

THE JOURNAL of the Acoustical Society of America

Vol. 107, No. 5, Pt. 1

May 2000

EDITORIAL

Literate writing and collegial citing

Allan D. Pierce

2303

SOUNDINGS SECTION**ACOUSTICAL NEWS—USA**

2313

USA Meetings Calendar

2314

ACOUSTICAL STANDARDS NEWS

2317

Standards Meetings Calendar

2317

REVIEWS OF ACOUSTICAL PATENTS

2321

GENERAL LINEAR ACOUSTICS [20]

Synthetic array measurements of acoustical waves propagating into a water-saturated sandy bottom for a smoothed and a roughened interface

Harry J. Simpson, Brian H. Houston

2329

Acoustic scattering from two circular apertures in a thick hard plane

Jin S. Seo, Hyo J. Eom, Haeng S. Lee

2338

Scale model experiments on the insertion loss of wide and double barriers

Glenn J. Wadsworth, James P. Chambers

2344

“Fast” quasilongitudinal sagittally polarized surface waves in layer-substrate structures

A. N. Darinskii, I. S. Didenko, N. F. Naumenko

2351

Circular asymmetric Helmholtz resonators

A. Selamet, Z. L. Ji

2360

Transfer and Green functions based on modal analysis for Lamb waves generation

Ismael Núñez, Ros K. Ing, Carlos Negreira, Mathias Fink

2370

Pseudostatic corrections for the forced vibroacoustic response of a structure-cavity system

Michel Tournour, Noureddine Atalla

2379

Calculation of acoustic radiation using equivalent-sphere methods

Laurent Bouchet, Thierry Loyau, Nacer Hamzaoui, Claude Boisson

2387

NONLINEAR ACOUSTICS [25]

Mass, momentum, and energy transfer by the propagation of acoustic solitary waves

N. Sugimoto

2398

UNDERWATER SOUND [30]

Three-dimensional acoustic scattering from a penetrable layered cylindrical obstacle in a horizontally stratified ocean waveguide

Gerassimos A. Athanassoulis, Aristides M. Prospathopoulos

2406

Time reversal in a waveguide: Study of the temporal and spatial focusing

Philippe Roux, Mathias Fink

2418

(Continued)

CONTENTS—Continued from preceding page

Analysis of resolution for an amplitude-steered array	Catherine H. Frazier, W. Jack Hughes, William D. O'Brien, Jr.	2430
ULTRASONICS, QUANTUM ACOUSTICS, AND PHYSICAL EFFECTS OF SOUND [35]		
A model for longitudinal and shear wave propagation in viscoelastic media	Thomas L. Szabo, Junru Wu	2437
Weakly localized waves on the corrugated surface of a medium of arbitrary anisotropy	A. N. Darinskii	2447
The quasistatic approximation for a cracked interface between a layer and a substrate	Claudio Pecorari, Piaras A. Kelly	2454
Integration of modeling and acoustic microscopy measurements for thin films	Zhiqi Guo, J. D. Achenbach, Anita Madan, Keith Martin, M. E. Graham	2462
The use of partially measured source data in near-field acoustical holography based on the BEM	Sung-Chon Kang, Jeong-Guon Ih	2472
Performance measurements on a thermoacoustic refrigerator driven at high amplitudes	Matthew E. Poese, Steven L. Garrett	2480
TRANSDUCTION [38]		
Thickness shearing vibration of the tangentially polarized piezoelectric ceramic thin circular ring	Lin Shuyu	2487
STRUCTURAL ACOUSTICS AND VIBRATION [40]		
Interaction between wave number pairs for a capped cylinder	J-F. Ille, J. H. Ginsberg	2493
Acoustic radiation from an elastic baffled rectangular plate covered by a decoupling coating and immersed in a heavy acoustic fluid	Olivier Foin, Alain Berry, Jeffrey Szabo	2501
On reconstruction of acoustic pressure fields using the Helmholtz equation least squares method	Sean F. Wu	2511
Active control of harmonic sound transmission into an acoustic enclosure using both structural and acoustic actuators	Sang-Myeong Kim, Michael J. Brennan	2523
NOISE: ITS EFFECTS AND CONTROL [50]		
Effects on sleep disturbance of changes in aircraft noise near three airports	Sanford Fidell, Karl Pearsons, Barbara G. Tabachnick, Richard Howe	2535
ACOUSTICAL MEASUREMENTS AND INSTRUMENTATION [58]		
Surface area measurement utilizing an acoustic bridge	Ippei Torigoe, Yasushi Ishii	2548
ACOUSTIC SIGNAL PROCESSING [60]		
Inverse problem solution techniques as applied to indirect <i>in situ</i> estimation of fish target strength	Andrzej Stepnowski, Marek Moszyński	2554
The matched-phase coherent multi-frequency matched-field processor	Gregory J. Orris, Michael Nicholas, John S. Perkins	2563
Use of higher order statistics in source signature estimation	Michael K. Broadhead, Lisa A. Pflug, Robert L. Field	2576
PHYSIOLOGICAL ACOUSTICS [64]		
Frequency responses of two- and three-tone distortion product otoacoustic emissions in Mongolian gerbils	David M. Mills	2586
Effects of acoustic trauma on acoustic enhancement of electrically evoked otoacoustic emissions	Hideko Heidi Nakajima, Allyn E. Hubbard, David C. Mountain	2603

