressure waveform recorded by a microphone. However, when compared to the simulations, the value of the tension obtained with this method leads to a systematic underestimation of the eigenfrequencies (about 5%). Therefore, starting from the estimated value, the tension is increased step-by-step in the simulation program until both measured and simulated  $f_{11}$ -frequency coincide in the pressure spectrum. In this case, a remarkable agreement is observed between the measured and simulated series of eigenfrequencies up to 700 Hz, which proves the validity of the method. The results shown in Table III are obtained through FFT analysis on a sample length of 9000 points (0.375 s at a sampling rate of 24 kHz) with a Hanning window, in the initial portion of the sound. This option follows from the necessity to use a relatively short window in order to measure the  $0n$ -modes with sufficient accuracy, since these modes are damped very rapidly. The uniform tension here is equal to 3325 N/m and the viscoelastic constant  $\eta$  is  $0.6 \times 10^{-6}$  s.

In Eq. (5), a viscoelastic constant  $\eta$  was introduced in order to account for the fact that the internal damping in the Mylar membrane usually increases with frequency, as it is also observed in nylon strings and xylophone bars. In these two latter cases,  $\eta$  was simply derived from measurements of the partials' decay times.<sup>17,19</sup> However, the problem is more complicated here since the decay times of the partials are not only governed by the internal damping but also by the radiation through the pressure jump term in Eq. (5). Therefore, the procedure used here was the following: from the series of measured decay times  $\tau_i$  for the lowest 20 to 30 eigenfrequencies  $f_i$  on a real kettledrum, an estimate for the viscoelastic constant  $\eta$  was derived, assuming that the radiation losses are negligible. In these cases,  $\eta$  is obtained from a curve-fitting between  $\tau_i$  and  $f_i$ , recalling that, from Eq. (5), one can easily show that the  $\tau_i$  are given by



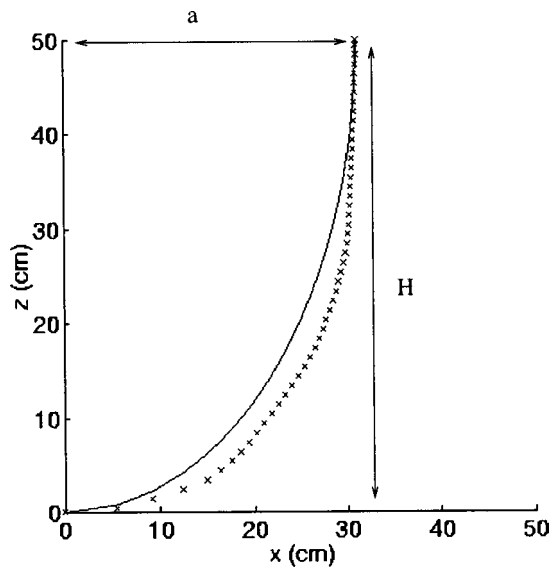


FIG. 10. Comparison between real and simulated kettle shape. Ellipsoidal shape (solid line); measured shape ("x").

$$\tau_i = \frac{1}{2\eta\pi^2 f_i^2} \quad (27)$$

The value obtained for  $\eta$  by means of Eq. (27) is used as a starting value for the simulations. Then, as for the tension, this parameter is adjusted by trial-and-error until the comparison between measured and simulated decay times is satisfactory enough. It must be emphasized that the model is quite sensitive to this parameter: differences of  $\pm 20\%$  in  $\eta$  are clearly audible.

In order to facilitate the calculation of the finite element's mesh, the shape of the cavity has been approximated by a half-ellipsoid given by the following equation:

$$z = -\frac{H}{a}\sqrt{a^2 - x^2 - y^2}, \quad (28)$$

$a$  being the radius of the membrane, and  $H$  the height of the cavity. The volume  $V$  of the kettle is then equal to  $2/3 \pi a^2 H \approx 0.1 \text{ m}^3$ . Figure 10 shows the comparison between this ellipsoidal shape and the shape of the real kettle-drum. It can be seen that this approximation leads to a slight underestimation of the cavity volume (nearly 10%).

### B. Comparisons between simulations and experiments

Figure 11 shows a comparison between measured and simulated accelerations of the mallet. The two curves are very similar in amplitude and shape, and the impact duration is fairly reproduced. The discrepancies ( $\pm 10\%$ ) are probably due to the relative simplicity of the model, where the mallet is reduced to its center of gravity, as well as to some errors in the estimation of the initial mallet's velocity and constants  $K$  and  $\alpha$ .

Similarly, Fig. 12 shows a comparison between measured and simulated pressure at the position of the microphone. In Fig. 12(a), the waveforms are compared during the first 100 ms of the sound, i.e., during the initial part which

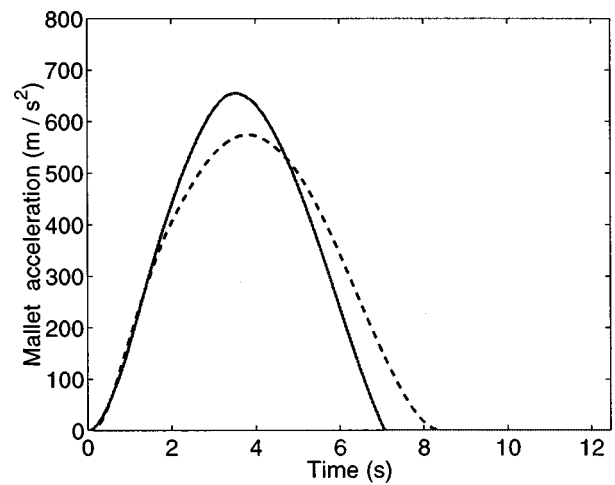
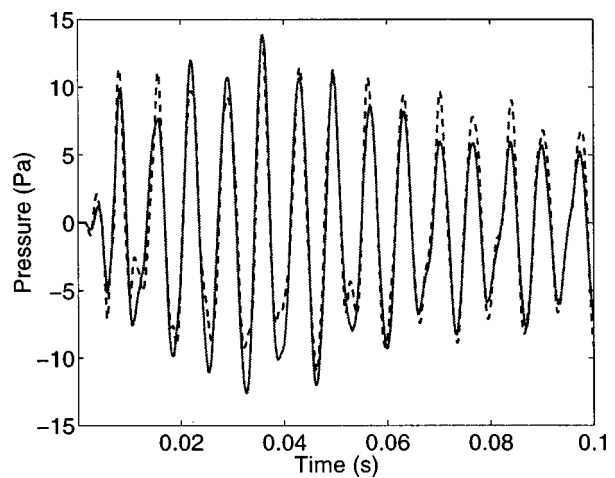
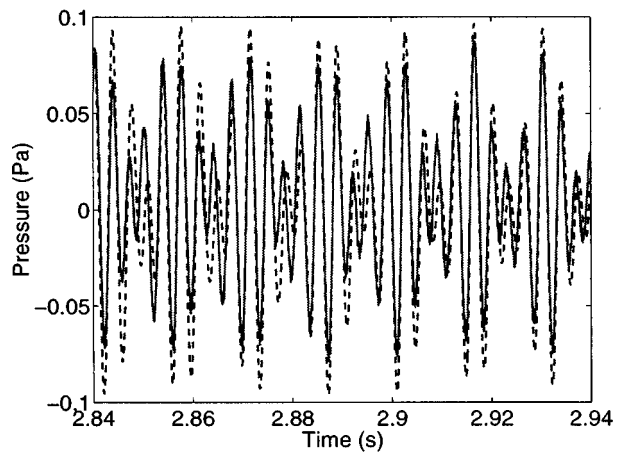


FIG. 11. Comparison between measured (solid line) and simulated (dashed line) acceleration of the mallet.

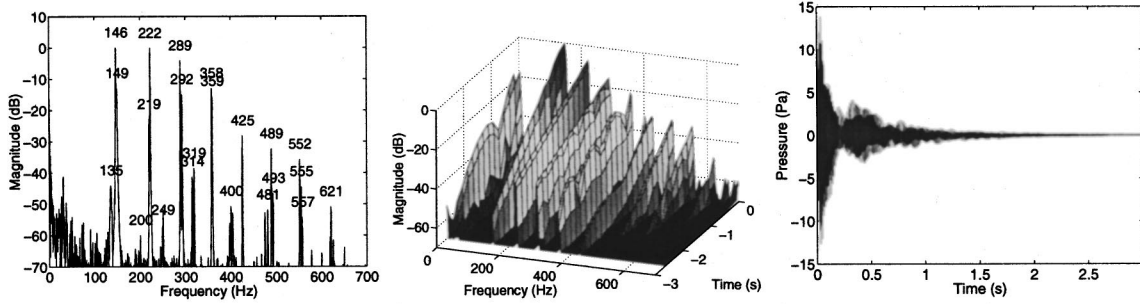


(a)

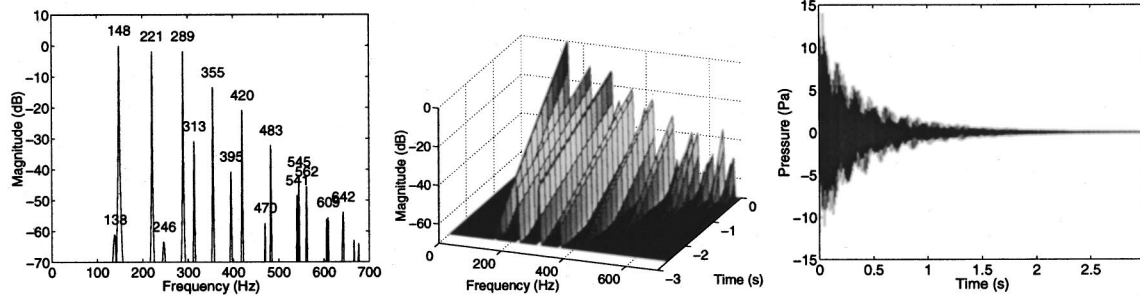


(b)

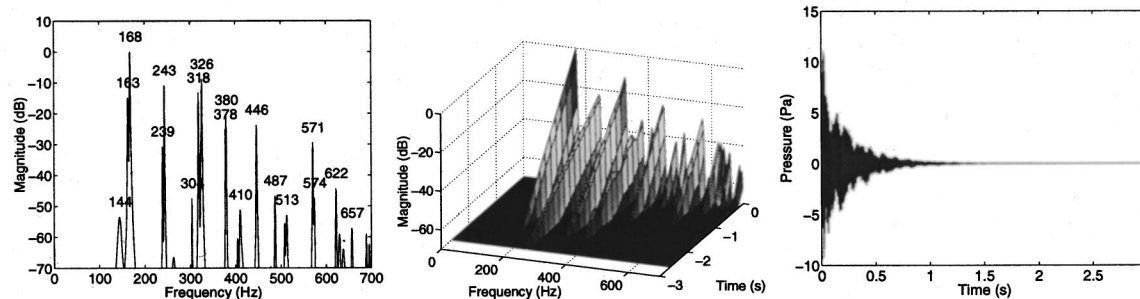
FIG. 12. Comparison between measured (solid line) and simulated (dashed line) sound pressure. (a) First 100 ms of the sound. (b) Free vibrations (2.84 s after the impact).



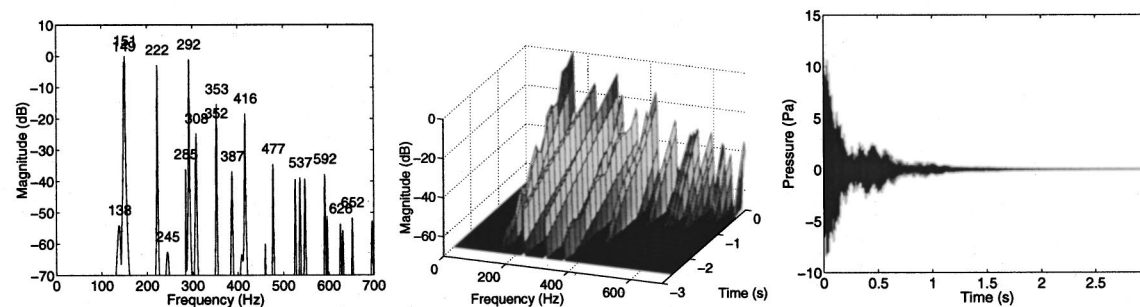
(a)



(b)



(c)



(d)

FIG. 13. Power spectra, 3D plots and waveforms of timpani sounds. (a) Real kettledrum. (b) Simulated kettledrum with uniform tension of 3325 N/m. (c) Simulated kettledrum with nonuniform tension (case 1). (d) Simulated kettledrum with nonuniform tension (case 2).

immediately follows the impact of the mallet. Figure 12(b) shows the waveforms nearly 3 s later, while the membrane vibrates freely. In both cases, it can be seen that the two sets of curves are almost superimposed.

The good similarity of the two signals is confirmed by spectral analyses [see Fig. 13(a) and (b)]. It is a well-established fact that timpani sounds have the property of containing partials which are almost multiple integers of a fundamental which is not present in the signal. These partials are musically very important, since they convey the pitch of the played note. They correspond to the  $(m,1)$  angular frequencies of the membrane.<sup>3</sup> In our simulations, these frequencies are very well reproduced, as seen in the first column of Fig. 13, which shows spectra averaged over the total duration of the sound (3 s).

The 3D plots in the second column of Fig. 13 are obtained through short-time Fourier transform on successive sample lengths of 9000 points (0.375 s) with Hanning window and 50% overlap.

One can notice some differences between measured and simulated pressure. Below 100 Hz, the spectrum of the measured signal in Fig. 13(a) shows some low-frequency noise, with a few salient spectral lines, which is not present in the simulations. It has been checked that this noise is due to the electronic channel and not to the sound field.

Above 600 Hz, the simulated signals seem to contain more high frequencies than the real one. In addition, in this spectral region, the magnitude of the peaks in the simulations are slightly higher ( $\approx 5$  dB) than the corresponding peaks in the real spectrum. This may be a consequence of the approximations made in the modeling of the damping causes, since the losses at the membrane's edge and in the mallet's felt are neglected.

Another obvious difference between measurements and simulations is the presence of double peaks in the spectrum of the real pressure. As a consequence of this peak doubling, a number of beats can be observed in the time-frequency spectrograms (see the central column in Fig. 13). These beats are also visible in the waveforms and are clearly heard.

The simulations were used in order to check whether or not this feature is due to imperfect tuning of the membrane, as it is usually assumed. For that purpose, simulations were made where the tension of the membrane is not uniform, but slightly higher around the tuning-screws than in the center. Figure 14 illustrates two examples of distribution for the membrane's tension which were used in the computations.

In the first case [see Fig. 14(a)], the uniform tension (3325 N/m) is increased arbitrarily by about 5% around four screws and by about 10% and 15% around the two remaining screws, respectively. As a result of this nonuniform tensioning, it can be seen in Fig. 13(c) that the eigenfrequencies are significantly increased (about 10%) and that multiple double peaks now appear in the spectrum. Consequently, multiple beats are visible in the 3D spectrogram. However, the values of the simulated eigenfrequencies do not correspond any longer to the one of the real kettledrum, even though the general shape of the spectrogram looks more realistic.

An appropriate method for obtaining simulations more

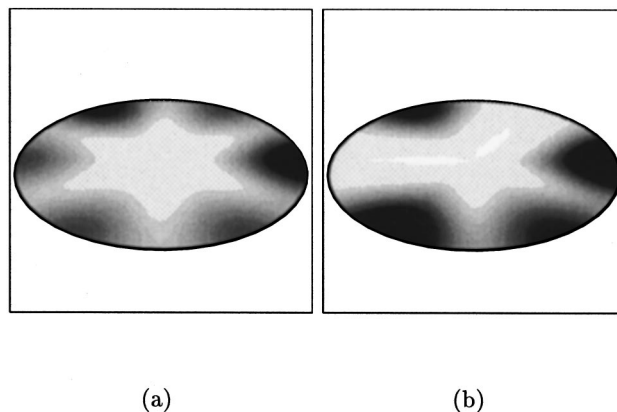


FIG. 14. Distribution of variable tension on the membrane. (a) Case 1: The tension varies from 3325 (white) to 3770 N/m (black). (b) Case 2: The tension varies from 3100 (white) to 3317 N/m (black).

similar to the measured pressure would have been to measure accurately the distribution of tension over the real membrane. These experiments were not conducted, due to their extreme difficulty. As an alternative, a trial-and-error procedure was used where a number of different simulations were made by changing the distribution of tension around the screws. Figure 14(b) shows an example of distribution (case 2) which yields a very satisfactory agreement between measured and simulated waveforms and spectra: it corresponds to a tension of 3100 N/m in the center of the membrane (in white in the figure) and to a maximum increase of 7% around four of the six tuning-screws (in black in the figure). The corresponding waveform and spectrum can be seen in Fig. 13(d). The simulated eigenfrequencies now agree with the measured ones within  $\approx 3\%$ . Another meaningful result is that the double peaks, the time evolution of the partials, and the envelope of the waveform are now well reproduced. The corresponding synthesized sounds are very close to the real one. In conclusion, this experiment shows, at least qualitatively, that these almost unavoidable features of timpani sounds can be reasonably attributed to irregularities in the tensioning process.

The tone quality of timpani is highly dependent on the decay times of the main partials, as it has been already mentioned by other authors.<sup>10</sup> This requires looking more closely at the comparison between measured and simulated time histories of these frequencies. For the sake of clarity in the figures, the comparison is made in what follows between the simulated spectrum obtained with uniform tension [see Fig. 13(b)] and the measured one. As a consequence, the simulated partials are decaying smoothly without beats, which makes them more recognizable. The decaying curves presented in Fig. 15 are obtained by selecting the most salient frequencies in the initial part of the 3D spectral plots. This figure shows that the decay times of the  $(m,1)$  modes are well reproduced, whereas the decay times of the  $(0,n)$  modes are slightly more damped in the simulations than in the reality. The corresponding frequencies are listed in Table III. One can notice differences of a few Hz between these frequencies and those presented on the spectra in the first column of Fig. 13(a) and (b). These small deviations are due to differences in the signal duration selected for the spectral analysis. In

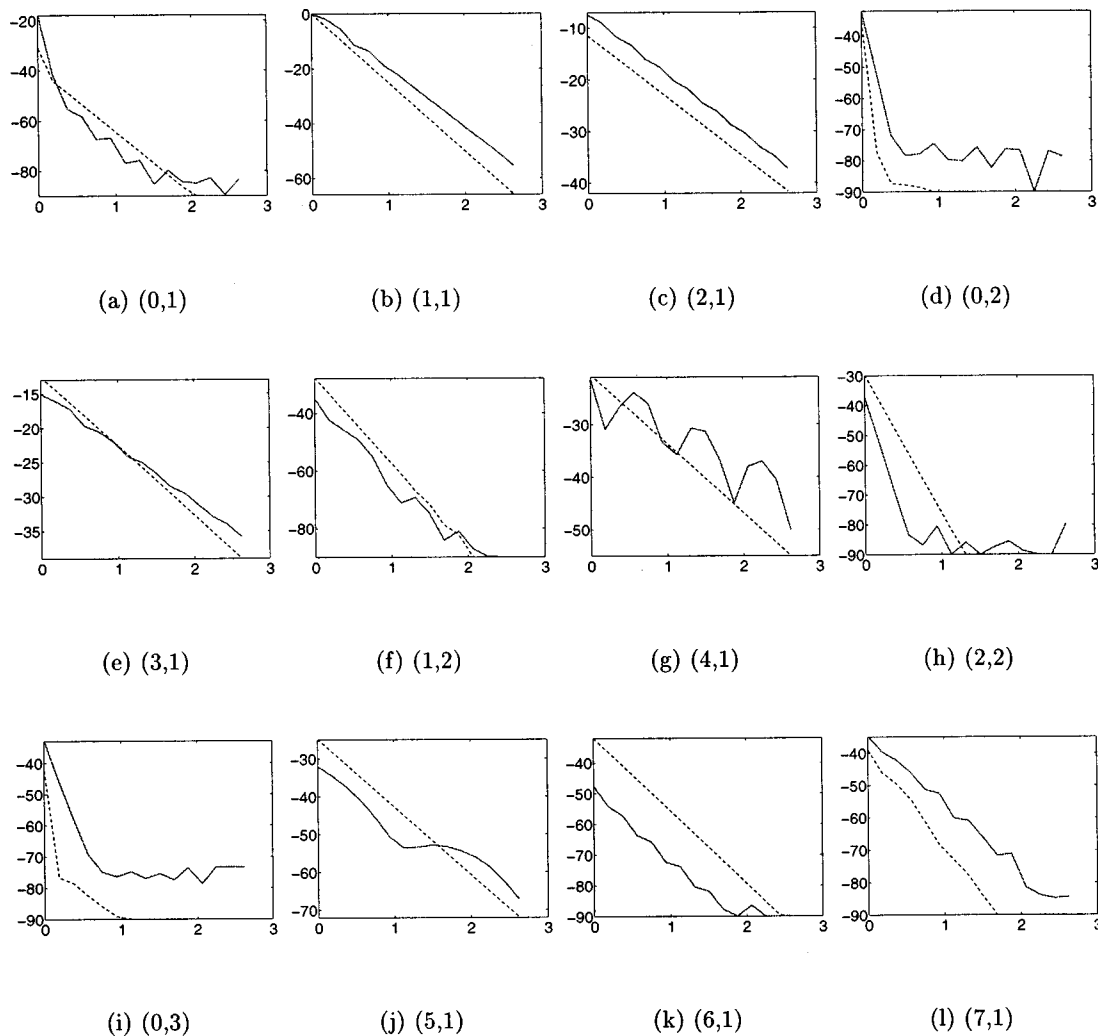


FIG. 15. Comparison between measured (solid lines) and simulated (dotted lines) decay curves of the main partials (magnitude in dB versus time in seconds). The corresponding mode is written below each plot.

Table III, the FFT is made on the initial portion of the sound (0.375 s), only, whereas the total length of the signal (3 s) is used for the averaged spectra shown in the first column of Fig. 13.