(Continued)

CONTENTS—Continued from preceding page

Auditory-nerve-fiber responses to high-level clicks: Interference patterns indicate that excitation is due to the combination of multiple drives	Tai Lin, John J. Guinan, Jr.	2615
PSYCHOLOGICAL ACOUSTICS [66]		
The case of the missing pitch templates: How harmonic templates emerge in the early auditory system	Shihab Shamma, David Klein	2631
A model of the perceptual asymmetry between peaks and troughs of frequency modulation	Alain de Cheveigné	2645
Frequency discrimination in budgerigars (<i>Melopsittacus undulatus</i>): Effects of tone duration and tonal context	Micheal L. Dent, Robert J. Dooling, Alisa S. Pierce	2657
Feedback path variability modeling for robust hearing aids	B. Rafaely, M. Roccasalva-Firenze, E. Payne	2665
Experience with a yes–no single-interval maximum-likelihood procedure	Marjorie R. Leek, Judy R. Dubno, Ning-ji He, Jayne B. Ahlstrom	2674
SPEECH PRODUCTION [70]		
Phonological conditioning of peak alignment in rising pitch accents in Dutch	D. R. Ladd, Ineke Mennen, Astrid Schepman	2685
SPEECH PERCEPTION [71]		
Neurophysiologic correlates of cross-language phonetic perception	Anu Sharma, Michael F. Dorman	2697
Speech recognition with reduced spectral cues as a function of age	Laurie S. Eisenberg, Robert V. Shannon, Amy Schaefer Martinez, John Wygonski, Arthur Boothroyd	2704
An investigation of current models of second language speech perception: The case of Japanese adults' perception of English consonants	Sugan G. Guion, James E. Flege, Reiko Akahane-Yamada, Jessica C. Pruitt	2711
MUSIC AND MUSICAL INSTRUMENTS [75]		
Construction and interpretation of equal-tempered scales using frequency ratios, maximally even sets, and P-cycles	Richard J. Krantz, Jack Douthett	2725
BIOACOUSTICS [80]		
Optimization of a low-frequency ultrasonic technique to monitor the change in physical states in viscoelastic media: Gelation process	G. Nassar, B. Nongaillard, Y. Noël	2735
One-hydrophone method of estimating distance and depth of phonating dolphins in shallow water	Roland Aubauer, Marc O. Lammers, Whitlow W. L. Au	2744
Classification of electronically generated phantom targets by an Atlantic bottlenose dolphin (<i>Tursiops truncatus</i>)	Roland Aubauer, Whitlow W. L. Au, Paul E. Nachtigall, Deborah A. Pawloski, Caroline M. DeLong	2750