### C. Comparison with a previous study by Davis<sup>20</sup>

Calculated and measured modal frequencies (see Table IV) and decay times (see Table V) are compared with the

results of a previous study by Davis. The theoretical results in the frequency domain were obtained by this author using Green function and boundary integral methods. The measurements were made on a 26-in.-diameter Ludwig kettledrum. This author investigated theoretically the influence of various kettle shapes.<sup>20</sup> Only the cylindrical and ellipsoidal shapes are discussed here.

The present work is dealing with a Premier 25-in. kettle-

TABLE IV. Modal frequencies  $f_{mn}$  of the first ten partials for the complete kettledrum. Comparison between the results obtained by Davis (Ref. 20) and by the present model, for comparable  $f_{11}$ .

Mode (mn)	Davis (calculated) cylindrical kettle (Hz)	Davis (calculated) ellipsoidal kettle (Hz)	Davis (measured) Ludwig 26-inch (Hz)	Present work (calculated) ellipsoidal kettle (Hz)	Present work (measured) premier 25-inch (Hz)
01	126	109	128	136	139
11	145	145	145	147	147
21	220	218	218	221	221
02	244	233	235	248	245
31	289	287	287	288	288
12	342	302	303	315	315
41	357	355	354	357	355
22	397	394	394	395	395
03	405	418	383	403	408
51	423	422	421	424	419

TABLE V. Decay times  $\tau_{60}$  of the first ten partials for the complete kettledrum. Comparison between the results obtained by Davis (Ref. 20) and by the present model, for comparable  $f_{11}$ .

Mode (mn)	$\tau_{60}$	$\tau_{60}$	$\tau_{60}$	$\tau_{60}$
	Davis (calculated) cylindrical kettle (s)	Davis (measured) Ludwig 26-inch (s)	Present work (calculated) ellipsoidal kettle (s)	Present work (measured) premier 25-inch (s)
01	0.06	0.4	0.8	0.4
11	1.4	2.3	2.4	2.8
21	4.7	3.7	5.0	4.2
02	0.2	0.3	0.25	0.5
31	15.4	4.6	6.0	8.4
12	0.1	2.5	2.0	2.4
41	31.6	4.3	4.8	4.6
22	1.0	0.9	1.3	0.9
03	1.1	0.5	0.6	1.0
51	33.5	4.1	3.4	5.2

drum. Fortunately, the tensions used (3710 for Davis, 3325 N/m here) are such that the frequency  $f_{11}$  is almost the same in both cases (145 for Davis, 147 Hz here). As a consequence, valuable comparisons can be made between the two different works.

The comparison between the two methods presented in Table IV shows that the results obtained for the first ten modal frequencies are very similar, within 2%, except for the 01, 12, and 03 modes. For the 01 mode, the cylindrical kettle model used by Davis yields a better prediction for the measured modal frequency than the ellipsoidal model. However, an opposite conclusion has to be drawn for the 12 mode. In addition, one can notice that the 01, 02, 12, and 03 modal frequencies measured on the Ludwig kettledrum are lower (between 4% and 9%) than the corresponding frequencies measured on the Premier kettledrum. This is certainly a consequence of the differences in size between the two kettle volumes (0.14 for Davis, 0.10 m<sup>3</sup> here).

The comparison between calculated and measured decay times  $\tau_{60}$  for both kettledrums is shown in Table V. This quantity is defined as the time for a 60-dB attenuation of the sound pressure. Here again, the theoretical values obtained by Davis, for a cylindrical kettle, are calculated by means of a frequency domain method. In our case, the decay times are estimated from the decay curves presented in Fig. 15.

The decay times measured on both kettledrums are similar (within 30%), except for the 02, 31, and 03 modes. This can be considered as a remarkable result, in view of the significant geometrical differences between the two instruments. However, the predictions are significantly better with the present model than with the model used by Davis, especially for the 01, 31, 41, and 51 modes. This follows essentially from the fact that our model includes a viscoelastic term to account for the losses in the membrane, whereas only sound radiation is included in the calculations made by Davis. In this context, the comparison between the two sets of calculated values yields useful information on the effect of radiation, compared to the effect of dissipation in the membrane, on the decay times.

## V. DISCUSSION AND CONCLUSIONS

The model of the kettledrum presented in this paper is used for illustrating the time-domain evolution of the membrane displacement, of the sound field and of the pressure jump, starting from the impact by the mallet. The time histories of mallet, membrane, and acoustic energies are also calculated, which yield interesting information on the energy exchange between the constituent parts of the instrument.

The simulated external pressure is compared to the sound pressure radiated by a real instrument. The values of the parameters are derived from measurements performed on a hand-tuned Premier kettledrum. The comparison between measured and simulated waveforms and spectra shows that the model succeeds in capturing the main features of real sounds: the eigenfrequencies are reproduced within a few percent and the decay times of the main partials look very similar. In addition, it has been shown to what extent an improper tensioning of the membrane accounts for the temporal envelope of the sound and for peak doubling in the spectrum.

The simulation results obtained for the modal frequencies of the kettledrum are in good agreement with theoretical frequency domain predictions obtained by Davis in a previous study.<sup>20</sup> Comparison of decay times between the present time-domain model and the frequency domain method by Davis shows the relative importance of radiation and viscoelasticity of the membrane, respectively, in the damping process.

In the present state of the model, a given sound is defined by nearly 15 physical and geometrical parameters (see Table I). However, through our simulations, it has been found that the model is more sensitive to some parameters than to the others: small modifications (of about 20%) in the viscoelastic damping factor  $\eta$ , for example, alter the decay times significantly and yield clearly perceivable differences between real and simulated sounds. It has been also observed that similar results in the spectrum could be obtained with different sets of parameters: small variations of the sound speed  $c_a$ , for example, can be compensated by changing slightly the tension of the membrane.

The modeling of the mallet is similar to the one previously used in the past for pianos and xylophone.<sup>12,17</sup> As for these two instruments, convincing results are obtained here despite the simplicity of the model. However, in the case of timpani, one may argue that, due to both the relative softness and dimensions of the head, it should be necessary to take into account the fact that the surface of contact varies with time during the impact. This question remains open.

The modeling of internal losses in the membrane with the help of only one coefficient (the viscoelastic loss factor  $\eta$ ) yields surprisingly good results, although an intuitive physical analysis of the problem suggests that other causes of losses may exist in the mallet's felt and at the rubber edge of the membrane. The results seem to indicate that these two latter effects are either of second order, compared to the viscoelasticity of the membrane, or that they can be compensated, to a first approximation, by increasing  $\eta$ .

The relevance of the kettle shape in terms of sound quality is a question of interest for timpanists and makers.<sup>8</sup> In this paper, an ellipsoidal shape has been selected for facilitating the definition of the mesh. It should be mentioned, however, that there are no basic limitations in the model with regard to the kettle shape. A comparison between existing and various mathematically defined shapes could be the goal of a future work. This work could also include a study on the function of the hole at the bottom of the kettle.

Finally, one can think of other possible extensions of the present model for time-domain investigation of acoustical systems made of a membrane stretched over an air cavity. This could be, for example, useful for the modeling of eardrum, microphone, or wall resonators. In this case, the impact excitation would have to be replaced by an acoustic excitation.

## APPENDIX A: FICTITIOUS DOMAIN FORMULATION

In this Appendix, the variational formulation of our problem, written in a fictitious domain, is presented and justified.

By multiplying Eq. (5) by a test function  $w^*$ , and integrating on  $\Sigma$ , one obtains

$$\begin{aligned} \frac{d^2}{dt^2} \int_{\Sigma} \sigma w w^* &= \int_{\Sigma} \operatorname{div}(T \nabla w) w^* + \eta \frac{d}{dt} \int_{\Sigma} \operatorname{div}(T \nabla w) w^* \\ &\quad - \int_{\Sigma} f(t) g w^* - \int_{\Sigma} [p]_{|\Sigma} w^*. \end{aligned}$$

After integration by parts, one easily checks that, when  $w^*$  is set equal to zero on  $\partial \Sigma$ , this equation becomes

$$\begin{aligned} \frac{d^2}{dt^2} \int_{\Sigma} \sigma w w^* + \int_{\Sigma} T \nabla w \nabla w^* + \eta \frac{d}{dt} \int_{\Sigma} T \nabla w \nabla w^* \\ = - \int_{\Sigma} f(t) g w^* - \int_{\Sigma} [p]_{|\Sigma} w^*. \end{aligned} \quad (\text{A1})$$

Multiplying Eq. (9) by a test function  $p^*$  and integrating on  $\Omega_j$ , yields

$$\frac{1}{c_a^2 \rho_a} \frac{d}{dt} \int_{\Omega_j} p_j p^* = - \int_{\Omega_j} \operatorname{div} \mathbf{V}_j p^* \quad \text{for } j=e, i. \quad (\text{A2})$$

Applying a similar treatment to Eq. (10) and using the Green formula,

$$\int_{\Omega_j} \nabla p_j \mathbf{V}^* + \int_{\Omega_j} p_j \operatorname{div} \mathbf{V}^* = \int_{\Gamma} \mathbf{V}^* \cdot \mathbf{n}_j p_j |_{\Gamma},$$

one obtains

$$\rho_a \frac{d}{dt} \int_{\Omega_j} \mathbf{V}_j \mathbf{V}^* - \int_{\Omega_j} p_j \operatorname{div} \mathbf{V}^* = - \int_{\Gamma} p_j |_{\Gamma} \mathbf{V}^* \cdot \mathbf{n}_j \quad \text{for } j=e, i. \quad (\text{A3})$$

The sum in Eqs. (A3) and (A2), for  $j=e$  and  $j=i$ , are calculated, and the functions  $p$  and  $\mathbf{V}$  on the complete space  $\Omega$  are defined by their restrictions to  $\Omega_j$  ( $j=e, i$ ):  $p|_{\Omega_j} = p_j$  and  $\mathbf{V}|_{\Omega_j} = \mathbf{V}_j$ . By combining the resulting equations together with Eq. (A1), the following system is obtained:

$$\begin{aligned} \text{a) } \frac{d^2}{dt^2} \int_{\Sigma} \sigma w w^* + \int_{\Sigma} T \nabla w \nabla w^* + \eta \frac{d}{dt} \int_{\Sigma} T \nabla w \nabla w^* \\ = - \int_{\Sigma} f(t) g w^* - \int_{\Sigma} [p]_{|\Sigma} w^*, \end{aligned} \quad (\text{A4})$$

$$\text{b) } \rho_a \frac{d}{dt} \int_{\Omega} \mathbf{V} \mathbf{V}^* - \int_{\Omega} p \operatorname{div} \mathbf{V}^* = \int_{\Gamma} [p]_{|\Gamma} \mathbf{V}^* \cdot \mathbf{n}, \quad (\text{A5})$$

$$\text{c) } \frac{1}{c_a^2 \rho_a} \frac{d}{dt} \int_{\Omega} p p^* + \int_{\Omega} \operatorname{div} \mathbf{V} p^* = 0. \quad (\text{A6})$$

In this formulation, the main difficulty follows from the term  $[p]$  which is, by definition, the pressure jump over the boundaries of the instrument. This term implies knowing both external and internal pressures on the boundaries and thus requires constructing a volume mesh that fits the instrument. The fictitious domain method consists of introducing the new unknown  $\lambda = [p]_{|\Gamma}$  defined on the surface  $\Gamma$  of the instrument. System (A4)–(A6) then becomes

$$\begin{aligned} \text{a) } \frac{d^2}{dt^2} \int_{\Sigma} \sigma w w^* + \int_{\Sigma} T \nabla w \nabla w^* + \eta \frac{d}{dt} \int_{\Sigma} T \nabla w \nabla w^* \\ = - \int_{\Sigma} f(t) g w^* - \int_{\Sigma} \lambda w^*, \end{aligned} \quad (\text{A7})$$

$$\text{b) } \rho_a \frac{d}{dt} \int_{\Omega} \mathbf{V} \mathbf{V}^* - \int_{\Omega} p \operatorname{div} \mathbf{V}^* = \int_{\Gamma} \lambda \mathbf{V}^* \cdot \mathbf{n}, \quad (\text{A8})$$

$$\text{c) } \frac{1}{c_a^2 \rho_a} \frac{d}{dt} \int_{\Omega} p p^* + \int_{\Omega} \operatorname{div} \mathbf{V} p^* = 0. \quad (\text{A9})$$

It remains to take into account the boundary conditions on the surface of the instrument,

$$\begin{aligned} \mathbf{V} \cdot \mathbf{n} &= \frac{\partial w}{\partial t} \quad \text{on } \Sigma, \\ \mathbf{V} \cdot \mathbf{n} &= 0 \quad \text{on } \mathcal{C}. \end{aligned} \quad (\text{A10})$$

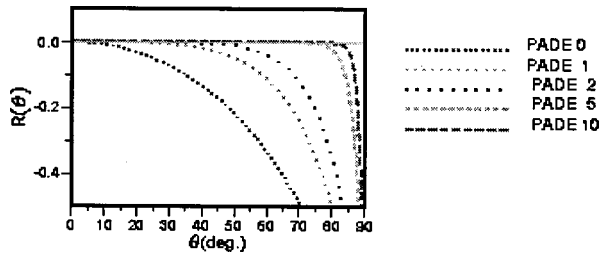


FIG. A1. Reflection coefficient of an incident plane wave on an absorbing boundary for various values of order  $L$  ( $L=0,1,2,5,10$ ).

Here, these conditions are taken into account in the weak sense, i.e., via an integration along  $\Gamma$ ,

$$\frac{d}{dt} \int_{\Sigma} w \lambda^* ds - \int_{\Gamma} \mathbf{V} \cdot \mathbf{n} \lambda^* ds = 0 \quad \forall \lambda^*.$$

This equation completes the system (A7)–(A9). Notice that, in this formulation, the space for  $\mathbf{V}$  (and therefore the mesh) can ignore the boundary of the instrument. This is a major advantage of this method, compared to more classical finite-element formulations.

## APPENDIX B: HIGHER-ORDER ABSORBING CONDITIONS

In this Appendix, the method used for simulating a free space for the sound field is presented. The basic principle of the method consists of adding artificial absorbing conditions on the boundaries of the computational domain.

### 1. The absorbing boundary condition for the half space

In order to reduce the computational domain of the wave equation,

$$\frac{\partial^2 p}{\partial t^2} - \left( \frac{\partial^2 p}{\partial x^2} + \frac{\partial^2 p}{\partial y^2} + \frac{\partial^2 p}{\partial z^2} \right) = 0, \quad \forall (x, y, z) \in \mathbb{R}^3,$$

to the half space  $z < 0$ , assuming that the sources are located in  $z < 0$ , an artificial boundary has to be imposed in the plane  $z = 0$ . (For simplicity, the speed of sound is supposed here to be equal to unity.)

The exact, or so-called “transparent,” condition is obtained via Fourier transform:  $p(x, y, z, t) \mapsto \hat{p}(k_x, k_y, z, \omega)$ . The solution satisfies

for  $z > 0$ ,

$$\hat{p}(k_x, k_y, z, \omega) = \hat{p}(k_x, k_y, 0, \omega) \exp \left( -i \omega \sqrt{1 - \frac{|k|^2}{\omega^2}} z \right),$$

$$|k|^2 = k_x^2 + k_y^2,$$

where the square root has been appropriately chosen to represent the outgoing waves. In  $z = 0$ , the transparent condition is written

$$\text{for } z = 0, \quad \frac{\partial \hat{p}}{\partial z} + i \omega \sqrt{1 - \frac{|k|^2}{\omega^2}} \hat{p} = 0.$$

The major drawback of this formulation is that it is non-local, both in space  $(x, y)$  and time. Therefore, the following approximation is preferred:

$$\frac{d \hat{p}}{dz} + i \omega \left( \gamma - \sum_{l=1}^L \frac{\beta_l}{\alpha_l} \frac{\omega^2}{\omega^2 - \alpha_l |k|^2} \right) \hat{p} = 0,$$

$$\gamma = 1 + \sum_{l=1}^L \frac{\beta_l}{\alpha_l}. \quad (\text{B1})$$

Equation (B1) corresponds to a rational approximation for  $u = (|k|^2/\omega^2) < 1$  of the function  $u \mapsto \sqrt{1-u}$ ,

$$\sqrt{1-u} \approx \left( 1 - \sum_{l=1}^L \beta_l \frac{u}{1 - \alpha_l u} \right)$$

$$= \gamma - \sum_{l=1}^L \frac{\beta_l}{\alpha_l} \frac{1}{1 - \alpha_l u} \quad \text{for } u \in [0, 1].$$

For stability reasons, the parameters  $(\alpha_l, \beta_l)$  must satisfy

$$\alpha_l > 0, \quad \beta_l \geq 0, \quad \sum_{l=1}^L \frac{\beta_l}{1 - \alpha_l} < 1.$$

Among the possible choices, the diagonal Padé approximation at  $u = 0$  can be used. This corresponds to

$$\beta_l = \frac{2}{2L+1} \sin^2 \left( \frac{l\pi}{2L+1} \right) \quad \text{and} \quad \alpha_l = \cos^2 \left( \frac{l\pi}{2L+1} \right). \quad (\text{B2})$$

By introducing auxiliary functions  $\psi_l^z(x, y, t)$  by means of their Fourier transforms,

$$\hat{\psi}_l^z = \frac{\omega^2}{\omega^2 - \alpha_l |k|^2} \hat{p}.$$

Then Eq. (B1) can be rewritten as the following system:

$$\begin{cases} \frac{d \hat{p}}{dz} + i \omega \left( \gamma \hat{u} - \sum_{l=1}^L \frac{\beta_l}{\alpha_l} \hat{\psi}_l^z \right) = 0 \\ \omega^2 \hat{\psi}_l^z - \alpha_l |k|^2 \hat{\psi}_l^z = \omega^2 \hat{p}, \quad |k|^2 = k_x^2 + k_y^2. \end{cases} \quad (\text{B3})$$

In the time-domain, this corresponds to

$$\begin{cases} \frac{\partial p}{\partial z} + \gamma \frac{\partial p}{\partial t} - \sum_{l=1}^L \frac{\beta_l}{\alpha_l} \frac{\partial \psi_l^z}{\partial t} = 0 \\ \frac{\partial^2 \psi_l^z}{\partial t^2} - \alpha_l \left( \frac{\partial^2 \psi_l^z}{\partial x^2} + \frac{\partial^2 \psi_l^z}{\partial y^2} \right) = \left( \frac{\partial^2 p}{\partial t^2} \right)_{/z=0}, \quad l = 1, \dots, L. \end{cases} \quad (\text{B4})$$

In Eq. (B4), the functions  $\psi_l^z$  are only defined on the boundary. The system can be seen as a transport equation in the  $z$  direction coupled to  $L$  bidimensional waves' equations on  $z = 0$  via the functions  $\psi_l^z$ .

To illustrate the performance of the method, the reflection of an incident plane wave, with angle of incidence  $\theta$  on such an absorbing boundary, is computed. The incident wave is reflected with a reflection coefficient  $R(\theta)$ . In Fig. A1, the

variations of  $R(\theta)$  as a function of  $\theta \in [0, (\pi/2)]$  are plotted, with Padé approximants and various values of  $L$  ( $L = 0, 1, 2, 5, 10$ ).

## 2. Simulation of a rectangular anechoic room

In order to reduce the computation to a finite rectangular room, one has to simulate an absorbing condition similar to Eq. (B4) on each wall. However, because of the presence of the tangential Laplacian, these equations have to be complemented by edge conditions. As an example, the edge condition on the line  $y=0, z=0$  is presented, with the computational domain in the quarter space  $y < 0, z < 0$ . The main problem is to determine compatibility conditions at the edge  $y=0, z=0$  between the auxiliary functions  $\psi_l^y(x, z=0, t)$  and  $\psi_l^z(x, y=0, t)$ . This leads to the introduction of  $2L^2$  edge functions  $\psi_{m,l}^{(y,x)}$  and  $\psi_{m,l}^{(z,x)}$ . Then, the edge conditions can be written

$$\begin{aligned} \frac{\partial \psi_l^y}{\partial z}(x, z=0, t) &= -\gamma \frac{\partial \psi_l^y}{\partial t}(x, z=0, t) \\ &+ \sum_{m=1}^L \frac{\beta_m \gamma_{l,m}}{\alpha_m^2} \frac{\partial \psi_{l,m}^{(y,x)}}{\partial t}(x, t) \\ &+ \sum_{m=1}^L \frac{\beta_m \gamma_{l,m}}{\alpha_m \alpha_l} \frac{\partial \psi_{m,l}^{(z,x)}}{\partial t}(x, t) \end{aligned}$$

and

$$\begin{aligned} \frac{\partial \psi_l^z}{\partial y}(x, y=0, t) &= -\gamma \frac{\partial \psi_l^z}{\partial t}(x, y=0, t) \\ &+ \sum_{m=1}^L \frac{\beta_m \gamma_{l,m}}{\alpha_m^2} \frac{\partial \psi_{l,m}^{(z,x)}}{\partial t}(x, t) \\ &+ \sum_{m=1}^L \frac{\beta_m \gamma_{l,m}}{\alpha_m \alpha_l} \frac{\partial \psi_{m,l}^{(y,x)}}{\partial t}(x, t), \end{aligned}$$

where  $\gamma_{l,m} = \alpha_l \alpha_m / (\alpha_l + \alpha_m - \alpha_l \alpha_m)$  and  $\psi_{m,l}^{(z,x)}$  and  $\psi_{m,l}^{(y,x)}$  satisfy the one-dimensional wave equations

$$\begin{aligned} \frac{\partial^2 \psi_{l,m}^{(z,x)}}{\partial t^2} - \gamma_{l,m} \frac{\partial^2 \psi_{l,m}^{(z,x)}}{\partial x^2} &= \frac{\partial^2 \psi_l^z}{\partial t^2}(x, y=0, t) \\ \frac{\partial^2 \psi_{l,m}^{(y,x)}}{\partial t^2} - \gamma_{l,m} \frac{\partial^2 \psi_{l,m}^{(y,x)}}{\partial x^2} &= \frac{\partial^2 \psi_l^y}{\partial t^2}(x, z=0, t). \end{aligned}$$

Finally, because of the second-order derivative along the edge, corner conditions are also needed. If, for example, the case of the domain  $x < 0, y < 0, z < 0$  is considered, then  $6L^2$  corner conditions at  $x=0, y=0, z=0$  are obtained (see Ref. 11). These conditions relate the auxiliary functions on the three edges.

The first  $L^2$  conditions are given by

$$\begin{aligned} \frac{\partial \psi_{l,m}^{(z,x)}}{\partial x}(0, t) &= -\gamma \frac{\partial \psi_{l,m}^{(z,x)}}{\partial t}(0, t) + \sum_{n=1}^L \frac{\beta_n \gamma_{l,m,n}}{\alpha_n^2} \frac{\partial \psi_{l,m}^{(z,x)}}{\partial t}(0, t) \\ &+ \sum_{n=1}^L \frac{\beta_n \gamma_{l,m,n} \gamma_{n,l}}{\alpha_n \alpha_l \gamma_{l,m}} \frac{\partial \psi_{n,l}^{(x,y)}}{\partial t}(0, t) \\ &+ \sum_{n=1}^L \frac{\beta_n \gamma_{l,m,n} \gamma_{n,l}}{\alpha_n^2 \gamma_{l,m}} \frac{\partial \psi_{l,n}^{(z,y)}}{\partial t}(0, t), \end{aligned}$$

where  $\gamma_{l,m,n} = \gamma_{l,m} \alpha_n / (\alpha_n + \gamma_{l,m} - \alpha_n \gamma_{l,m})$ . By interchanging the roles of  $x$  and  $z$ ,  $y$  remaining constant,  $L^2$  additional conditions are obtained, and the remaining  $4L^2$  conditions are obtained by circular permutation on  $x, y, z$ . In the present calculations, the computational domain is a rectangular room. In this case, edge and wedge conditions, similar to those presented above, must be fulfilled.

In practice, the interior scheme is used for computing the values of the pressure  $p$  in the cubes which are strictly situated in the room, while the absorbing conditions yield the pressure in the cubes connected to the boundaries.

Readers who are interested in more details may download the technical report written by Collino<sup>11</sup> via ftp at the following address: <http://www-rocq.inria.fr/ondes/publi.html>.

- <sup>1</sup>T. D. Rossing, "Acoustics of percussion instruments: Part II," *Phys. Teach.* **15**, 278–288 (1977).
- <sup>2</sup>T. D. Rossing, "The physics of kettledrums," *Sci. Am.* **247**, 172–178 (1982).
- <sup>3</sup>N. H. Fletcher and T. D. Rossing, *Physics of Musical Instruments*, 2nd ed. (Springer-Verlag, New York, 1998).
- <sup>4</sup>L. Meirovitch, *Analytical Methods in Vibrations* (Macmillan, New York, 1967).
- <sup>5</sup>K. F. Graff, *Wave Motion in Elastic Solids* (Dover, New York, 1991).
- <sup>6</sup>R. S. Christian, R. E. Davis, A. Tubis, C. A. Anderson, R. I. Mills, and T. D. Rossing, "Effects of air loading on timpani membrane vibrations," *J. Acoust. Soc. Am.* **76**, 1336–1345 (1984).
- <sup>7</sup>A. Chaigne and A. Ramdane, "Numerical simulations of membrane-mallet interaction in kettledrums," Technical report, ENST, Paris, France, 1998.
- <sup>8</sup>A. Tubis and R. Davis, "Kettle-shape dependence of timpani normal modes," in *Proceedings of the 12th Int. Congress on Acoustics, Toronto* (1986), Vol. 3, pp. K2–7.
- <sup>9</sup>I. Bork, "Measuring the acoustical properties of mallets," *Appl. Acoust.* **30**, 207–218 (1990).
- <sup>10</sup>Donald L. Sullivan, "Accurate frequency tracking of timpani spectral line," *J. Acoust. Soc. Am.* **101**, 530–538 (1997).
- <sup>11</sup>F. Collino, "Conditions absorbantes d'ordre élevé pour l'équation des ondes 3-D," Technical Report 2932, INRIA, Rocquencourt, France, 1996.
- <sup>12</sup>A. Chaigne and A. Askenfelt, "Numerical simulations of piano strings. Part I. A physical model for a struck string using finite difference methods," *J. Acoust. Soc. Am.* **95**, 1112–1118 (1994).
- <sup>13</sup>F. Millot, F. Collino, and P. Joly, "Fictitious domain method for unsteady problems," *J. Comput. Phys.* **138**, 907–938 (1997).
- <sup>14</sup>R. Glowinski, T. W. Pan, and J. Periaux, "A fictitious domain method for Dirichlet problem and applications," *Comput. Methods Appl. Mech. Eng.* **111**, 283–304 (1994).
- <sup>15</sup>J. C. Nédélec, "Mixed finite elements in  $\mathbb{R}^3$ ," *Numer. Math.* **35**, 315–341 (1980).
- <sup>16</sup>L. Rhaouti, P. Joly, and A. Chaigne, "Numerical simulation of timpani," INRIA Sics Communication scientifique multimedia, Rocquencourt, 1998 (videotape).



- <sup>17</sup>A. Chaigne and V. Doutaut, "Numerical simulations of xylophones. I. Time-domain modeling of the vibrating bars," *J. Acoust. Soc. Am.* **101**, 539–557 (1997).
- <sup>18</sup>D. Hall, "Piano string excitation. Part VI: Nonlinear modeling," *J. Acoust. Soc. Am.* **92**, 95–105 (1992).
- <sup>19</sup>A. Chaigne, "On the use of finite differences for musical synthesis. Application to plucked stringed instruments," *J. Acoust.* **5**, 181–211 (1992).
- <sup>20</sup>R. E. Davis, "Mathematical modeling of the orchestral timpani," Ph.D. thesis, Physics Department, Purdue University, 1988.

# Preferred self-to-other ratios in choir singing

Sten Ternström<sup>a)</sup>

*Department of Speech, Music and Hearing, Royal Institute of Technology, Drottning Kristinas Väg 31, SE-100 44 Stockholm, Sweden*

(Received 13 August 1998; revised 24 February 1999; accepted 26 February 1999)

Choir singers need to hear their own voice in an adequate self-to-other ratio (SOR) over the rest of the choir. Knowing singers' preferences for SOR could facilitate the design of stages and of choral formations. In an experiment to study the preferred SOR, subjects sang sustained vowels together with synthesized choir sounds, whose loudness tracked that of their own voice. They could control the SOR simply by changing their distance to the microphone. At the most comfortable location, the SOR was measured. Experimental factors included unison and four-part tasks, three vowels and two levels of phonation frequency. The same experiment was run four times, using sopranos, altos, tenors, and basses, with stimulus tones adapted for each category. The preferred self-to-other ratios were found to be similar to SORs measured previously in actual performance, if a little higher. Preferences were quite narrow, typically  $\pm 2$  dB for each singer, but very different from singer to singer, with intrasubject means ranging from  $-1$  to  $+15$  dB. There was no significant difference between the unison and the four-part tasks, although this might have been caused by systematic differences in the stimulus sounds. Some effects of phonation frequency and vowel were significant, but interdependent and difficult to interpret. The results and their relevance to live choir singing are discussed. © 1999 Acoustical Society of America. [S0001-4966(99)01306-5]

PACS numbers: 43.75.Rs [WJS]

## INTRODUCTION

For a choir singer, it is important to hear one's own voice (self) as well as the rest of the choir (other), in some suitable proportion. If one's own voice is masked, errors may occur, both in pitch (Naylor, 1987) and in vowel timbre (Ternström and Sundberg, 1989). One's own voice can be masked in very reverberant rooms and/or close choir formations (Ternström and Sundberg, 1988). If the rest of the choir is masked, timing becomes a problem. This can be the case in open-air places or very spread-out formations.

The most thorough treatment of hearing of self/other of which I am aware is that of Naylor, who studied this problem in orchestras. His work supports the view that sound levels are more important than the temporal distribution of returned energy (i.e., early reflections versus reverberation).

In another major study, Daugherty (1996) studied the preferences and perceptions of choristers and auditors with regard to choral spacing and choral formation. This issue is closely related to hearing of self. A choir was asked to sing the same excerpt of music in systematically varied formations. Immediately afterwards, the choir singers filled out a questionnaire. The 160 auditors made paired comparisons of high-fidelity stereo recordings taken in an audience position in the auditorium. The experiment was exceptionally well controlled, with great attention to detail. For example, a videotaped conductor was used so as to minimize irrelevant variation between trials. From Daugherty's results, we note that choir singers preferred a wider spacing to the normal spacing, and that this preference was stronger than that for mixed formation over sectional formation. The auditors showed some preference for the choral sound produced in wider spacing.

The present investigation is concerned only with the sound-pressure levels of self (airborne) and other. The difference in level between self and other will be called the self-to-other ratio (SOR), expressed in decibels. A positive SOR means that a singer's ears are subjected to a higher sound pressure from his/her own voice than from the rest of the choir.

The individual choir singer can do little to influence the SOR at his own position. Raising one's own voice is rarely permissible, with regard to the music and to one's fellow singers. One might use the sheet music as a sound reflector and thereby gain a few decibels of feedback at high frequencies. It is also possible to block an ear with a fingertip, thereby increasing low-frequency feedback and also attenuating the choir; but this has significant disadvantages, both perceptual and aesthetic.

The SOR is, however, also influenced by choral formation and by the room acoustics. In theory, the SOR can be predicted for a given singer position, using computer models or even standard formulas of statistical acoustics. (The data needed for such a prediction would be the number of singers, their relative locations, their voice output powers, and the so-called equivalent absorption area of the room.) Therefore, if the SORs preferred by singers were known, guidelines might be suggested for stage design and choral formation. With knowledge about these relationships, choir directors would be better equipped to optimize conditions, e.g., by changing the spacing between singers and/or by enforcing alterations to the room's absorption.

The objectives of the present investigation were (1) to find out how large an SOR is preferred by choir singers, and (2) to determine whether such preferences are strong (manifest in a small range of acceptable SORs) or weak (large range).

<sup>a)</sup>Electronic mail: sten@speech.kth.se

## I. METHOD

### A. Overview

The experimental approach was to let individual subjects sing sustained vowels together with a synthesized choir, in various conditions of vowel,  $F_0$ , etc. While singing, they were to adjust the level of that choir for an optimal self-to-other ratio, which was then measured.

For the task of the subjects to be realistic yet simple, a fairly complex experimental setup was devised. Subjects stood alone, at least 1.5 meters in front of a wall, onto which four loudspeakers were mounted, equidistant from the subject's head. A microphone stood on a floor stand straight ahead of the subject. The sound level at this microphone was used to control the volume of a synthesized choir coming through the speakers. The subject's own voice was not fed to the loudspeakers.

The synthesis remained silent while the subject was silent. The sound-pressure level (SPL) of the "choir" was controlled one-to-one by the subject's voice SPL. As soon as the subject started to sing (sustained vowels), the choir would automatically follow suit. Any change in the subject's SPL would be accompanied by a corresponding change in the choir SPL. Hence, the SOR at the subject's ears was, in effect, independent of the subject's voice SPL. This is important, since it would be a difficult and musically irrelevant task to maintain a stable SPL of one's own voice.

By walking closer to the microphone, the subject could raise the relative level of the choir, thereby decreasing the SOR. Conversely, by backing away from the microphone, the subject could lower the level of the choir and increase the SOR. Subjects were asked to seek out the distance from the microphone at which the self and other sounds were perceived to be most suitably balanced in loudness. Having found the preferred position, the subject would hold down a button and continue singing until told to stop (at least 5 s). During this time, the SOR would be measured, using a separate pair of binaural microphones worn by the subject.

This measurement method has some convenient properties that make it robust:

- (i) It requires a fairly reverberant room rather than an anechoic room. Anechoic rooms are usually disliked by singers.
- (ii) Stimuli are presented over hi-fi loudspeakers, not headphones. Headphones disrupt the normal auditory feedback and fix the placement of the stimulus sound field relative to the subject's head.
- (iii) The front microphone signal is used only for acquiring the subject SPL, not for recording the subject's voice. The SOR control mechanism works regardless of how close the subject is to the microphone, and regardless of whether or not the subject is singing the intended vowel and the intended pitch.

In the rest of this methods section, I shall describe the technique for measuring the SOR; the experimental factors; the stimulus sounds; the signal processing; the subjects; the procedure; and the analysis.

### B. Measuring the self-to-other ratio

The measurement technique is described in detail by Ternström (1994), and relies on the use of lightweight binaural microphones worn in the subject's ears. The method allows the SOR experienced by a singer to be measured during normal singing, with little or no inconvenience to the singer. The method does not account for the contribution of bone-conducted sound; some consequences of this will be dealt with in the discussion. A brief description of the method will be given here.

The objective is to measure the sound-pressure levels at the singer's ears, both of the singer's own voice (self) and of the rest of the choir (other). Since these signals are always mixed, and are often quite similar, the task is not trivial. With one microphone at each ear, however, we note that the signal from one's own voice in the left and right channels will be practically identical and in phase at all frequencies. The sounds of the rest of the choir and of room reflections, on the other hand, are not in phase. In fact, they tend to be slightly out of phase, because sung sounds have an average spectral peak near the first formant at about 500 Hz, and at this frequency the propagation distance for sounds arriving sideways differs by about half a wavelength (out of phase) for the left and right ears.

These circumstances enable us to separate the SPLs of self and other with reasonable accuracy. By subtracting the right-ear signal  $R$  from the left-ear signal  $L$ , a new difference signal  $S$  is created in which the self component is entirely canceled. The level of this difference signal is proportional to the SPL of other and is essentially independent of the SPL of self. If instead the left and right signals are added together into a sum signal  $M$ , self will gain exactly twofold, or +6 dB, but other will gain only about 3 dB. Expressed in equations,

$$L = \text{self}_L + \text{other}_L, \quad R = \text{self}_R + \text{other}_R, \quad (1)$$

where

$$\text{self}_L = \text{self}_R, \quad \text{other}_L \neq \text{other}_R, \quad (2)$$

and from which we have

$$\begin{aligned} M = L + R &= \underbrace{2 \times \text{self}}_{+6 \text{ dB}} + \underbrace{\text{other}_L + \text{other}_R}_{+3 \text{ dB}} \\ S = L - R &= \underbrace{\text{self} - \text{self}}_0 + \underbrace{\text{other}_L - \text{other}_R}_{+3 \text{ dB}} \end{aligned} \quad (3)$$

The symbols  $M$  for mono and  $S$  for stereo are conventionally used by broadcasting engineers for denoting the sum and difference signals, respectively, of a stereo program. Provided that self is stronger than other, which is most often the case, the level of the  $M$  signal will be determined mostly by the SPL of self. We can thus approximate the self-to-other ratio with the difference in level between  $M$  and  $S$ , less 3 dB. This approximation can be refined to deal also with weak self signals, if the cross correlation of other (left) to other (right) can be measured and taken into account. The method

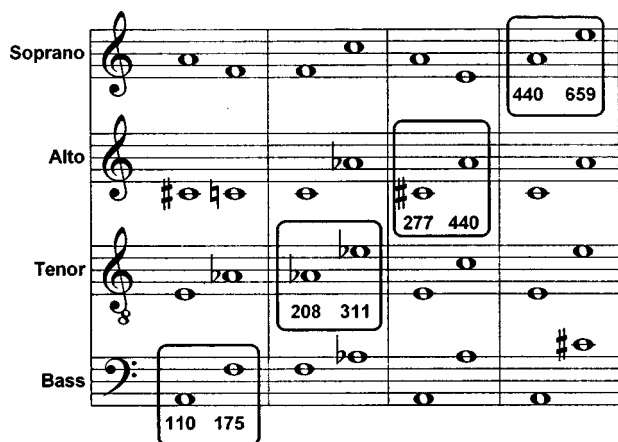


FIG. 1. Notes in boxes indicate the low and high tones sung by the subjects in both unison and chord contexts. Their fundamental frequencies are given in Hz. Notes outside boxes indicate the tones produced by the synthesized choir in the chord context. All tones were repeated on three vowels, each of which was replicated three times.

will henceforth be called the M/S measurement method for SOR.

Because of the random beating that is inherent in choral sounds, the  $M$  and  $S$  levels fluctuate considerably over the short term. With the integration time constants used here, the levels were averaged over at least 5 s to obtain stable readings. This in turn means that the method is not particularly sensitive to incidental extraneous sounds, such as brief traffic sounds and doors slamming elsewhere in the building. The operator could sit in the same room with the subjects.

### C. Experimental factors

The stimulus sounds representing the rest of the choir could not be taken from recordings, since the duration of each stimulus would have to depend on the time taken by a subject to decide on the preferred self-to-other ratio. Instead, the stimulus sounds were synthesized in real time. The digital synthesis facilities at our disposal are very flexible and allow for a real-time synthesis of up to three simultaneous sung vowel sounds, each available on a separate analog output.

This limitation to three synthesized voices suggested one series of stimuli with unison vowel sounds, and another series of “music-minus-one” stimuli of four-part chords, in which the subject sings the fourth part. Hence, one experimental factor was *context*, with two levels: unison and four-part, or chord. The hypothesis for this factor was that a higher SOR would be preferred in the unison case, in which the masking effect of other is stronger.

In order to introduce some variation in the task of the subjects, two levels of fundamental frequency  $F_0$  and three vowels were chosen. There was one  $F_0$  level in the lower half of the voice type  $F_0$  range and one in the upper half (Fig. 1). The three vowels were [u:], [a:] and [ɛ:].

For realism, tailor-made other sounds had to be produced for each of the four choral voice types, i.e., soprano, alto, tenor, and bass. A soprano will not find it meaningful to sing in a context designed for a bass. This separation of stimulus sets has the formal consequence of splitting the ex-

TABLE I. The experimental factors and their levels; also the types and number of replications.

Factor	Levels	No. of levels
Context	unison, chord	2
$F_0$	low, high	2
Vowel	[u:], [a:], [ɛ:]	3
Replications		
Loop gain (replicas)	-3, 0, +3 dB	3
Section (voice type)	S,A,T,B	4

periment into four subexperiments, which are different with regard to the experimental factors of vowel and  $F_0$ . These four will henceforth be called the experimental *sections*, one for each voice category. In comparisons across sections, data points may not be broken down by vowel and  $F_0$ , since the stimuli were not the same in these respects. Only the context factor (unison or not unison) may be considered consistent across sections.

In order to check for reproducibility, each stimulus sound was replicated three times, although with a *loop gain* that was modified by -3, 0, or +3 dB. To achieve the same SOR on all three, the subject would have to modify the microphone distance significantly for each one. This variation was introduced to encourage the subjects to reassess each token and not just stand still.

Since each subject had only one voice type, this makes a total of  $2 \times 2 \times 3 \times 3 = 36$  tokens per subject. The stimulus sounds were tailor-made for all factors and for voice types, making a total of  $2 \times 2 \times 3 \times 4 = 48$  combinations of synthesis parameters (Table I).

Only stationary sounds were used; that is, the task of the subjects was to sing not music but just sustained vowel sounds in unison or in a four-part chord. This constraint was adopted for several reasons:

- (i) it focuses the experiment on the question of sound levels only, and excludes such aspects of ensemble acoustics as may be related to the propagation timing of direct and reflected sounds,
- (ii) it avoids the complication of incidental variations in SOR that are concomitant with text-induced changes in vowel and voice effort, as for example when other and self are not perfectly synchronized,
- (iii) the SOR in stationary conditions can be assumed to be very closely related to the average SOR when performing actual music,
- (iv) it eliminates the need for subjects to learn to sing a particular excerpt of music,
- (v) it lets the subjects concentrate as purely as possible on balancing the level of their own voice (and, to a lesser extent, the timbre) against the rest of the ensemble.

Some further aspects of choosing sustained vowels will be mentioned in the discussion.

### D. Synthesis of stimulus sounds

Each of the three “voices” in the stimulus sounds was synthesized as follows. A conventional source-filter model

was used. The glottal source-pressure waveform was simulated using a new real-time algorithm that will be the subject of a forthcoming paper. The algorithm has one spectrum slope parameter, which corresponds well perceptually to the degree of vocal “press;” and an  $F_0$  parameter. The  $F_0$  of each source oscillator was perturbed independently by a flutter generator of a type described earlier by Ternström and Friberg (1989). The vocal tract filter was implemented as a chain of seven formant filters, followed by a +6 dB/octave pre-emphasis.

The formant frequencies, formant bandwidths, and source spectrum slope factors were found by starting with values known from the literature and from personal experience, and then adjusting them by ear. This work was assisted by an experienced singing teacher passing judgments on the subjective quality of the synthesis. The  $F_0$  and gain parameters were given by the experimental factor levels. The stimulus sounds were not equalized for SPL, since this would remove natural differences between vowels. The unison synthesis had 14 parameters and the three-part synthesis had 20 parameters. Complete listings of the parameter values used are available from the author, as are the setup script files and the parameter files that controlled the models. Readers who have access to the same signal-processing system can thereby recreate the setup.

### E. Signal processing

An overview of the experimental setup is given in Fig. 2. A commercial system for digital signal processing (DSP) was used to simultaneously generate and measure all the signals involved (*Aladdin Interactive DSP*<sup>TM</sup> from AB Nyvalla DSP). A signal processor add-on board (Loughborough Sound Images PC/C31) with four channels of true 16-bit analog input and output was mounted in a PC. The sampling rate was 16 kHz. The Aladdin model running on the DSP performed the following actions continuously and in parallel:

- (i) Synthesize the three sustained vowel sounds under parameter-file control, as described above.
- (ii) Advance to a new set of parameter values when the operator presses a button. This caused the next stimulus sound to be played for 1.6 s as a cue to the subject.
- (iii) Acquire the front microphone signal (subject’s voice) on one A/D converter, compute its level envelope, and use this level signal to control the level of the synthesis that is presented back to the subject.
- (iv) Acquire the state of the button that the subject holds down to indicate approval.
- (v) Acquire the left and right binaural microphone signals on two more A/D converters, and compute the self-to-other ratio as described above.
- (vi) Output results to a disk file with five channels, each sampled at 100 Hz: (1) the SOR, (2) the token number 1,...,18, (3) the sound level of the subject’s voice as registered by the front microphone, (4) the level of the *M* signal, approximating the *self* level, and (5) the “subject approves” switch signal.

An annotated screen image of the DSP block diagram is shown in Fig. 3.

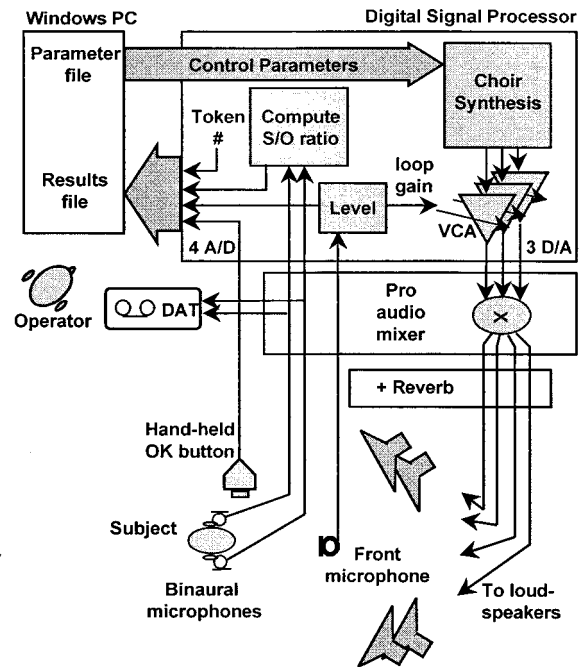


FIG. 2. Experimental setup, incorporating a PC complemented with a DSP board, studio quality audio equipment, four loudspeakers, and three microphones.

### F. Stimulus presentation

The experiment took place in a moderately reverberant classroom with a volume of about 300 m<sup>3</sup> and a reverberation time of 0.76 s. The stimuli were presented over four studio-quality self-powered loudspeakers arranged symmetrically in a vertical plane on a wall (Figs. 4 and 5). The wideband noise output level of all four loudspeakers was adjusted to be the same to within 0.5 dB on the central axis shown in Fig. 4. The M/S method cannot discriminate between the subject’s own voice and other sound sources located in the vertical plane of symmetry. Therefore, the three synthesized voices were counter phased across the loudspeakers such that there would be some stereo effect yet no apparent source in the center of the stereo image. This arrangement also improved the “surround-sound” effect of standing inside a choir, and helped reduce the risk for regenerating feedback via the front microphone, which in a pilot experiment was seen to be a potential problem. The counter phasing was also intended to somewhat increase the proportion of diffuse field sound at the subject’s ears. This is desirable since the M/S method works best when the left-ear and right-ear other signals are uncorrelated with each other. Finally, a small amount of artificial stereophonic reverberation was added, both for its decorrelating effect and to increase the realism of the synthetic choir. However, some correlation remained, and was dealt with as described below.

The order of presentation was randomized differently for each section and for both contexts, but not for each subject. That is, all sopranos received their stimuli in the same order, as did altos, tenors, and basses. The unison context task with 18 stimuli was performed first, followed by a short rest, and then the chord context task, also with 18 stimuli. The procedure took 18–25 min per subject.

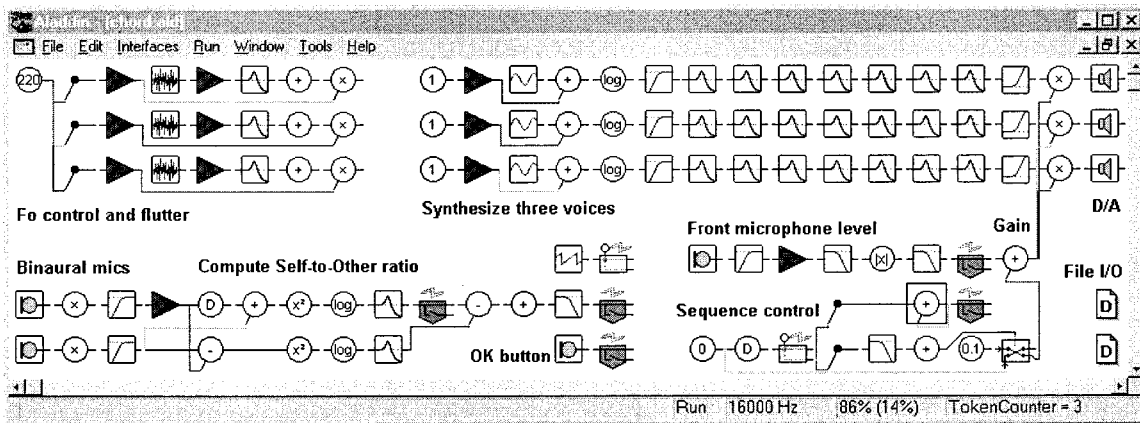


FIG. 3. DSP block diagram, showing the synthesis of three voices, *M/S* processing of the binaural microphone signal, and level processing of the front microphone.

### G. Subjects

Twenty-three subjects took part in the experiment. There were six each of basses and tenors, all males; and six altos and five sopranos, all females. All had at least 5 years of choral experience. They were aged 19–54 and were recruited from various choirs around Stockholm. They were paid for participating.

### H. Procedure

On arrival, each subject was given a sheet to read while waiting for their turn. The sheet carried some introductory information about hearing oneself, and a description of the task to come.

The subject was first fitted with lightweight binaural microphones (Sennheiser MKE2002) worn in the outer ears, and with a hand-held button for signaling approval. The archiving digital audiotape (DAT) recorder was then started and the subject was asked to sing, alone, a few bars of a simple song of his or her own choice. This was done to enable matching of the channel gains of the left and right microphone signals, for each subject. Then the DSP was started and the synthetic choir was demonstrated. The subject was asked to sing in unison with the synthesis and to explore the control of SOR by varying the distance to the front microphone. Once the subject appeared to understand the behavior of the control mechanism, the sequence of 18 unison sounds was started.

For each sound, the subject would first hear the synthesis on its own for about 2 s. Then the subject would sing the same sound, thereby invoking the choir again. He or she would approach the front microphone or recede from it, searching for a satisfactory balance between the own voice and the synthesized choir. At the optimum position, the subject would hold down the button, thereby illuminating a lamp at the operator's desk. The operator would count at least 5 s and then ask the subject to stop singing. The subject was allowed to take a new breath at any time.

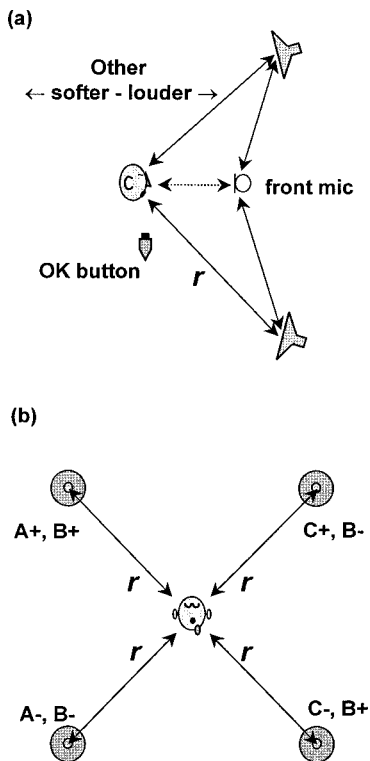


FIG. 4. Schematic arrangement of loudspeakers, front microphone, and singer subject. (a) Side view. The level picked up by the front microphone controls the level of other. (b) Front view. The 3 synthesized voices *A*, *B*, *C* were counterphased as shown, for reasons explained in the text.

### I. Analysis

An example printout of a small portion of an output file is shown in Fig. 6. Two such output files were created for each subject: one for the unison trial, and one for the chord

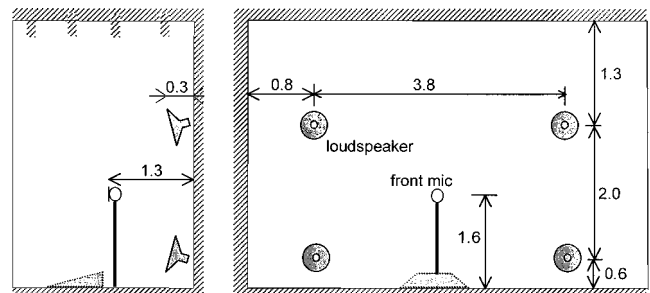


FIG. 5. Physical dimensions, in meters, of the loudspeaker setup and microphone arrangement. (a) Right-hand view; (b) front view.

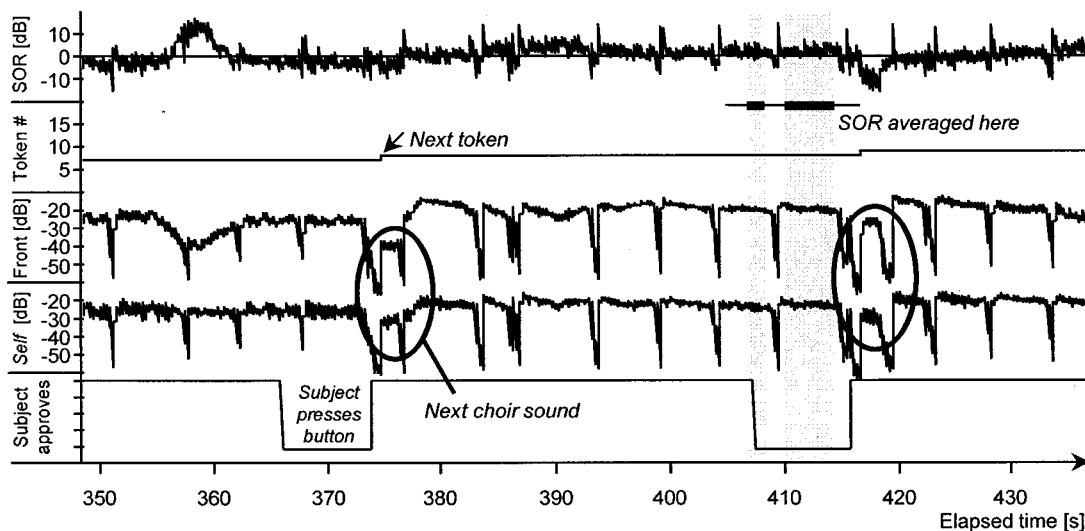


FIG. 6. Sample excerpt from a result file. The abrupt dips in self level and front mic level occur each time the singer pauses for breath. In the time interval 355–363 s, the singer recedes from the front microphone and then approaches it again, while maintaining a constant voice-effort level. This causes a drop in the front mic level, which lowers the level of other and thereby causes a corresponding rise in the self-to-other ratio (SOR). At 375 and 417 s, the operator advances to a new stimulus. The shaded areas mark the parts of the SOR curve that would be averaged for that token. The ellipses enclose the prompting choir sounds, during which the subject is silent. The rapid down-up transients in the SOR curve are due to the time constant of the level envelope that controls the level of the synthetic choir. Because the choir lags slightly behind the singer’s voice, the SOR makes excursions when the subject rapidly changes voice level. The sampling rate per channel is 100 Hz.

trial. The output files were analyzed for SOR, by averaging the SOR curve over the portions that had been approved by the subjects. These portions were marked manually. The averages of the marked sections were computed by an accessory program and entered into a spreadsheet.

### J. Accounting for left–right cross correlation of the other signal

The binaural method for separating the level of other from the level of self relies in the first approximation on the assumption that the two signals other (left) and other (right) are uncorrelated with each other. Several measures were taken to decorrelate these two signals, i.e., independent  $F_0$  randomization of the voices, counter phasing of the direct sounds (to reduce the ratio of direct to diffuse sound), and addition of some artificial reverberation. With only three voices in the synthesized choir, however, there might still be systematic variations in the cross correlation, due to changes in  $F_0$  and in the spectrum shape (vowel). Such variations would introduce systematic errors in the SOR readings. To study the effects of these two factors, therefore, some independent assessment was required of the left–right cross correlation in other.

When self is silent, any positively cross-correlated components of other will appear as a ‘phantom’ self. In a binaural signal that contains only other and no self, the cross correlation of left-to-right can thus be obtained directly from the SOR of that signal, henceforth called  $SOR_{\text{other}}$ .

Figure 7 shows how vector arithmetic can be applied to describe this quantitatively. With an SOR in decibels, this gives a cross-correlation estimate of

$$r = \frac{4}{\pi} \tan^{-1} (10^{(SOR+3.01)/20}) - 1, \quad (4)$$

and, with a little trigonometric exercise, the magnitude of the  $M$  vector becomes

$$|M| = \sqrt{2 + 2 \sin\left(\frac{\pi}{2} \cdot r\right)}, \quad (5)$$

while  $|S|$  is obtained if  $-r$  is substituted for  $r$ . The ratio  $|M| \div |S|$ , expressed in dB, is subtracted from the ‘raw’ SOR reading to correct it for the cross correlation in other.

In the experiment, each token started with a prompting choir sound, during which the singer nearly always remained silent, and which was identical to the subsequent stimulus. The  $SOR_{\text{other}}$  was measured for all these prompt sounds ( $N = 828$ ), yielding an average of at least 15  $SOR_{\text{other}}$  readings for each stimulus, from which the cross correlation was estimated.

There were indeed systematic variations in  $SOR_{\text{other}}$  with vowel and with  $F_0$ , but not with the loop-gain replications, nor with subject. In other words, each stimulus sound had an individual amount of cross-correlation  $r$ , which was quite stable across repeated presentations. In a few cases,  $r$  was rather far from zero, with extremes observed at  $-0.46$  and  $+0.65$ . When averaged over vowels and  $F_0$ , however, the  $SOR_{\text{other}}$  was  $-3$  to  $-6$  dB, corresponding to cross-correlation values of  $0$  to  $-0.2$ . This matches very closely the range of values observed in measurements on silent singers in choirs performing real music with lyrics and melody (Ternström, 1995). It also strengthens the earlier contention that the left and right other signals tend to be somewhat negatively correlated.

All SOR readings were compensated for the cross correlation in other before further processing.

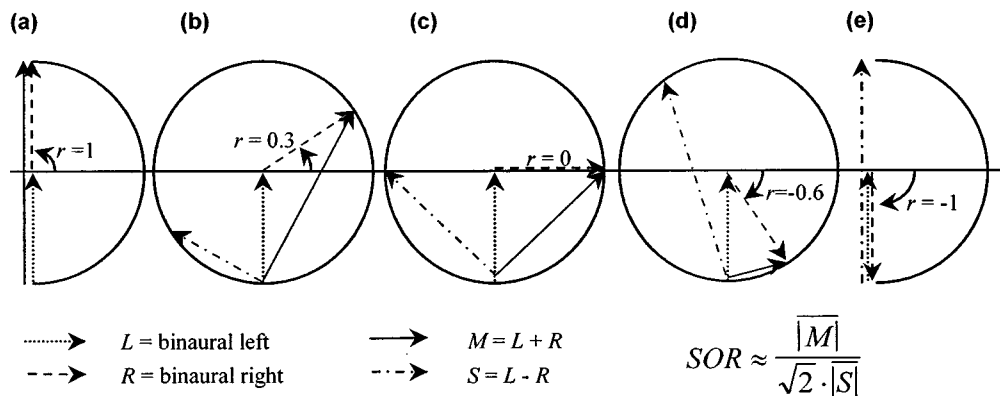


FIG. 7. Diagrams of the average sound-intensity vectors of the left and right signals at the singer's ears, with angle representing cross correlation; to illustrate the relationship between the cross-correlation  $r$  and the SOR. (a) is the case for the self signal, which is identical in both channels. (c) is a common configuration of the other signal, uncorrelated between  $L$  and  $R$ ; and  $S$  and  $M$  are of equal magnitude. (b) and (d) show how the levels of  $S$  and  $M$  become biased when  $L$  and  $R$  are partially correlated. In (e), the left and right signals are in exactly opposite phase, which never occurs in reality. The SOR is estimated as the SPL of  $M$  less the SPL of  $S$  less 3 dB ( $|M|/\sqrt{2} \cdot |S|$ ). Hence, the  $r$  of a binaural stereo signal can be computed trigonometrically from its SOR (assuming that the magnitudes of the  $L$  and  $R$  signal are approximately equal). A potential problem is that a truly null  $S$  signal would give an infinite SOR, but this can occur only in anechoic conditions. Elsewhere, any sound is reverberated in the room and will be interpreted as a weak other signal.

## II. RESULTS

### A. Main result

The average preferred SOR is shown in Fig. 8, broken down by subject and by context (unison/chord). Each point is an average over 18 tokens (3 vowels  $\times$  2  $F_0$  levels  $\times$  3 loop-gain levels). The grand total average preferred SOR ( $N = 828$ ) was +6.1 dB, with a mean intrasubject standard deviation of 3.1 dB.

### B. Validity

There was a remarkable amount of variation between subjects: intrasubject averages ranged from  $-1$  to  $+15$  dB. However, it is clear from Fig. 8 that most subjects are quite consistent across the unison/chord factor. It introduces much less variation than does a change of subject, even though the stimulus sounds were quite different for the unison and chord contexts. Hence, we may safely conclude that the observed intersubject variations in preferred SOR, though large, are real and reliable.

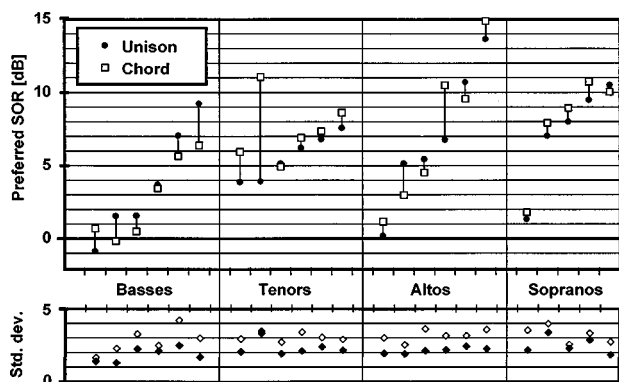


FIG. 8. Main result of experiment. Upper frame: Preferred self-to-other ratios (SOR), broken down by subject and by context (unison/chord). Each point is an average of 18 SOR readings (3 vowels  $\times$  2  $F_0$  levels  $\times$  3 replications). Lower frame: per-subject standard deviations. The larger values (open diamonds  $\diamond$ ) are the standard deviations across all tokens. The smaller values (filled diamonds  $\blacklozenge$ ) show the standard deviations due only to replications, with context, vowel, and  $F_0$  effects removed (see the text).

The choir was synthesized using a bare minimum of three voices generated using a fairly basic source-filter model running at a modest sampling rate. The sounds were recognizably those of a small choir, but their naturalness was in a few cases questionable. This may raise concerns about the realism of the subjects' task. It was therefore reassuring that several aspects of the results matched the SOR values found earlier in live choirs in actual performance:

- (i) the range of the preferred SORs was similar, if somewhat larger,
- (ii) the observed phantom SORs for silent subjects were quite the same,
- (iii) the tendency for lower SOR in the bass section and higher SOR in the soprano section was similar.

### C. Reproducibility and reliability

Over the course of a trial, each stimulus would occur a total of three times, identically presented except for the loop gain, which was modified by  $-3$ ,  $0$ , or  $+3$  dB. This variation was introduced as a control of reproducibility (are the subjects able to reproduce their own preferences?) and reliability (are we measuring the right thing?).

If the preferences for SOR are narrow *and* the subjects' control of SOR is adequate, then we would expect little variation in preferred SOR when only the loop gain is varied. Large variations, conversely, would be an indication of weak preferences and/or of poor control. Reproducibility was assessed by computing the variance  $V$  of the three replications  $x_i$  as

$$V = \frac{1}{3} \left[ \frac{(x_2 - x_1)^2}{2} + \frac{(x_3 - x_2)^2}{2} + \frac{(x_1 - x_3)^2}{2} \right]. \quad (6)$$

For each subject, this gave 12 observations of  $V$ , which were averaged into one per-subject variance  $V_{\text{subj}}$ . The square root of this value corresponds to a standard deviation in decibels, shown as s.d.<sub>rep</sub> in Fig. 8. This standard deviation represents the dispersion caused by the replications alone, while the dispersion due to the experimental factors is elimi-



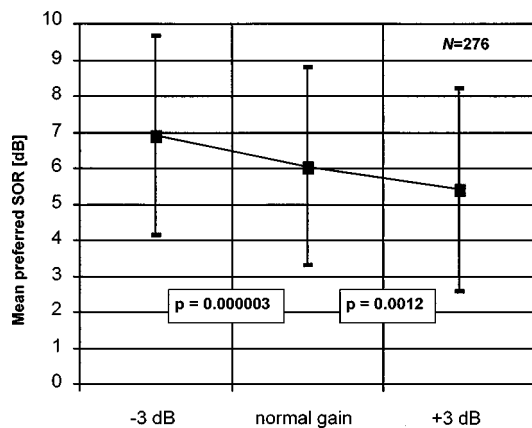


FIG. 9. The effect of changing the loop gain by +3 and -3 dB for the three replications (276 observations per gain level). On the average, the subjects compensated for the 3-dB changes in loop gain to within 1 decibel. However, there remains a small but highly significant effect of loop gain. This may be a side effect of small systematic changes in the voice effort level (see Sec. III D).

nated. The rms mean over all subjects of  $s.d._{rep}$  was 2.2 dB. This means that when conditions were replicated, subjects would reproduce the same SOR to within  $\pm 2.2$  dB, two times out of three.

If the subjects were to compensate perfectly for the changes in loop gain, the average preferred SOR would be the same for the three gain levels. The average effect of the loop gain therefore gives an indication as to what extent the subjects had understood the task. Figure 9 shows the SOR statistics broken down by loop-gain conditions -3, 0, and +3 dB. Although there was indeed a highly significant effect of loop-gain changes, its magnitude was less than 1 dB for a 3-dB change in gain. The task thus appears to have been very well understood. An explanation for the small effect of the loop gain will be suggested in the discussion.

#### D. Effects of vowel and $F_0$ were complex

For analyzing the effects of context, vowel and  $F_0$ , one analysis of variance was made for each of the four sections. These were three-factor analyses of variance (ANOVAs), using the three loop-gain settings as replications. The resulting significance levels of the factor effects are shown in Table II.

Surprisingly, there was no effect of context. I was expecting a higher SOR to be preferred in the unison context, in which one's own voice is more efficiently masked by other than in the chord context. However, this result may be in error, as will be discussed below. There were several significant effects on the preferred SOR of vowel and  $F_0$ , but their interactions were strong and diverse, making it hard to generalize. I was expecting, for example, the preferred SOR for the closed vowel [u:] to be different from that of the two open vowels, because of the stronger bone-conducted component of self for closed vowels (Letowski and Caravella, 1994). The observed effects were, however, quite different from section to section, and, when significant, they interacted strongly with context and  $F_0$  level. This is not really surprising, as a soprano tone is acoustically very different from a

TABLE II. Summary of ANOVA results: Probability levels  $p$  for rejection of the effects of the factors and their interactions.  $p < 0.05$  is deemed significant (boldface).

Factor	Sopranos	Altos	Tenors	Basses
Context	0.10	0.75	0.15	0.18
$F_0$	0.29	0.51	<b>0.001</b>	<b>0.009</b>
Vowel	0.07	<b>0.024</b>	<b>0.004</b>	<b>0.020</b>
$C^*F$	0.50	<b>0.009</b>	0.57	0.15
$C^*V$	0.23	<b>0.000 03</b>	<b>0.000 2</b>	<b>0.029</b>
$F^*V$	0.76	<b>0.028</b>	<b>0.002</b>	<b>0.034</b>
$C^*F^*V$	<b>0.001</b>	<b>0.013</b>	<b>0.000 01</b>	0.16

bass tone; and the masking effect of a unison ensemble sound is rather different from that of a three-part sound.

Nonetheless, two general observations could be made.

- (1) The effect of  $F_0$  level was significant for basses and tenors only. For both, the preferred SOR was higher on the higher  $F_0$ . The reason for this escapes me.
- (2) For sopranos, no single factor had any significant effect on the preferred SOR. This could be due to the fact that soprano sounds vary less with vowel, especially at high  $F_0$ , where all soprano vowels tend to approach [a:].

#### E. Experimental limitations

##### 1. SPL calibration was lost

The objective of this experiment was to measure preferred self-to-other ratios, without regard for absolute SPL. Although SPL calibration tones were recorded, the gain settings could not be stringently maintained throughout. With a few very loud voices, the gain had to be reduced on the spur of the moment to prevent clipping in the audio system. This was somewhat unfortunate, since it is not possible to report the absolute SPL of self produced by the subjects, which, in retrospect, could have been interesting. The absolute SPL of self may have some effect on SOR, given the possible interactions of SPL, bone conduction, and spectral slope, as discussed below.

##### 2. Bone conduction was not accounted for

The relative importance of the airborne and the bone-conducted feedback was investigated by Békésy (1949), who concluded that they contribute about equally to perceived loudness. Although the mechanisms of bone conduction have been researched in detail by Tonndorf (1972), he reports little data on humans as regards the relation between the airborne and the bone-conducted sound. A more recent overview is given by Howell (1985). Garber *et al.* (1981) investigated the amplification effect, i.e., the tendency of a speaker to lower the voice effort in the presence of amplification. Using low- and high-passed filtered feedback, they showed that this effect changes with the frequency content of the airborne signal. This change was ascribed to the low-frequency character of the bone-conducted sound.

In the present experiment, bone conduction was not accounted for. It is likely that the air-to-bone ratio increases with the the voice effort level, with bone conduction dominating in soft voice and/or on closed vowels, and airborne

feedback dominating in loud voice and open vowels. Hence, we may expect some effect of the SPL of self on the preferred SOR, but its investigation would require a different experiment.

For the purpose of describing the relationship between SOR and room acoustics for choirs, the omission of bone conduction is not a serious constraint, since the bone-conducted component of self does not change with the room acoustics. Although the SOR values reported here presumably would be higher if bone conduction had been taken into account, I would suggest that this fact is of limited practical use.

### 3. Spectrum slope was not accounted for

Although the synthesized choir tracked the sound level produced by the subject, its spectrum envelope remained constant. Since the spectrum slope of the subject's voice is not constant but varies, from soft to loud phonation, this may have some effect on the preferred SOR if the subjects choose to sing at very different voice-effort levels. It is important also to realize that the M/S method estimates the SOR in that frequency region which happens to dominate the spectrum. For vocal sounds this is the region of the first formant, which is vowel-dependent but usually near 500 Hz. This is not necessarily the frequency region in which a singer is most particular about the SOR.

## III. DISCUSSION

### A. Differences between choir sections

It is evident from Fig. 8 that the preferred SOR is different in the different sections. These section differences are similar to those observed when measuring long-term average SOR values inside real choirs in live performance (Fig. 10). The averages for the four sections are of the same magnitudes, and in both cases there is an overall trend of higher SOR values for sopranos and lower for basses. This suggests that the observed differences between the soprano, alto, tenor, and bass sections are probably not due to systematic errors in the synthesized stimulus sounds.

### B. Sustained chords as test sounds

In normal choir music, with constantly changing levels, pitches, and vowels, the SOR varies rapidly. In that situation, it is likely that a singer will be able to pick up more auditory clues regarding his or her own voice production than in sustained chords. On the other hand, the singer performing normal music is likely to be paying more attention to the execution of the music and the lyrics than to a precise balance of levels. When considering hearing-of-self in choirs, the sustained chord or sustained unison note can be said to represent a kind of "worst case" in which the overall audibility of one's own voice would rest almost entirely on the SOR, rather than on musical timing or spectral variation. Furthermore, auditory feedback is pivotal to correct and precise intonation (e.g., Elman, 1981), which, from a choral point of view, is particularly important in sustained chords. There-

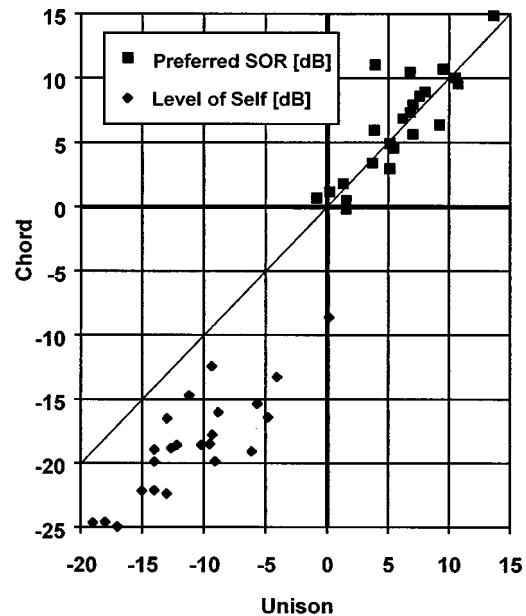


FIG. 10. Comparison of the relative level of self and of the preferred SOR between the unison and chord contexts. Each point is a mean of one subject over 18 tokens. All subjects sang softer (average  $-7.6$  dB) in the chord context (dots), probably because the loop gain of the synthesized choir was about 14 dB lower. However, the preferred SOR was much the same as in the unison context (squares).

fore, even if sustained chords are a special case, they are especially relevant when considering the SOR.

### C. Habit may be of importance

The large intrasubject variation in preferred SOR could be attributable in part to the habitual position of the singer in the choir. Although this issue was not formally investigated, the subject with the largest preferred SOR was used to standing at one end of the choir, while the subjects that preferred a negative SOR said they normally stand near the center of the choir. It seems possible, then, that SOR preference is at least in part acquired, and not determined by acoustical factors alone. One subject was particularly experienced, and freelances as a supporting choir singer in many elite choirs. He had very precise preferences, even though he frequently changes choir and position in the choir.

### D. Possible influences of the spectral variation in self

Even though the SPL calibration was lost (Sec. II E 1), it remained possible to measure the intrasubject *changes* in self level over the course of the two trials, by examining the *M* signal in the result files. This revealed that the mean self level was 7.6 dB lower in the chord context than in the unison context. The reason for this was probably that the loop gain for the chord stimuli was fully 14 dB lower than for the unison stimuli (through a technical oversight). Therefore, the prompting choir sound that preceded each token was noticeably softer in the chord trial, a circumstance which probably induced the subjects to sing in a softer nuance overall. By approaching the microphone, however, they could still obtain the same SPL of other and the same SOR as in

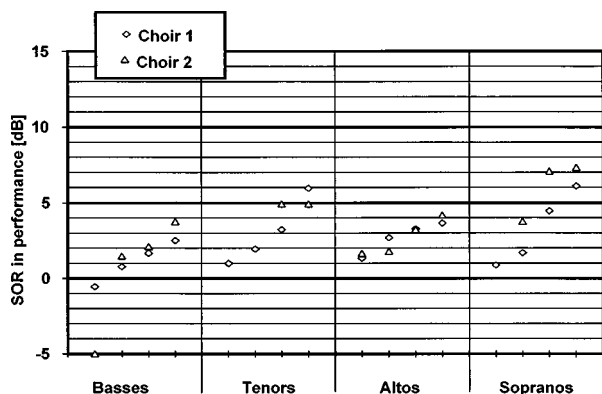


FIG. 11. SOR values encountered in actual performance inside live choirs, where singers have little control over SOR (adapted from Ternström, 1995). For comparison with Fig. 8.

the unison context, if they so wished. In fact, Fig. 11 shows that while the context factor did have a large effect on the SPL of self, it had hardly any effect on the preferred SOR.

A perceptive reviewer of this paper suggested that the effect of context on SOR might yet exist, but be somehow masked by this change in SPL of self. Also, there remains to account for the small but significant effect of the loop gain (Sec. II C). This prompted me to also examine the effect of the loop gain on the level of the  $M$  signal, being the best estimate of the SPL of self. On the average, a 3-dB increase of the loop gain was accompanied by an increase in SPL of self of about 0.6 dB, and a decrease in preferred SOR of 0.75 dB. Furthermore, this effect was much the same both in the unison and chord contexts.

If the spectrum slope of self were constant, as for other, we would expect no effect at all of the SPL of self. However, the spectrum slope of self is not only more negative than that of other, but also varies with the SPL of self, with the spectrum level gaining about 2 dB at 3 kHz for each 1-dB increase of voice SPL, and only about 0.5 dB in the region of the fundamental frequency (Fant, 1960). This spectral behavior was charted in detail for choral voices by Ternström (1993). As the subjects tended to raise their voice SPL a little with increasing loop gain (supposedly because of the louder prompt), the level of self at 2–4 kHz would increase more. This relative improvement in the feedback might explain some of the observed reduction in preferred SOR.

If this spectral variation of self is relevant, it may also explain to some extent why no effect of context on SOR was observed. The observed decrease in level of self going from unison to chord ( $-7.6$  dB) implies a larger decrease in the high-frequency region of self (and a smaller decrease in the level of the frequency region dominated by bone-conducted feedback). The larger loss of high frequencies in self might then cause the subjects to prefer a higher SOR; on the other hand, going to the chord context should lead to lower SOR preferences. Perhaps these effects canceled out; although this would imply a very strong preference for balance at high frequencies. Since the spectrum slope of self and the relative contribution of bone conduction were not known, it can be inferred only that the actual effect of context on SOR should have been no less than half of this change, or at least 4 dB down. This would be consistent with our expectations.

## E. SOR preferences are narrow

The standard deviation in SOR preference across replications was only 2.2 dB, including the loop-gain effect discussed in the previous section. This narrowness of preferences in SOR is remarkable, not least in view of the fact that the SPL of choral sounds always has large, random short-term variations due to beats.

## F. Some variance remains unexplained

A theoretical prediction of the SPL of other can be made by summing the acoustic power of many sources at known distances. For example, if the output power of all singers is the same, and if the choir stands in two rows, as is the most common case in Sweden, the level of other should be only 2–3 dB higher in the center of a choir compared to one end of the choir. The voice output power tends to be higher for sopranos and lower for basses than for the other voice types. If the singers are not crowded together, this difference would typically lead to higher SOR values inside a soprano section and lower SOR values in a bass section. These effects, however, do not seem sufficient to account for the large variations in preferred SOR observed here.

Small individual variations in hearing might account for some of the variability. Although no subject reported having hearing problems, a person can have a hearing loss of 5 or even 10 decibels and not know it. If the loss were sensorineural, it would affect airborne and bone-conducted sounds alike, and so should not affect SOR preferences. However, with a conductive hearing loss, SOR preferences should go down, since other would become weaker, while the bone-conducted portion of self would be largely unaffected. It seems that performing standard audiometry on the subjects would not shed much light on this, since its precision is not much higher than the SOR variability observed here.

## G. Implications for stage design, choral practice, and choral formation

The two major parameters that control the level of other in real-life situations are the equivalent absorption area of the room, which affects the intensity of the diffuse field; and the intersinger spacing, which affects the intensity of the direct sound from the other singers. There are also influences of local voice power variations within the choir and of the total number of singers in the room. The average preferred SOR in this experiment was +6.1 dB. This is somewhat higher than a mean actual SOR of +3.9 dB that was found earlier (Ternström, 1994) in the following conditions: a choir of 20 singers spaced by about 0.7 m in a single-row crescent, in a room with an equivalent absorption area of 65 m<sup>2</sup>. The SOR increases when the absorption increases, when the spacing increases, and when the number of singers decreases. For similar SOR values, a large choir is likely to require more floor area *per singer* than a small choir. This of course runs contrary to what is usually practical.

As shown by Coleman (1994), singers even in one and the same choir can vary greatly in their voice output power. The present experiment has shown that the SOR preferences, too, vary considerably. While these are complications, the

fact remains that singer preferences in a controllable situation are narrow, suggesting that there is something to be gained by trying to optimize for SOR. Choir directors might discuss the issue of hearing oneself with the singers in some detail, and perhaps ask one singer at a time to explore where in the choir he or she prefers to stand. Singers with soft voices will be at a somewhat greater advantage when placed at the ends of the choir. It is possible, if unconventional, to experiment with varying amounts of spacing in the center and at the ends of the choir. The more rows in which the choir is standing, the greater will be the influence of position within the choir.

#### H. Alternative measurement methods

For measuring SOR, the *M/S* method has several good properties: it is not affected by the subject's exact position and head orientation, nor by room acoustical details. The absolute levels of self and other do not have to be controlled. The method can be used in live performance, with little inconvenience to the measured singer and none to the rest of the choir. Since I had already collected data in live choirs using this method, it was the obvious choice also for the present production experiment. The main disadvantages of the *M/S* method are that the self/other discrimination depends somewhat on frequency and that the left-right correlation of other may need to be accounted for. However, both of these complications can be managed. Also, bone conduction is neglected, as discussed earlier.

When the other signal is known beforehand, as in this experiment, other techniques for measuring SOR are, of course, possible. Two examples will be given.

- (1) Place a microphone at one or both ears, and let the subjects stand still or sit throughout. Let them control the loop gain directly, e.g., using a rotary knob with a bias that changes randomly for each token. Calibrate the other level separately and account for the actual loop gain. The contribution of the subject's voice to the total ear-level SPL can be inferred if the other SPL is known for each stimulus. Since variations in head shape and exact microphone placement would affect the readings, this method would require the other SPL alone to be measured with a silent subject for all subjects and all unreplicated stimulus sounds. This would significantly increase the time required per subject.
- (2) Since the other signal is known, it is, in principle, possible to achieve a perfect *post hoc* separation of self and other. For each subject, one would measure the transfer functions (or impulse responses) all the way from synthesizer via speakers and room to ear-level microphones; then record simultaneously the stereo signal of the synthesis and that picked up by the microphones; and finally deconvolve the other sound from the self sound. This would have to be done for both the left and right channels. This method would require the subject to remain quite still throughout so as not to perturb the transfer functions. The data processing need not be done in real time, but would nevertheless be a substantial undertak-

ing for the long durations in this experiment; about 20 min of sound per subject.

#### IV. SUMMARY

The self-to-other ratio (SOR) preferred by choir singers when singing stationary tones was measured in a production experiment, with 23 individual choir singers, using as other a three-voice ensemble that was synthesized in real time. The singer subjects could control the SOR while singing. The task was well understood and the reproducibility was high.

Each subject made 36 adjustments of the SOR, half of which were sung in a unison ensemble context and the other half in a four-part chord context. In each context, six stimulus sounds differing in vowel and  $F_0$  were replicated three times each.

The key finding was that the choir singers had quite narrow preferences for SOR, with a standard deviation across replications of only  $\pm 2.2$  dB. Personal preferences, however, differed greatly. The average preferred SOR was +6.1 dB, with extremes at  $-1$  and  $+15$  dB. This is a few dB higher than average SORs observed in live performance.

Changing from a unison to a four-part chord context seemed to have no significant effect on the preferred SOR. However, such an effect may have been masked by systematic differences in voice-effort level, which was 7.6 dB lower in the chord context.

For sopranos, there was no effect of vowel nor of  $F_0$  level. In the other voice types, some vowel effects were significant, but interacted in complex ways with context and  $F_0$  level. Tenors and basses preferred a somewhat higher SOR on the upper  $F_0$  compared to the lower  $F_0$ .

#### ACKNOWLEDGMENTS

The participating subjects are gratefully acknowledged. Monica Thomasson kindly and competently assisted in judging the synthesis quality. Joakim Westerlund aided us with the statistical analysis. Several colleagues at KTH contributed with valuable discussions. My thanks go also to an anonymous reviewer for diligence and for pointing out the possible counteraction between the context factor and the SPL of self. This work was generously supported by the Swedish Natural Science Research Council and the Bank of Sweden Tercentenary Foundation.

- Coleman, R. (1994). "Dynamic intensity variations of individual choral singers," *J. Voice* 8(3), 196–201.
- Daugherty, J. F. (1996). "Spacing, formation and choral sound: preferences and perceptions of auditors and choristers," Ph.D. thesis, Florida State University at Tallahassee, School of Music.
- Elman, J. L. (1981). "Effects of frequency-shifted feedback on the pitch of vocal productions" *J. Acoust. Soc. Am.* 70, 45–50.
- Fant, G. (1960). *Acoustic Theory of Speech Production* (Mouton, The Hague).
- Garber, S. R., Siegel, G. M., and Pick, H. L. (1981). "Regulation of vocal intensity in the presence of feedback filtering and amplification." *J. Speech Hear. Res.* 24(1), 104–108.
- Howell, P. (1985). "Auditory feedback of the voice in singing," in *Musical Structure and Cognition*, edited by P. Howell, I. Cross, and R. West (Academic, London), pp. 259–286.
- Letowski, T., and Caravella, J. M. (1994). "Sound levels produced at and in the occluded ear of the talker," *Arch. Acoust.* 19(2), 139–146.

- Naylor, G. M. (1987). "Musical and Acoustical Influences upon the Achievement of Ensemble," Ph.D. thesis, Heriot-Watt University, Edinburgh, UK.
- Ternström, S. (1993). "Long-time average spectrum characteristics of different choirs in different rooms," *Voice (United Kingdom)* **2**, 55–77.
- Ternström, S. (1994). "Hearing myself with the others-sound levels in choral performance measured with separation of the own voice from the rest of the choir," *J. Voice* **8**(4), 293–302.
- Ternström, S. (1995). "Self-to-Other ratios measured in choral performance," in *Proceedings of 15th International Congress on Acoustics, ICA 95*, Trondheim, Norway, June 1995, Vol. II, pp. 681–684.
- Ternström, S., and Friberg, A. (1989). "Analysis and simulation of small variations in the fundamental frequency of sustained vowels," *Speech Transmission Laboratory Quarterly Progress and Status Report 3/1989*, pp. 1–14.
- Ternström, S., and Sundberg, J. (1988). "Intonation precision of choir singers," *J. Acoust. Soc. Am.* **84**, 59–69.
- Ternström, S., and Sundberg, J. (1989). "Formant frequencies of choir singers," *J. Acoust. Soc. Am.* **86**, 517–522.
- Tonnendorf, J. (1972). "Bone Conduction," in *Foundations of Modern Auditory Theory*, edited by J. V. Tobias (Academic, New York), Vol. 2, pp. 197–237.
- von Békésy, G. (1949). "The structure of the middle ear and the hearing of one's own voice by bone conduction," *J. Acoust. Soc. Am.* **21**, 217–232.

# The underwater audiogram of the West Indian manatee (*Trichechus manatus*)

Edmund R. Gerstein and Laura Gerstein

Florida Atlantic University, Department of Psychology, P.O. Box 3091, Boca Raton, Florida 33431-0991

Steven E. Forsythe

Navy Undersea Warfare Center Division Newport, Newport, Rhode Island 02841-1708

Joseph E. Blue

Leviathan Legacy, Inc., 3313 Northglenn Drive, Orlando, Florida 32806

(Received 16 October 1996; revised 27 December 1996; accepted 6 February 1999)

The hearing thresholds of two adult manatees were measured using a forced-choice two alternative paradigm and an up/down staircase psychometric method. This is the first behavioral audiogram measured for any Sirenian, as well as the first underwater infrasonic psychometric test with a marine mammal. Auditory thresholds were obtained from 0.4 to 46 kHz, and detection thresholds of possible vibrotactile origin were measured at 0.015–0.2 kHz. The U-shaped audiogram demonstrates an upper limit of functional hearing at 46 kHz with peak frequency sensitivity at 16 and 18 kHz (50 dB *re*: 1  $\mu$ Pa). The range of best hearing is 6–20 kHz (approximately 9 dB down from maximum sensitivity). Sensitivity falls 20 dB per octave below 0.8 kHz and approximately 40 dB per octave above 26 kHz. The audiogram demonstrates a wider range of hearing and greater sensitivity than was suggested from evoked potential and anatomical studies. High frequency sensitivity may be an adaptation to shallow water, where the propagation of low frequency sound is limited by physical boundary effects. Hearing abilities of manatees and other marine mammals may have also been shaped by ambient and thermal noise curves in the sea. Inadequate hearing sensitivity at low frequencies may be a contributing factor to the manatees' inability to effectively detect boat noise and avoid collisions with boats. © 1999 Acoustical Society of America. [S0001-4966(99)02005-6]

PACS numbers: 43.80.Lb, 43.80.Jz [FD]

## INTRODUCTION

Sirenian populations are threatened with extinction throughout the world. Only four species survive today: the dugong (*Dugong dugong*), the West African manatee (*Trichechus senegalensis*), the Amazonian manatee (*T. inunguis*), and the West Indian manatee (*T. manatus*). The current population of the Florida manatee, a West Indian manatee subspecies (*T. manatus latirostris*), is estimated to be little more than 2500 animals (Florida Department of Environmental Protection, unpublished 1996 manatee census).

With the population threatened, much of the research on living animals has focused on population biology. Sirenians have a unique phylogeny and aquatic lifestyle. Being the only obligate herbivores among the predacious marine mammals, they are of special interest, particularly with regard to the ontogeny and emergent capabilities of their sensory systems. Although manatees are readily observed in the wild and in captivity, few controlled or quantitative behavioral studies have been undertaken. With the exception of a few average evoked potential studies (Bullock *et al.*, 1980, 1982; Klishin *et al.*, 1990), no controlled quantified tests have been undertaken to measure the hearing sensitivity of manatees.

Although manatees have the cognitive and physical prowess to recognize and avoid boats (Gerstein, 1994, 1995; Hartman, 1979), individuals are injured repeatedly and often killed as a consequence of collisions with boats. A manatee audiogram is thus not only of academic interest, but could

prove essential to developing a strategy to protect this endangered marine mammal.

## I. METHODS

### A. Subjects

Two captive-born (8 and 9 year old) adult male manatees (Stormy and Dundee) were used for this study. They were both approximately 3 m long and each weighed about 400 kg. Complete medical records indicated neither animal was ever treated with ototoxic medications. Each subject received 45 kg of romaine lettuce per day, supplemented with vitamins and hydroponic sprouts. Preferred food items (monkey chow biscuits, carrots, and assorted aquatic plants) were effective as positive reinforcers during training and testing sessions.

### B. Facilities

The research was conducted at the Lowry Park Zoo in Tampa, Florida. The manatee facility consisted of five distinct pools connected by a water filled channel which enabled the segregation of Stormy and Dundee from other manatees and from each other during their individual training and testing. Watertight wooden gates helped to acoustically insulate the test pool from the adjoining channel and remote pools. The 492 050 l main exhibit pool served as the test tank. It had an irregular elongated shape, and a variable

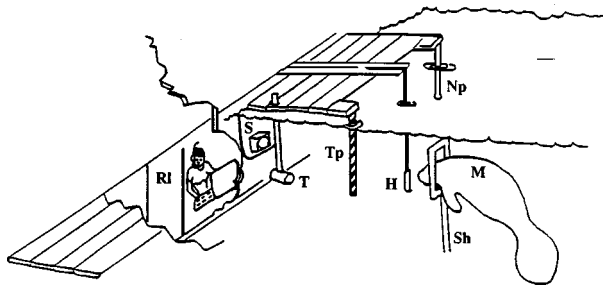


FIG. 1. Diagram of the experimental apparatus and test setup. Manatee (M); Stationing hoop (Sh); Hydrophone (H); Transducer (T); Strobe light (S); Research lab (RI); Tone paddle (Tp); No-tone paddle (Np).

water depth (1m–3m). Five underwater viewing panels provided excellent visual recording capabilities, while the irregular sides and bottom surface contour of rocks, ledges, and logs reduced reverberation levels in the pool to negligible levels for test purposes. Tests were conducted at mid-depth in 3 m of water, and where all lateral distances exceeded 3 m. Skimmers, return lines, and drains were turned off prior to and throughout testing to maintain consistent low ambient sound pressure levels and eliminate equipment related transients. Tests were conducted in the evening and early morning hours, while the zoo was closed, to minimize operational conflicts and sources of acoustic contamination. The in-ground facility was approximately 3 miles from the public highway. Infrasonic tests were restricted to very late evening and early morning periods in an effort to minimize contamination from distant ground sources. Signal levels received at the listening station were stable with fluctuations less than 1 dB. The resulting ambient sound pressure levels in the test tank were below Sea State 0 and proved to be within tolerances necessary to conduct all of the measurements for an audiogram.

An adjoining 23 m<sup>2</sup> underwater viewing area served as the observation and data collection lab, where all the controlling electronics were housed. This research lab had a 0.9×3 m underwater viewing window as well as interior conduits for electro-acoustical cabling.

### C. Apparatus

The test apparatus consisted of five components (Fig. 1): (1) a water filled PVC “stationing” hoop in which the manatee positioned its head toward the transducer; (2) a USRD H56 hydrophone, used to measure stimulus and background noise levels inside the hoop, around the subject’s head; (3) a USRD I-9 or I-13 projector, positioned 1.5 m from the stationing hoop at a depth of 1.5 m; (4) a structural platform above the water’s surface from which the transducer and paddles were suspended into the water; and (5) two submerged PVC response paddles for the animals to push to indicate the presence or absence of a test signal (each paddle was an equal distance from the stationing hoop). The underwater apparatus was positioned to direct the axis of sound propagation in the long direction of the pool. This, along with the irregular nonparallel sides of the pool, helped to minimize reflections. The stationing hoop positioned the subject’s head in-line with the transducer with the hydrophone

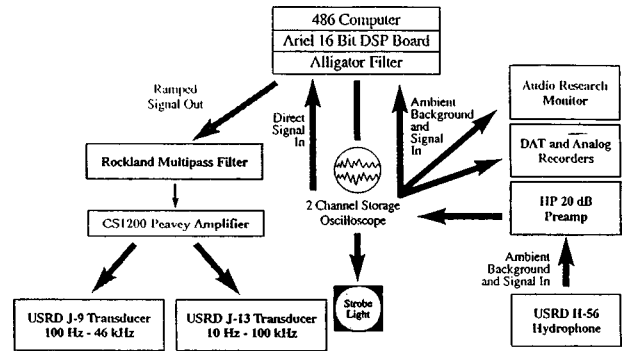


FIG. 2. Signal generation system used in all experiments.

approximately 20 cm from the manatee’s meatus. With respect to accurately measuring the sound pressure levels received at the subject’s ears, the low frequency portion of the test (below 1 kHz where the wavelength is greater than 1.5 m) presented no challenge. In this frequency region the sound pressure level does not vary rapidly laterally so that measurements made within 20 cm of the manatee head are accurate measures of the sound at the manatee’s ear. However, because the sound pressure gradient field near a low frequency source is enhanced over that for free field, animal sensations close to a source may be felt rather than heard. This phenomenon is discussed as a possible interpretation for improved responses by one subject at frequencies below 0.4 kHz. Regarding the hydrophone position, at higher frequencies the sound from the J9 becomes directional (e.g., at 6 kHz the half-power is 100 degrees) so that no significant reflection from tank boundaries arrives at the hydrophone and the sound at the hydrophone is taken to be the sound heard by the manatee. Real-time spectral monitoring helped to further insure the integrity of the localized acoustic measurements about the subject’s head.

### D. Signals and signal generation

The signal generation and recording equipment (Fig. 2) was controlled by a 486 DX computer, configured with an Ariel DSP card and an Alligator Technologies AAF-3, two channel low-pass filter board to generate and automate the delivery and recording of acoustic tones and ambient conditions. In addition to low-pass filtering, an in-line external bandpass filter (Rockland model 852) shaped outgoing signals. Filtered signals were sent through a Peavey CS1200 power amplifier and projected to the manatees through a Navy USRD J9 transducer (for signals from 0.1 to 46 kHz), and a USRD J13 (for signals between 0.015 and 0.1 kHz). Pure tone 500-ms sinusoidal pulses having a 100-ms rise time, a 300-ms steady state peak, and a 100-ms fall time were repeated twice per second for 4 s resulting in a total of eight pulses per trial. The slow rise/fall times helped to insure a narrow bandwidth of transients, guarding against speaker artifacts or “popping.” A USRD H56 hydrophone recorded ambient and test sound pressure levels at the manatee’s stationing point during each trial. The signals and noise were monitored with a Tektronix T912 dual channel storage

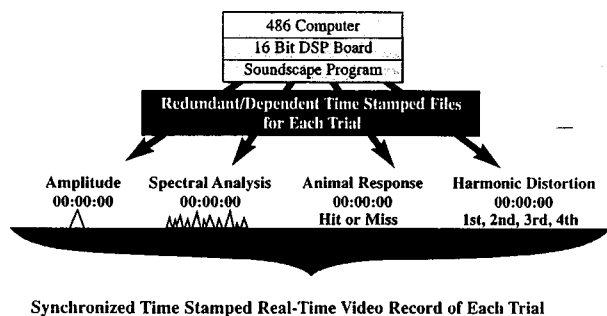


FIG. 3. Flow diagram of data collection and documentation.

oscilloscope, and were simultaneously recorded by the computer and with time a coded Nagra IV or Sony DAT10 recorder.

### E. Data collection

A dedicated audiogram program (“Soundscape”), developed for this study, was used to document behavioral and acoustic events synchronously (Fig. 3). This program provided a real-time spectral and harmonic analysis for each trial. Monitoring for harmonic distortion was especially pertinent when measuring lower frequencies at higher amplitudes to ensure that the subjects did not respond to a harmonic of the test tone’s fundamental frequency. All manatee responses were entered into the computer where along with the corresponding sound pressure levels and spectral information they were automatically time stamped. Test sessions were recorded using a Sony Hi8 video camera (CCD-TR81) time synchronized with the computer. The videotapes were indexed by date, test session number, corresponding counter number, and behavioral comments.

### F. Training procedures

Operant conditioning using positive reinforcement was used to train the manatees to discriminate between the tone and no-tone paddles and to learn the procedural chain of behaviors. Using an unambiguous two paddle presentation with fixed locations, distinct light and dark pattern differences, and tactual surfaces facilitated a shorter discrimination training period (Gerstein, 1994). The naïve manatees required six months to learn the necessary paddle discriminations and procedural behaviors to perform the threshold tests.

To convey a sense of the pace and complexity of the chain behaviors required, listed are the steps executed during each test trial for a single reward opportunity: (1) swim to the experimenter at the surface of the water; (2) station and wait for a hand signal to start the trial; (3) upon receiving the hand signal turn, dive and swim to the stationing hoop, 6 m away and 1.5 m below the water’s surface; (4) place head in hoop facing the transducer; (5) station in place, listen, and wait for the strobe light; (6) after strobe flash, back out of the hoop and select one of two paddles (tone or no-tone); and swim back to the experimenter at the surface for a reward and/or the start of the next trial. Each trial, from start to finish averaged 2 min, not including time spent rewarding the subject between trials. Sessions consisted of 30–80 trials

lasting 1–2 h. Both animals had to be conditioned physically as well as behaviorally to endure and accept the long test sessions.

### G. Testing procedure

A forced-choice two alternative paradigm was used. This test required an equal demonstrative action by the manatees to select either the no-tone or the tone paddle, which helped minimized selection bias. The experiment used a double blind presentation of randomized tone and no-tone trials. Computer-generated lists of modified Gellerman series (Gellerman, 1933) determined the ordered sequence of on-off trials. The experimenter working with and reinforcing the subject at the surface was unaware of the trial sequence and communicated with the computer operator in the underwater lab via radio headsets. During testing, the paddles were unmanned to insure against inadvertent cueing. Once the subject was stationed inside the stationing hoop, the computer initiated a randomized tone or no-tone trial sequence. A strobe light pulsed 2.5 s after the onset of the sequence. The strobe light signaled the animal to leave the stationing hoop and select a paddle. Since the manatees required 2–5 s to back out of the hoop, the delay provided adequate time to hear a signal, while not penalizing or disqualifying an over anxious animal which might try to leave prematurely in response to the first detectable pulse. After backing out of the hoop each subject required approximately 5 s to reach either paddle. When a paddle was pushed with sufficient force, the subject’s choice was recorded on the computer.

To safeguard against motivational artifacts, a system of five “warm-up” and five “cool-down” trials was a useful check on the animals’ reliability. During these trials, the tonal amplitudes were approximately 15 dB above the animal’s estimated threshold for each frequency. These indicator trials revealed the subjects’ motivational state at the beginning and end of long sessions. If the subject performed with 80% or greater accuracy, then he passed this motivational check and the preceding data were considered accurate. If during warm-up trials the subject’s accuracy was <80% accurate in paddle selections, then the session was used for training only. If the subject scored <80% during cool-off trials, then the session was discarded. Other audiogram studies have used similar motivational checks and performance criteria (Aubrey *et al.*, 1988; Hall and Johnson, 1972; Johnson, 1967; Schusterman *et al.*, 1972; Schusterman and Moore, 1978; Thomas *et al.*, 1988, 1990).

Once a session began, and the animal passed the criterion on the warm-up trials, each successive signal was attenuated in 3-dB steps until the first incorrect response (a “miss”) occurred. Following a miss, the sound pressure levels were increased in 1-dB steps until the subject responded correctly (a “hit”). Sound pressure levels were subsequently presented in  $\pm 1$ -dB steps. This up/down staircase method (Robinson and Watkins, 1973) yielded a series of reversals. Sessions ranged from 30 to 80 trials until the subject completed at least 5 reversals. A minimum of 12 reversals with <3-dB differences measured from two or more consecutive sessions were averaged to calculate the threshold for each test frequency. Only one threshold session per subject was



TABLE I. Pure tone underwater thresholds for Stormy.

Frequency (kHz)	Mean threshold (dB <i>re</i> : 1 $\mu$ Pa)	Ambient noise (1-Hz band)	Number of reversals	Std. deviation (dB <i>re</i> : 1 $\mu$ Pa)	Std. error (dB <i>re</i> : 1 $\mu$ Pa)	Percent correct control trials	Gellerman series of each test session
0.015	111	68	27	1.46	0.28	85	417,418,420a,420b
0.05	98	68	31	2.62	0.47	88	409,413,415,416
0.1	93	43	33	2.25	0.39	93	303,304,402,404
0.2	93	36	12	1.53	0.44	91	402,404
0.4	102	43	29	1.84	0.34	89	300,301,302,303
0.5	102	43	32	2.20	0.39	90	558a,558b,559a,559b
0.8	82	41	27	1.84	0.35	94	233,234,236a,236b
1.6	72	25	53	2.55	0.37	95	18,191,196,198,526,582a,582b
3	67	25	28	1.97	0.37	92	44,4238,603,604,605
6	58	26	28	1.98	0.45	98	286,288a,288b,289
10	56	26	32	2.52	0.45	98	441a,441b,444,445
12	52	27	31	1.60	0.29	98	291,292,293,294
16	50	28	34	3.25	0.56	98	437,438a,438b,439a,439b
18	50	25	33	3.01	0.52	100	447,449a,449b,450
20	58	26	30	1.68	0.31	95	295a,295b,297,312
26	66	31	13	1.90	0.53	95	309,310,311
32	77	31	30	2.83	0.52	95	315,317,320,325,328
38	88	32	31	3.29	0.59	90	425,426,433a,433b
46	112	33	16	1.94	0.49	85	434a,434b,435

completed per day to avoid behavioral artifacts associated with successive, repeated measurements (Johnson, 1967).

Since the manatees' hearing range and sensitivity were previously unknown, the first measurements focused on one subject, Stormy. Thresholds for 19 test frequencies over a range of 0.15–46 kHz were obtained. After Stormy's audiogram was completed, select frequencies along significant points of the audiogram were later measured with Dundee. A total of 8967 test trials were conducted using both subjects; of these, 1005 were discarded due to session interruptions from sudden thunderstorms, zoo emergencies, or failure of the animals to reach behavioral criteria during warm up or cool down trials. The remaining 7962 trials were used to estimate the hearing thresholds of both subjects.

## II. RESULTS

The thresholds from both subjects were averaged from a total of 824 reversals (Tables I and II). The variability between a hit and miss within each frequency was less than 3 dB for each reversal. The average standard deviation about each mean threshold is less than 3 dB with a standard error of less than 1 dB. Both subjects averaged greater than 90%

selection accuracy during control trials (warm-up, cool-down, and no-tone trials).

The resulting audiograms for both subjects are U-shaped curves (Fig. 4). Dundee's hearing thresholds are plotted from 0.5 to 38 kHz and Stormy's from 0.4 to 46 kHz. Both subjects demonstrate a maximum hearing sensitivity of 50 dB *re*: 1  $\mu$ Pa at 16 kHz. Stormy also exhibits the same sensitivity at 18 kHz. The frequency range of best hearing for both subjects is 6–20 kHz (defined as 9 dB within maximum sensitivity). Sensitivity decreased approximately 40 dB per octave above 26 kHz and 20 dB from 0.8 to 0.4 kHz. At 0.4 kHz, thresholds exceed 100 dB *re*: 1  $\mu$ Pa (60 dB over ambient).

After three months of probing lower frequencies, Stormy detected tones lower than 0.4 kHz at thresholds lower than 100 dB *re*: 1  $\mu$  Pa. These detection thresholds may be the result of vibrotactile sensations caused by acoustic particle velocities which are larger (particularly for lower frequencies) in a spherically spreading sound field than in a planar field. A conservative interpretation of the maximum hearing range is 0.4–46 kHz, approximately 6.5 octaves. A more liberal interpretation would extend from 0.015 to 46 kHz, approximately 11 octaves.

TABLE II. Pure tone underwater thresholds for Dundee.

Frequency (kHz)	Mean threshold (dB <i>re</i> : 1 $\mu$ Pa)	Ambient noise (1-Hz band)	Number of reversals	Std. deviation (dB <i>re</i> : 1 $\mu$ Pa)	Std. error (dB <i>re</i> : 1 $\mu$ Pa)	Percent correct control trials	Gellerman series of each test session
0.5	101	43	40	3.27	0.52	90	559a,559b,3828a,3828b
1.6	76	25	30	4.70	0.75	96	555a,555b,103a,103b
3	67	25	26	2.23	0.44	97	552a,552b,2575a,2575b
6	63	26	22	1.96	0.42	100	550,551,4136
12	55	29	47	3.05	0.44	100	541a,541b,542,2202a,2202b
18	53	25	36	2.70	0.45	100	543a,543b,4887a,4887b
26	68	31	34	2.35	0.40	100	545,547a,547b,4668
38	94	31	40	3.28	0.52	87	101,548,549,2547

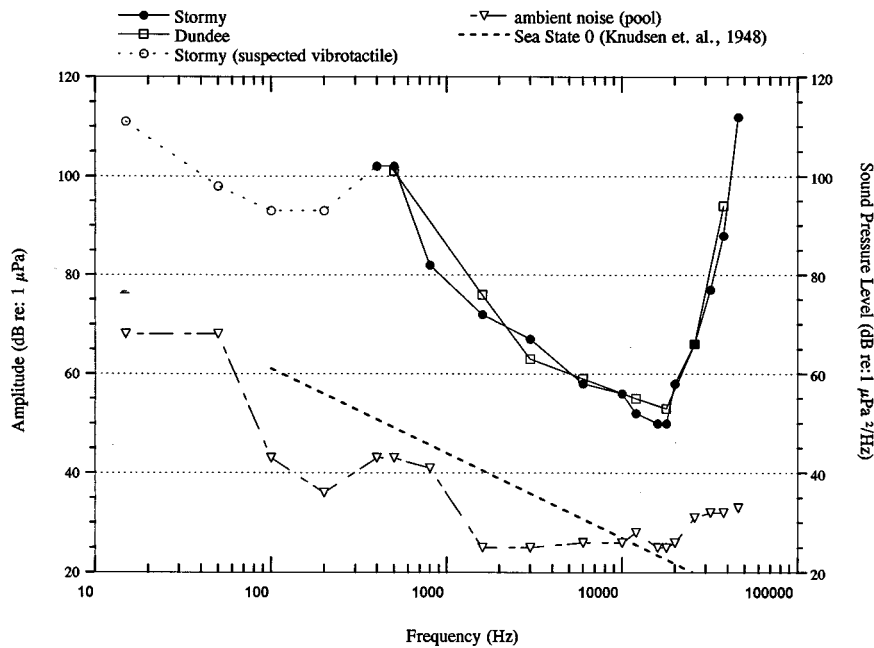


FIG. 4. Underwater audiograms. The audiograms are plotted for Dundee (□) and Stormy (●). Stormy's suspected "vibrotactile" detections below 0.4 kHz are indicated (○, dotted line). The accompanying ambient noise in the pool is plotted in 1-Hz bands and referenced with Sea State 0.

### III. DISCUSSION

#### A. Absolute hearing limits

The audiograms and corresponding ambient levels are presented with Sea State 0 for reference. The ambient noise was below Sea State 0 and at least 22 dB lower than any threshold, suggesting little probability of masking at the frequencies presented. The audiograms are good estimates of absolute hearing.

#### B. Anatomical and evoked potential studies

The behavioral audiograms demonstrate a higher range of hearing and greater sensitivity than was predicted from the anatomical investigations conducted by Ketten *et al.* (1992). Evaluations of cochlea and middle ear anatomy led to conclusions that manatees have a low frequency, "nonacute"

ear, with peak frequency sensitivity near 5 kHz and poor sensitivity throughout their hearing range. The behavioral audiogram provides no support for the anatomical based assessments; in fact, manatees have peak frequency sensitivity at 16–18 kHz with a significant gradient of sensitivity throughout their range of hearing and extended ultrasonic hearing up to 46 kHz. Furthermore, comparison with other published marine mammal underwater audiograms suggests that the manatees have greater peak frequency sensitivity than most pinnipeds (Fig. 5) and comparable peak sensitivity with some odontocetes (Fig. 6).

While the manatees' high frequency sensitivity extends beyond the estimates of Ketten *et al.* (1992), ultrasonic sensitivity is not entirely surprising in light of some anatomical, behavioral, and physiological observations. The manatees' upper frequency limit defined by the behavioral audiogram is

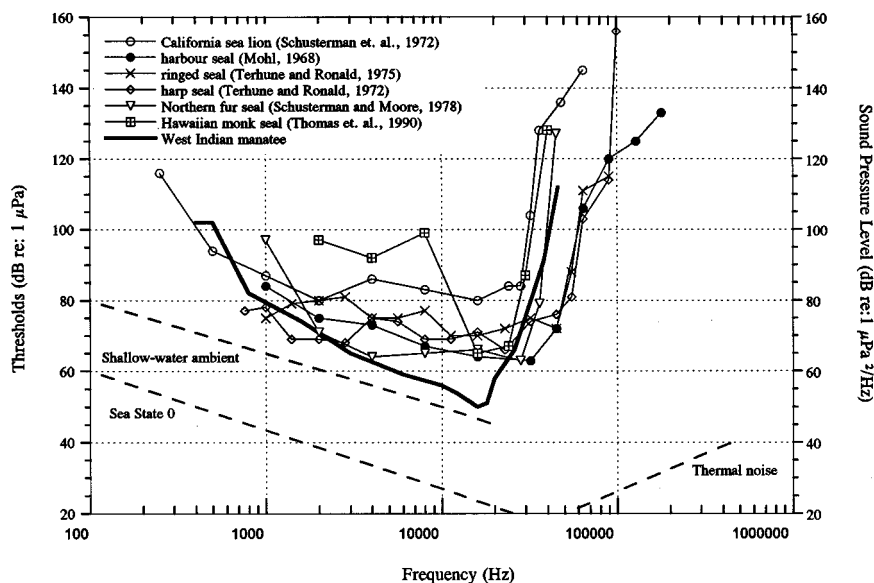


FIG. 5. Comparison of manatee and pinniped underwater audiograms. The manatee audiogram drawn from Stormy and Dundee is presented with published pinniped audiograms. The shallow water ambient curve is a compilation of measurements from varied coastal areas and harbors presented in Urlick (1983). Sea State 0 and thermal noise curves are referenced to illustrate low noise levels in shallow and deep water environments.

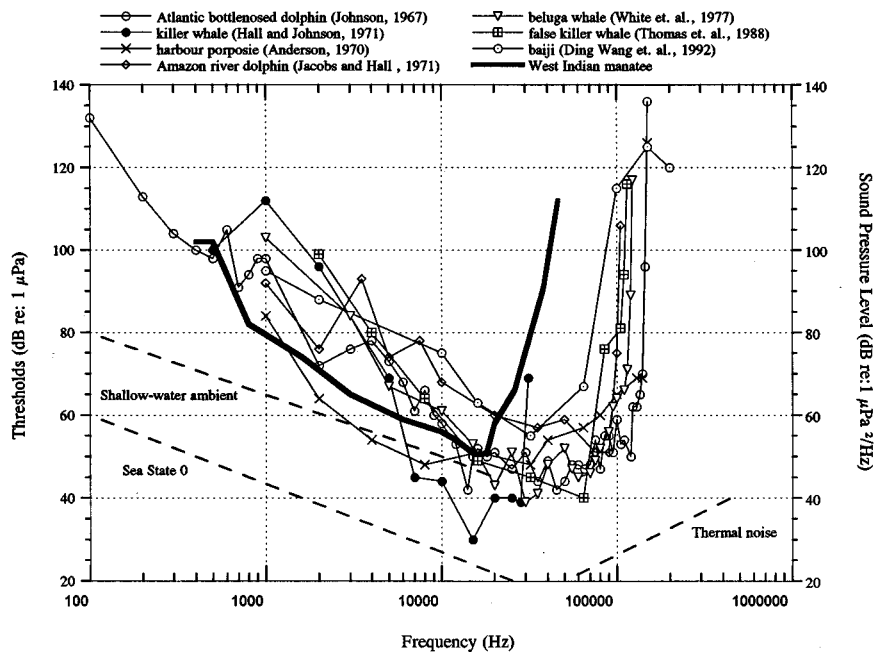


FIG. 6. Comparison of manatee and odontocete underwater audiograms. The manatee audiogram is averaged from Stormy and Dundee and presented with published pinniped audiograms. The shallow water ambient curve is a compilation of measurements from varied coastal areas and harbors presented in Urlick (1983). Sea State 0 and thermal noise curves are referenced to illustrate low noise levels in shallow and deep water environments.

consistent with the Heffner and Masterton (1980) regression which successfully plots the behavioral upper frequency limits of an array of mammalian species. Both manatee subjects have the same size heads with intermeatal distances of 390 mm. Using the intermeatal distance (divided by the speed of sound in water) the functional interaural distance or estimated  $\Delta t$  for both subjects is 260  $\mu$ s. The Heffner and Masterton line predicts the manatees' upper hearing limit would approximate 46 kHz. Functional ultrasonic hearing is also suggested by the spectral content within some manatee vocalizations. Amazonian manatees have calls with harmonics as high as 32 kHz (Bullock *et al.*, 1980), and more recent recordings of West Indian manatee vocalizations reveal significant acoustic energy above 22 kHz (Gerstein, 1995). Bullock *et al.* (1982) also provided direct physiological evidence for ultrasonic sensitivity in West Indian manatees. They measured the auditory evoked potentials (AEP) of four West Indian manatees from 0.4 to 35 kHz. Bullock noted that the headphones used for the study were frequency limited and that sensitivity could have been higher than 35 kHz. While the AEP data and behavioral audiogram are consistent with regard to hearing at 35 kHz, Bullock *et al.* (1982) also noted the most effective frequency was only 1–1.5 kHz in the West Indian manatee and 3 kHz in the Amazonian manatee (Bullock *et al.*, 1980). These frequencies are significantly lower than 16- and 18-kHz peak sensitivity measured in the behavioral audiograms. However, since the evoked potential waveforms were not the same across frequencies, relative sensitivity estimates using AEPs are not reliable. The authors conceded that their AEP data had greater utility as a gross indicator of conspicuous hearing limits than use as a scalar of relative frequency sensitivity. Bullock *et al.* (1982) note that their "most notable finding" was ultrasonic detection "clearly to 35 kHz" and possibly higher.

Following the AEP work, Klishin *et al.* (1990) recorded auditory brain-stem responses (ABR) from an Amazonian manatee (*Trichechus inunguis*) up to 60 kHz. Klishin *et al.*

(1990) unlike Bullock *et al.* (1980, 1982) did not use contact headphones and projected auditory stimuli underwater in a small water filled bath where the manatee was suspended. An ABR audiogram was obtained from 5 to 40 kHz. While the general shape of the ABR audiogram and the range of best hearing from 5 to 20 kHz are consistent with the West Indian manatee behavioral audiogram, the overall sensitivity is significantly less. The Amazonian manatee exhibited weak ABRs at significantly higher signal thresholds than odontocetes and pinnipeds (Forbes and Smock, 1981; Popov and Supin, 1990; Ridgway *et al.*, 1981; Watkins and Wartzok, 1985). Although ABRs are weaker than those obtained with odontocetes and pinnipeds, varied test conditions make relative magnitude based comparisons ambiguous especially between different taxa. Attenuation of electrical fields through the manatees' denser skull, as well as differences in electrode placement with respect to the manatees' peculiar brain organization (Reep and O'Shea, 1989), may significantly effect ABR amplitudes. However, such comparisons among individuals of the same or closely related species may be useful, especially when consistent measurements are recorded. Despite methodological differences between the Bullock and Klishin investigations, both far-field and cortical electrical potentials were conspicuous at ultrasonic frequencies (up to a minimum of 35 kHz) in both West Indian and Amazonian manatee subjects.

### C. Interspecies comparisons

A comparison of the manatee behavioral audiogram with other marine mammals shows a shared increase in sensitivity with pinnipeds and odontocetes to frequencies between 10 and 26 kHz. This shared sensitivity among marine mammals could be a convergent adaptation to exploit low noise conditions underwater. Although active echolocating odontocetes hear frequencies >46 kHz, they share a common sensitivity

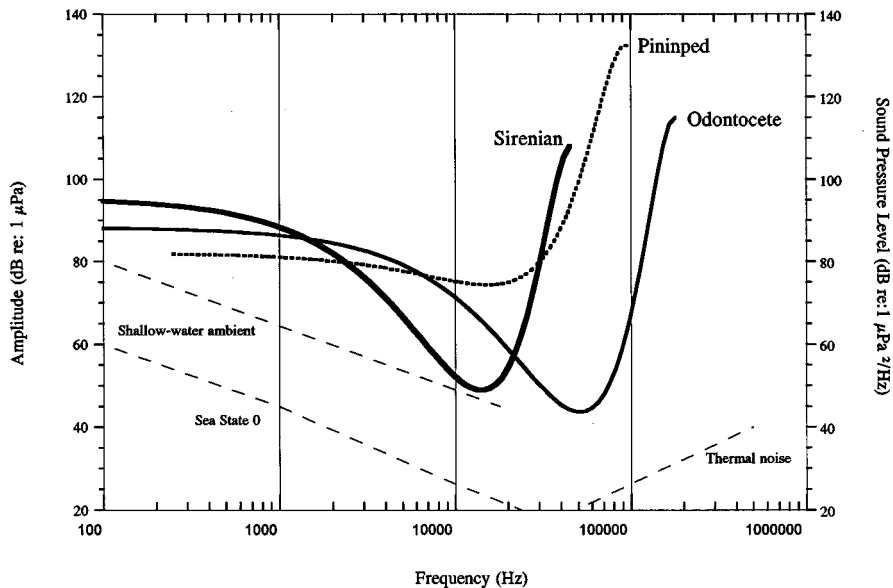


FIG. 7. Best-fit curves of sirenian, pinniped and odontocete audiograms. Third order polynomial curves are drawn from the audiograms referenced in Figs. 5 and 6. These curves are accompanied by shallow water and noise curves from Urick (1983).

with nonecholocating marine mammals. The killer whale (*Orcinus orca*) and manatee audiograms demonstrate respective peak hearing sensitivity near 16 kHz. When looking at Fay's (1988) compilation on hearing in vertebrates, one is struck by the similarities rather than the differences in marine mammal hearing. Therefore, it is not unreasonable to expect manatee hearing to be similar to other marine mammals. In Fig. 7 we have drawn third order polynomial curves for published underwater audiograms of cetaceans and pinnipeds. These arbitrary curves are meant to illustrate a general trend in sensitivity, are not drawn to represent absolute values, and should be viewed as a gross best fit estimate of the hearing curves for these marine mammals. The pinnipeds are amphibious marine mammals and, although they have extended high frequency hearing, their peak sensitivity is not as acute as the obligate marine mammals. Manatee hearing lies between that of the pinnipeds and fully aquatic echolocating odontocetes. Convergence in sensitivity may be a function of similar background levels underwater. An additional phenomenon that might have shaped the hearing of coastal marine mammals is the *Lloyd Mirror Effect*. In moderate to calm seas and shallow water areas, this underwater sound interference pattern can significantly attenuate propagation of sound near the surface. Destructive and constructive waves generated near and reflected off the surface can effectively cancel low frequencies. The *Lloyd Mirror Effect* predicts the sound pressure level at the surface approximates (0) zero, because the ocean's surface is a pressure-release boundary that is free to move in response to pressure in the water. The increase of pressure away from the ocean's surface is proportional to frequency, with pressure at shallow depths being inversely proportional to wavelength and thus proportional to frequency (the lower the frequency the smaller the acoustic pressure near the surface). The details of these fluctuations at short distances depend on many factors, most important of which are ocean depth, bottom shape and density, and surface roughness. Marine mammals repeatedly surface to breathe, and many rest and cavort near the surface

for extended periods in relatively shallow coastal zones. If near-surface underwater acoustic communication had selective advantages, the *Lloyd Mirror Effect* might also have pressured marine mammals, and particularly manatees, to evolve higher frequency hearing. Acoustic pressure-based communication, particularly low frequency vocalizations generated near the surface, would be ineffective over moderate distances, and would require higher frequency sounds in order to propagate greater distances near the surface. Although manatees are herbivorous, and their unique lifestyle exposed them to different selective pressures than other marine mammals, their passive hearing sensitivity at higher frequencies may reveal a convergent adaptation with other marine mammals to the shared challenges of a noisy environment.

#### D. Low frequency detection

While both manatees demonstrated moderate low frequency sensitivity, their hearing sensitivity fell approximately 20 dB per octave below 0.8 kHz, and at 0.4-kHz amplitudes had to be >100 dB re: 1 μPa. Because of the phylogenetic relationship which manatees have to elephants (Simpson, 1932) and speculation of shared low frequency hearing (Reynolds and Odell, 1991), we were particularly interested in infrasonic frequencies. Consequently, we spent several months probing both subjects' hearing below 0.4 kHz. Stormy was able to detect frequencies below 0.4 kHz only after months of repeated trials. The improved sensitivity at frequencies <0.4 kHz suggests that Stormy either switched detection strategies from hearing to feeling, or somehow became even better at listening. Stormy was retested at 1.6, 0.8, 0.5, and 0.4 kHz to see if he would demonstrate increased sensitivity for these frequencies as well. However, his threshold at these frequencies did not change, and indicated that his performance at 0.2, 0.1, 0.05, and 0.015 kHz was not an artifact of improved task proficiency. It is possible that he switched from hearing to feeling. Al-

though both animals were motivated to receive their much prized monkey chow rewards, Dundee—who was trained to listen—never made the cognitive leap to switch detection modes.

Reep *et al.* (1989) recently discovered cortical specializations unique to Sirenia, which are believed to be associated with perioral tactile function and sensation. In light of this proposed cortical specialization, and the manatees' readily observed exploratory tactile behavior, neural commitment for tactile sensory input and processing is probably highly developed. Manatees also have short hairs uniformly spaced over their body, and while no sensory function has been demonstrated, mechanoreception possibilities have been suggested (Reynolds, 1979). If these hairs can detect particle displacement near the projector where acoustic particle velocity is larger than in a plane wave field, it could be one possible explanation for Stormy's behavioral response and improved performance below 0.4 kHz. The video log revealed that Stormy consistently rotated his body on axis ~45 degrees to the left, and bent his head down while listening in the hoop during these low frequency sessions. This posturing was not exhibited for frequencies >0.4 kHz. Since sessions began with warm-up trials, which were approximately 15 dB above suspected thresholds, the subjects were sensitized at the start of each session as to what frequency they would need to listen for. Therefore, if Stormy had indeed learned to switch detection strategies, he did so selectively at frequencies below his hearing, and did not exhibit this turning posture during a session testing tones >0.4 kHz. Perhaps Stormy was able to utilize a vibrotactile response, which is stimulated by acoustic particle velocity.

Manatees inhabit shallow water where particle motion detection is more useful near the waters' surface, where sound pressures are low due to the *Lloyd Mirror Effect*. In the near field, vibrotactile receptors may be stimulated, causing one to assume that sound receptors have been stimulated. A good method of determining if vibrotactile effects are present in audiograms is to measure at two different transducer distances, as particle velocity falls off more rapidly than sound pressure in diverging acoustic fields and consistent results may not be obtained. For this study, a J13 transducer provided a technical solution to projecting in-water low frequency and infrasonic signals in a shallow (3-m) pool as opposed to generating in-air signals and projecting them down through the water (e.g., Aubrey *et al.*, 1988). The tolerances of the J13, combined with the distance the subjects were from the source and the distance from the surface, helped avoid harmonic and reflective distortions, but at the same time made increased distance measurements impractical.

Aside from vibrotactile responses, other possible explanations for Stormy's increased detection could be related to the selective low frequency resonance of air-filled sinuses in his head, body cavity, or lungs, or perhaps, enhanced bone conduction (via the manatees' spongy zygomatic process), as was first suggested by Bullock *et al.* (1982) and later Ketten *et al.* (1992).

It is of coincidental interest that Stormy's improved detection was best at 0.1 and 0.2 kHz, the range for lateral line

particle displacement detection in fish (Fay, 1988; Tavalga and Wodinsky, 1963; Cahn *et al.*, 1969). Although shifts in audiograms from hearing to feeling are common in fish, few observations have been reported in marine mammals; Turl (1993) with a bottlenose dolphin and most recently by Kastak and Schusterman (1998) in two sea lions.

## E. Ecological consequence

Manatees may be well adapted for listening in the sea and, in particular, shallow water environments, but they appear poorly suited to detect the dominant low frequency sounds produced by boats near the surface. Given the manatees' limited low frequency hearing sensitivity, it is likely that manatees have difficulty detecting, as well as locating approaching boats from safe distances. At only 0.5–2 m from the waters' surface (a zone at which manatees are vulnerable to collisions with boats), low frequencies generated near the surface are significantly attenuated and can drop below or become indistinguishable from background levels. The complications of the *Lloyd Mirror Effect* for manatee hearing and awareness of sound are serious: The manatee's ability to hear low frequency sounds from a boat, especially frequencies generated by the blade rate of a propeller, is at its worst when the manatee is at the surface and thus most vulnerable. Exploitation of the manatees' optimum hearing sensitivity with a high frequency directional acoustic beacon would provide both directional and distance cues to the manatee and therefore might help reduce mortalities associated with boats. Additionally, since both manatees and other marine mammals (including large whales) must cope with the *Lloyd Mirror Effect*, it is possible that ship collisions with large whales at sea are related to the attenuation of the very low frequency propeller noise generated near the surface by commercial and military ships. Similar exploitation of a high frequency beacon could provide valuable acoustic spatial and distance cues to other marine mammals, since they are exposed to the same phenomenon and related ecological dangers as manatees near the water's surface in shallow coastal zones.

## IV. CONCLUSIONS

The West Indian manatee exhibits the typical mammalian U-shaped audiogram. Frequency sensitivity increased from 0.4 kHz to a trough of maximum sensitivity between 6 and 20 kHz (9 dB from maximum sensitivity). From these data a conservative estimate of the maximum hearing range for the West Indian manatee is from 0.4 to 46 kHz—the points where hearing sensitivity is 60 dB down from peak sensitivity. Hearing sensitivity dropped approximately 40 dB per octave above 26 kHz and approximately 20 dB per octave below 0.8 kHz.

There remain many endangered animals for which behavioral audiograms do not exist but for which management decisions need to be or are being considered. As this study demonstrates, assessments and predictions of functional hearing based upon morphometric data should be viewed cautiously, as the sound reception pathways and physical impedance characteristics of anatomical structures, as well as

neural processes, have not been identified or adequately measured in the manatee or the large whales. Results of this study remain our best estimate of what manatees hear. Forthcoming measurement of the manatees' critical ratios, critical bands, and directional sensitivity will provide further understanding of the sensory abilities of this unique marine mammal, and perhaps offer a means by which to help protect them in the wild.

## ACKNOWLEDGMENTS

This work is dedicated to Stormy and Dundee two very patient and trusting ambassadors of "*manateekind*." We wish to thank Dena Dickerson, Coastal Ecology Group of the Army Corps of Engineers Waterways Experiment Station, in Vicksburg, Mississippi, and Art Wilde of the Florida Inland Navigation District in Jupiter, Florida, for their support; Ingrid Johanson, Betty Tuller, and Michael Salmon of Florida Atlantic University for their assistance and review of the manuscript. Thanks to Lex Salisbury and the Lowry Park Zoo for hosting this study at their manatee facility in Tampa, Florida. Thanks to Richard Pierce and Bill Tavalga at the Mote Marine Lab and Ted Bullock at the Scripps Institute for their encouragement and assistance with permitting. Thanks to on-site volunteers Anthony Finney, Becky Dunbar, and Janel Troide. Finally, a special acknowledgment to Geoffrey Patton of the University of South Florida for his inspiration and significant efforts which were instrumental in initiating this study. This study was funded in part by the U.S. Army Corps of Engineers, Waterways Experiment Station in Vicksburg, Mississippi, the Florida Inland Navigation District, Jupiter, Florida, and the Florida Department of Environmental Protection. Manatee training and testing activities were conducted under the authorization of the U.S. Fish and Wildlife Service, Office of Management Authority, Endangered Species Permit PRT-761873.

Aubrey, F., Thomas, J., and Kastelein, J. R. (1988). "Low-frequency underwater hearing sensitivity in belugas, *Delphinapterus leucas*," J. Acoust. Soc. Am. **84**, 2273–2275.

Bullock, T. H., Domning, D. P., and Best, R. (1980). "Evoked brain potential demonstrate hearing in a manatee (*Sirenia: Trichechus inunguis*)," J. Mammal. **61**, 130–133.

Bullock, T. H., O'Shea, T. J., and McClune, M. C. (1982). "Auditory evoked potentials in the West Indian Manatee (*Sirenia: Trichechus manatus*)," J. Comp. Physiol. **148**, 547–554.

Cahn, P. H., Silver, W., and Wodinsky, J. (1969). "Acoustico-lateralis system of fishes: tests of pressure and particle-velocity sensitivity in grunts (*Haemulon sciurus*) and (*Haemulon parrai*)," J. Acoust. Soc. Am. **46**, 1572–1578.

Fay, Richard R. (1988). "Hearing in mammals," in *Hearing in Vertebrates: A Psychophysics Databook* (Hill-Fay Associates, Winnetka, Ill), pp. 391–392, Figure M33/0, Table m33/0, pp. 393–394, Figure N34/0, Table N34/0.

Forbes, J. L., and Smock, C. C. (1981). "Sensory capacities of marine mammals," Psychol. Bull. **89**, 288–307.

Gellerman, L. W. (1933). "Chance orders of alternating stimuli in visual discrimination experiments," J. Gen. Psychol. **42**, 206–208.

Gerstein, E. R. (1994). "The manatee mind: discrimination training for sensory perception testing of West Indian Manatees (*Trichechus manatus*)," Marine Mammals: Public Display and Research **1**, 10–21.

Gerstein, E. R. (1995). "Underwater Audiogram of the West Indian manatee (*Trichechus manatus latirostris*)," Masters Thesis, Florida Atlantic University, Boca Raton, FL.

Hall, J. D., and Johnson, C. S. (1972). "Auditory thresholds of a killer whale (*Orcinus orca*) Linnaeus," J. Acoust. Soc. Am. **51**, 515–517.

Hartman, D. S. (1979). "Ecology and behavior of the manatee (*Trichechus manatus*) in Florida," Special publication no. 5, American Society of Mammalogists.

Heffner, R. S., and Masterton, R. B. (1980). "Hearing in Gliries: domestic rabbit, cotton rat, feral house mouse, and Kangaroo rat," J. Acoust. Soc. Am. **68**, 1584–1599.

Johnson, C. S. (1967). "Sound detection thresholds in marine mammals," in *Marine Bioacoustics II*, edited by W. N. Tavalga (Pergamon, Oxford), pp. 247–260.

Kastak, D., and Schusterman, R. J. (1998). "Low-frequency amphibious hearing in pinnipeds: Methods, measurements, noise and ecology," J. Acoust. Soc. Am. **103**, 2216–2228.

Ketten D. R., Odell, D., and Domning, D. (1992). "Structure, function and adaptation of the manatee ear," in *Marine Mammal Sensory Systems*, edited by J. A. Thomas, R. A. Kastelein, and A. Ya. Supin (Plenum, New York), pp. 77–95.

Klishin, V. O., Diaz, R. P., Popov, V. V., and Supin, A. Y. (1990). "Some characteristics of hearing of the Brazilian Manatee (*Trichechus inunguis*)," Aquatic. Mamm. **16**, 140–144.

Knudsen, V. O., Alford, R. S., and Emling, J. W. (1948). "Underwater ambient noise," J. Marine Research **7**, 410–429.

Popov, V. V., and Supin, A. Y. (1990). "Electrophysical studies on hearing in some cetaceans and a manatee," in *Sensory Abilities of Cetaceans: Laboratory and Field Evidence*, edited by J. A. Thomas and R. A. Kastelein (Plenum, New York), pp. 405–416.

Reep, R. L., Johnson, J. I., Switzer, R. C., and Welker, W. I. (1989). "Manatee cerebral cortex: Cytoarchitecture of the frontal region in (*Trichechus manatus latirostris*)," Brain, Behav. and Evol. **34**, 365–386.

Reep, R. L., and O'Shea, T. J. (1990). "Regional brain morphometry and lissencephaly in the sirenian," Brain, Behavior and Evolution **35**, 185–194.

Reynolds III, J. E. (1979). "The semisocial manatee," Natural History **88**, 44–53.

Reynolds III, J. E., and Odell, D. K. (1991). *Manatees and Dugongs* (Facts on File, New York), pp. 42–43.

Ridgway, S. H., Bullock, T. H., Carder, D. A., Seely, R. L., Woods, D., and Galambos, R. (1981). "Auditory brainstem response in dolphins," Proc. Natl. Acad. Sci. USA **78**, 1943–1947.

Robinson, D. E., and Watson, C. S. (1973). "Psychophysical methods in modern Psychoacoustics," in *Foundations of Modern Auditory Theory*, edited by J. V. Tobias (Academic, New York), Vol. 2, pp. 99–131.

Schusterman, R. J., Balleiet, R. F., and Nixon, J. (1972). "Underwater audiogram of the California sea lion by the conditioned vocalization technique," J. of Exp. An. Behav. **17**, 339–350.

Schusterman, R. J., and Moore, P. (1978). "Underwater audiogram of the northern fur seal (*Callorhinus urisnu*)," J. Acoust. Soc. Am. **64**, S87.

Simpson, G. G. (1932). "Fossil Sirenia of Florida and the evolution of the Sirenia," Bull. Am. Mus. Nat. Hist. **59**, 419–503.

Tavalga, W. N., and Wodinsky, J. (1963). "Auditory capacities in fishes: Pure tone thresholds in nine species of marine teleosts," Bull Am. Mus. Nat. Hist. **126**, 179–239.

Thomas, J., Chun, N., Au, W., and Pugh, K. (1988). "Underwater audiogram of a false killer whale (*Pseudorca crassidens*)," J. Acoust. Soc. Am. **84**, 936–940.

Thomas, J., Moore, P., Withrow, R., and Stoermer, M. (1990). "Underwater audiogram of a Hawaiian monk seal (*Monachus schauinslandi*)," J. Acoust. Soc. Am. **87**, 417–419.

Turl, C. W. (1993). "Low frequency sound detection by a bottlenose dolphin," J. Acoust. Soc. Am. **94**, 3006–3008.

Urick, R. J. (1983). *Principles of Underwater Sound* (McGraw-Hill, New York).

Wang, D., Wang, K., Xiao, Y., and Sheng, G. (1992). Auditory Sensitivity of a Chinese River Dolphin.

Watkins, W. A., and Wartzok, D. (1985). "Sensory biophysics of marine mammals," Marine Mammal Sci. **1**, 219–260.

White, M. J., Norris, J., Ljungblad, D., Barron, K., and Di Sciara, G. (1978). "Auditory thresholds of two beluga whales," Hubbs/Sea World Research Inst. Tech. Rept. No. 78–109.

# LETTERS TO THE EDITOR

This Letters section is for publishing (a) brief acoustical research or applied acoustical reports, (b) comments on articles or letters previously published in this Journal, and (c) a reply by the article author to criticism by the Letter author in (b). Extensive reports should be submitted as articles, not in a letter series. Letters are peer-reviewed on the same basis as articles, but usually require less review time before acceptance. Letters cannot exceed four printed pages (approximately 3000–4000 words) including figures, tables, references, and a required abstract of about 100 words.

## Characteristics of thermoacoustic resonance in a ducted burner

Sangyeon Cho and Soogab Lee<sup>a)</sup>

Department of Aerospace Engineering, Seoul National University, Seoul, Korea

(Received 4 February 1998; revised 12 September 1998; accepted 8 March 1999)

Combustion instability accompanied with low-frequency oscillations is one of the typical phenomena in a ducted flame burner. This is due to the interaction between unsteady heat release and the sound pressure field known as thermoacoustic feedback. A parametric study was carried out through measurements of acoustic pressure and flame radiation at various conditions not only to elucidate the effect of Reynolds number and equivalence ratio on thermoacoustic oscillation, but also to understand the mechanism of the onset of thermoacoustic resonance in a ducted flame burner. The results explain that equivalence ratio has significant effects on the magnitude of pressure and heat fluctuation as well as the fundamental frequency. It is also found that the onset of thermoacoustic resonance is strongly affected by the interacting mechanism between acoustic energy and heat release. © 1999 Acoustical Society of America. [S0001-4966(99)03306-8]

PACS numbers: 43.28.K+, 43.50.Nm [LS]

### INTRODUCTION

Thermoacoustic oscillation is one of the typical phenomena in a ducted flame burner, and is usually accompanied with severe vibration and noise. This low-frequency oscillation with high amplitude can result in structural damage or deterioration of the performance of the subsidiary system. Thus the understanding of physics in thermoacoustic oscillation has great importance in both combustor design and noise control.

Different from turbulent combustion noise, the spectrum of the thermoacoustic oscillation shows discrete frequency behavior. The noise is mainly caused by the interaction between the unstable heat release and the sound field—which is called thermoacoustic feedback. If the energy release by combustion and the pressure fluctuation are in phase, thermodynamic and acoustic energy are fed to each other and then the amplitude of oscillation increases remarkably. Rayleigh<sup>1</sup> explained that thermoacoustic resonance occurs if the Rayleigh index becomes positive. The Rayleigh index  $G(x)$  can be expressed as

$$G(x) = \frac{1}{T} \int_T q'(x,t)p'(x,t)dt, \quad (1)$$

where  $T$  is the instability period,  $q'$  is the unsteady heat release,  $p'$  is the pressure fluctuation,  $x$  is the axial location, and  $t$  is the time. Although many papers<sup>2–6</sup> concerning this

phenomenon have already been published, some fundamental questions regarding the effects of flow and combustion parameters on the onset of thermoacoustic resonance still remain to be investigated. For example, Heito *et al.*<sup>6</sup> found that the equivalence ratio has a crucial effect on the onset of oscillation, but no further explanation about the relationship between equivalence ratio and heat release or pressure fluctuation was given. Understanding the onset of thermoacoustic resonance requires detailed measurements of acoustic pressure and unsteady heat release with variations of physical parameters. Therefore, in the present study, a parametric study was carried out through measurements of acoustic pressure and flame radiation at various conditions, to elucidate the effect of physical parameters, such as Reynolds number and equivalence ratio, on the onset of thermoacoustic oscillation in a ducted premixed flame burner.

### I. BRIEF DESCRIPTION OF EXPERIMENT

A schematic sketch of the experimental setup to simulate thermoacoustic oscillation is shown in Fig. 1. It is a Rijke-type burner with a premixed flame across the entire cross section in the lower half (approximately at the quarter position) of the pipe having a diameter of 60 mm. The total length of the pipe was 1 m, and a short quartz pipe was installed at the location of the flame to observe flame oscillation and measure the light emission. Bundles of fine stainless steel pipes 2 mm in diameter were used for a flow straightener as well as flame holder. A decoupling chamber (60l) was installed for stabilizing airflow and for sufficient

<sup>a)</sup>Electronic mail: solee@plaza.snu.ac.kr

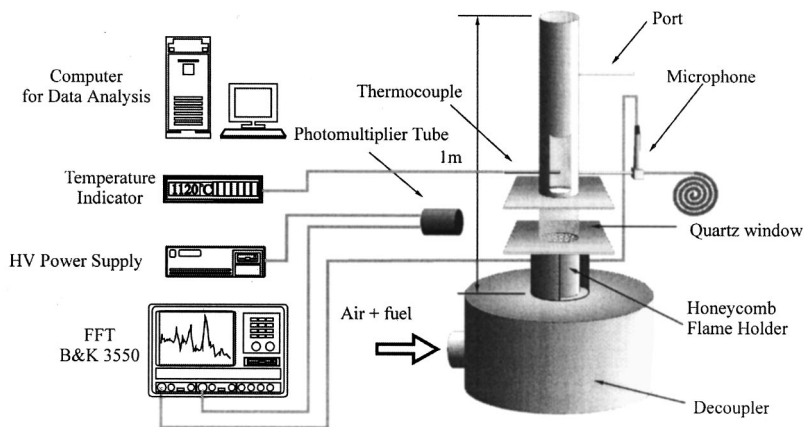


FIG. 1. Schematic sketch of the experimental setup.

premixing of the fuel (propane) and oxidizer. Sound pressure level was measured by a B&K 4134 condenser microphone located at the middle of the pipe. Since the temperature of the pipe was so high, an acoustic waveguide was utilized instead of flush mounting of the microphone to the wall of the tube. The acoustic signal measured by the sensor was transferred via a preamplifier to a B&K 3550 FFT analyzer. A thermocouple was also installed to measure the time-averaged (mean) temperature at the same time.

In the combustion region, there exist light emissions from free radicals. Among them,  $C_2$  light emission intensity from the flame can be interpreted as an extent of unsteady heat release.<sup>7,8</sup> To measure the  $C_2$  emission intensity only, a Hamamatsu R943-02 photomultiplier tube (PMT) was utilized with an optical filter, whose peak transmission wavelength is 516.4 nm.

## II. RESULTS AND DISCUSSION

Figure 2 shows the pressure spectra of acoustic oscillation as a variation of Reynolds number at a fixed equivalence ratio of  $\Phi=1.0$ . The measured fundamental frequency was far from the theoretical value ( $f_c = c/2L$ ) evaluated approximately with the geometry and a given temperature, which implies that this phenomenon is obviously nonlinear.

The largest magnitude of oscillation was obtained for the fundamental mode, and was over 20 dB stronger than the following higher harmonics. With a limited variation of Rey-

nolds number (from 3000 to 9000 due to the limitation of the burner capacity), the magnitudes of pressure peaks were not very sensitive to the flow rate. The broadband noise increased slightly by turbulence as Reynolds number increased, and as a consequence, total sound pressure level increased by 1–2 dB.

Figure 3 shows the variation of the mean temperature (time-averaged temperature measured by a thermocouple at the middle of the pipe) at the onset of strong thermoacoustic oscillation with respect to equivalence ratio ( $\Phi=0.8-1.6$ ). This figure clearly shows that at lean mixtures the oscillation starts at relatively higher mean temperatures than at rich mixtures. In the process of initiating a resonant oscillation, the mean temperature should be raised to a minimum (or threshold) value as shown in Fig. 3. To understand the detailed mechanism, the time traces of the pressure spectra during the process of starting a resonance were measured as shown in Fig. 4. For a lean mixture ( $\Phi=0.8$ ), resonance started at a relatively higher mean temperature (about 1000 °C at  $L/2$ ). With an onset of oscillation, the mean temperature suddenly dropped down to 750–800 °C, and then recovered again until a terminal mean temperature was attained. During this process, the fundamental frequency and its harmonics also suddenly shifted down and came back, as shown in Fig. 4(a). Before the thermoacoustic resonance initiated, there existed weak tonal noise (at least 30 dB weaker than in the resonance case). In this case, flame oscillations

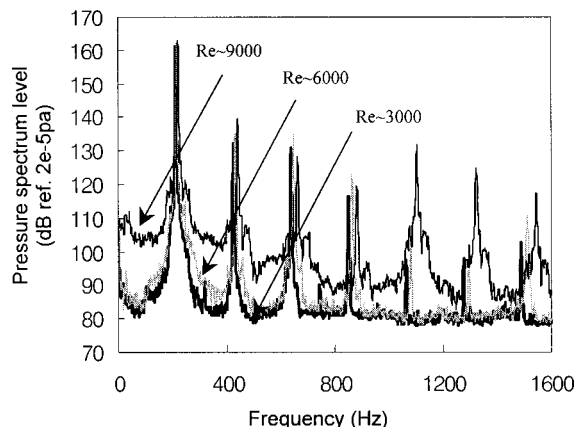


FIG. 2. Pressure spectra at the middle of the pipe ( $\phi=1.0$ ).

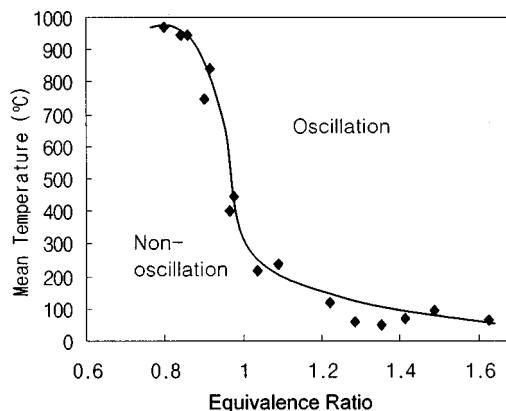


FIG. 3. Resonance-starting mean temperature at the middle of pipe with respect to equivalence ratio.



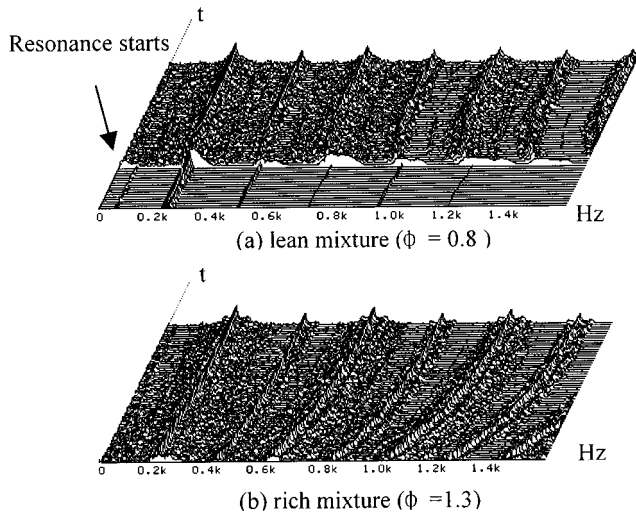


FIG. 4. Time histories of pressure spectra during resonance-starting process (time step: 0.5 s).

and PMT signals are rarely observed, therefore, it can be regarded as a transient state rather than a driving state.

As the equivalence ratio increased, the mean temperature required for the development of oscillations decreased rapidly. Since the resonance-starting temperature was relatively low for a rich mixture ( $\Phi=1.3$ ), no temperature drop was found in this range. As shown in Fig. 4(b), frequencies of pressure spectra shifted up monotonously with a temperature rise. The above results explain that the mechanism of an onset of thermoacoustic oscillation is strongly affected by equivalence ratio. It is also worthwhile to note that a critical requirement for the development of thermoacoustic oscillation is that the heat energy should be sufficient to overcome the energy loss from the system. At a lean mixture, the energy gain through the cycle is smaller (than at a rich mixture) since heat release is smaller. Therefore, the lean mixture case takes more time to reach the critical energy level to excite a resonant oscillation. Consequently, the oscillation-starting temperature of a lean mixture is higher than that of a rich mixture case. From Fig. 4, one can notice that, at a lean mixture, the system needs an additional energy shift from heat energy to excite resonance at an earlier stage of oscillation. This is the reason that the temperature (frequency) decreased at first, then increased again as the system gained energy. On the other hand, in the rich mixture case, the system did not need a shift from heat energy to mechanical energy for exciting a resonance since the system already had a sufficient amount of energy.

To verify this idea, the unsteady heat release was also measured since  $C_2$  emission intensity can be interpreted as an index of unsteady heat release. Results of microphone and photomultiplier measurements are shown in Fig. 5(a). It confirms that the two signals oscillate with the same fundamental frequency and its harmonics. Figure 5(b) shows the variations of unsteady heat release, pressure fluctuation, and Rayleigh index with respect to equivalence ratio. To calculate the Rayleigh index  $G(x)$  from the measurements, the integral of Eq. (1) has been expressed in the frequency domain:

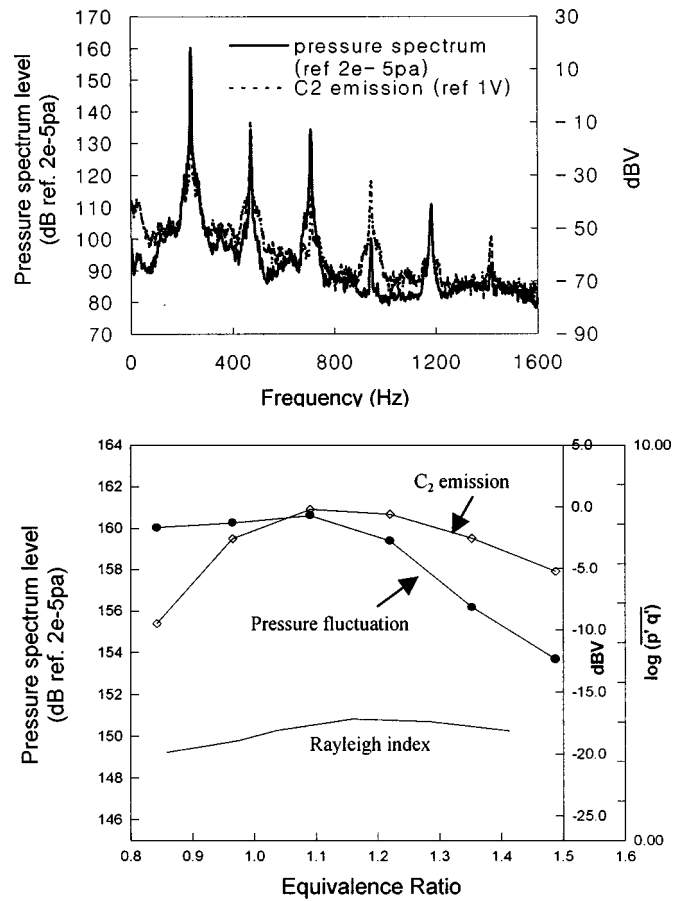


FIG. 5. (a) Spectra of pressure fluctuation and  $C_2$  emission rate ( $\phi=1.0$ ). (b) Variation of pressure fluctuation,  $C_2$  emission, and Rayleigh Index with respect to equivalence ratio.

$$G(x) = \int |S_{pq}| \cos \theta_{pq} dV. \quad (2)$$

In this equation,  $S_{pq}$  and  $\theta_{pq}$  are the cross spectrum and phase difference between the pressure and heat release, respectively, and the  $V$  is the volume of the flame.

Heat release was relatively small for a lean mixture, whereas the pressure fluctuation is large. It showed a reversed tendency for the rich side. In other words, pressure fluctuation is more dominant than heat fluctuation for a lean mixture, whereas heat fluctuation is more dominant for a rich mixture. This is a slightly surprising result because, in a lean mixture, a higher sound pressure was obtained in spite of smaller heat release than at a rich mixture. It could be explained as follows: the system energy gain is a product of pressure fluctuation and unsteady heat release. With a fixed condition, the total energy gain or Rayleigh index should be maintained at almost a constant value. Otherwise, the system gains more energy until it bursts (the index keeps increasing) or loses energy to be out of resonance (the index keeps decreasing).

### III. CONCLUSIONS

The results of this study are summarized as follows:

- (1) The Reynolds number is not a sensitive parameter for both the maximum pressure peak and the fundamental

frequency of thermoacoustic oscillation in a Rijke-type burner. On the other hand, equivalence ratio has a strong effect on the maximum pressure fluctuation and the fundamental frequency of thermoacoustic resonance.

- (2) To initiate a thermoacoustic resonance, a certain level of system energy should be required. This is why, for a lean mixture, oscillation starts from a higher mean temperature (than at a rich mixture) and then a sudden temperature drop exists in its initial phase of thermoacoustic resonance.
- (3) The product of acoustic pressure fluctuation and unsteady heat release (the Rayleigh index) should be maintained in the thermoacoustic resonance process.

## ACKNOWLEDGMENTS

This work was sponsored by Korea Science and Engineering Foundation under Contract No. 95-0200-07-03-3.

The authors want to express thanks to Professor I. S. Jeung and Professor C. S. Bae for their valuable comments.

<sup>1</sup>A. P. Dowling, and J. E. Ffowcs Williams, *Sound and Sources of Sound* (Ellis Horwood Limited, Chichester, 1983).

<sup>2</sup>G. J. Bloxidge, "Reheat buzz. An acoustically driven combustion instability," Ph.D. Thesis, Cambridge University (Cambridge, England, 1987).

<sup>3</sup>M. Fleifil, A. M. Annaswamy, Z. A. Ghoneim, and A. F. Ghoniem, "Response of a laminar premixed flame to flow oscillation. A kinematic model and thermoacoustic instability results," *Combust. Flame* **106**, 487–510 (1996).

<sup>4</sup>P. J. Langhorne, "Reheat buzz. An acoustically coupled combustion instability. Part 1. Experiment," *J. Fluid Mech.* **193**, 417–443 (1988).

<sup>5</sup>U. G. Hedge, D. Reuter, and B. T. Zinn, "Sound generation by ducted flames," *AIAA J.* **26**, 532–537 (1987).

<sup>6</sup>M. V. Heito, A. M. K. P. Taylor, and J. H. Whitelaw, "Influence of confinement on combustion instabilities of premixed flames stabilized on axisymmetric baffles," *Combust. Flame* **57**, 109–121 (1984).

<sup>7</sup>B. N. Shivashankara, W. C. Strahle, and J. C. Hardley, "Evaluation of combustion noise scaling laws by an optical technique," *AIAA J.* **13**, 623–627 (1975).

<sup>8</sup>I. R. Hurlle, R. B. Price, T. M. Sugden, and A. Thomas, "Sound emission from open turbulent premixed flame," *Proc. R. Soc. London, Ser. A* **303**, 409–427 (1968).