CONTENTS—Continued from preceding page

LETTERS TO THE EDITOR

Kirchhoff evaluation of scattered elastic wavefields in anisotropic media [20]	Martin Spies	2755
The noise emitted from vehicles at roundabouts [50]	W. M. To, T. M. Chan	2760
Effect of membrane motor on the axial stiffness of the cochlear outer hair cell [64]	K. H. Iwasa	2764

ACOUSTICS RESEARCH LETTERS ONLINE

Time-frequency representation of Lamb waves using the reassigned spectrogram	Marc Niethammer, Laurence J. Jacobs, Jianmin Qu, Jacek Jarzynski	L19
An integrated Printed Circuit Board (PCB) microphone	Jarmo Hietanen, Nick Zacharov	L25
Selective attention to the parameters of a physically informed sonic model	Stephen Lakatos, Perry C. Cook, Gary P. Scavone	L31

CUMULATIVE AUTHOR INDEX		2767
--------------------------------	--	------

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>).

CONTENTS

	Page
Technical Program Summary	A8
Schedule of Technical Session Starting Times	A9
Map of Atlanta Metro Area	A10
Map of Downtown Atlanta	A11
Maps of Meeting Rooms	A12
Calendar—Technical Program	A14
Calendar—Other Events	A20
Meeting Information	A22
Guidelines for Presentations	A32
Dates of Future Meetings	A34
50- and 25-Year Awards	A35
Technical Session (1a_), Tuesday Morning	2771
Technical Sessions (1p_), Tuesday Afternoon	2773
Technical Sessions (2a_), Wednesday Morning	2783
Technical Sessions (2p_), Wednesday Afternoon	2810
Technical Sessions (3a_), Thursday Morning	2832
Technical Sessions (3p_), Thursday Afternoon	2860
Plenary Session, Business Meeting and Awards Ceremony, Thursday Afternoon	2869
R. Bruce Lindsay Award Encomium	2869
Helmholtz-Rayleigh Interdisciplinary Silver Medal Award Encomium	2869
Gold Medal Award Encomium	2869
Technical Sessions (4a_), Friday Morning	2870
Technical Sessions (4p_), Friday Afternoon	2891
Technical Sessions (5a_), Saturday Morning	2911
Author Index to Abstracts	2923
Sustaining Members	2928
Application Forms	2930
Regional Chapters	2935
Index to Advertisers	2936

Time-frequency representation of Lamb waves using the reassigned spectrogram

Marc Niethammer, Laurence J. Jacobs

School of Civil and Environmental Engineering,
Georgia Institute of Technology, Atlanta, Georgia 30332-0355
mn@kyb.uni-stuttgart.de; laurence.jacobs@ce.gatech.edu

Jianmin Qu, Jacek Jarzynski

G.W. Woodruff School of Mechanical Engineering,
Georgia Institute of Technology, Atlanta, Georgia 30332-0405
jianmin.qu@me.gatech.edu; jacek.jarzynski@me.gatech.edu

Abstract: This brief note reports on a study that applies the reassigned spectrogram (the reassigned energy density spectrum of the short-time Fourier transform [STFT]) to develop the dispersion curves for multimode Lamb waves propagating in an aluminum plate. The proposed procedure first uses the spectrogram to operate on a *single*, laser-generated and detected waveform to develop the dispersion relationship for this plate. Next, a reassignment procedure is used to refine the time-frequency resolution of the calculated dispersion curves. This reassignment operation clarifies the definition of the measured modes. This study demonstrates that the reassigned spectrogram is capable of distinguishing multiple, closely spaced Lamb modes in the ultrasonic frequency range.

© 2000 Acoustical Society of America

PACS number: 43.20.Mv, 43.35.Cg

1. Introduction

This research demonstrates the effectiveness of using the reassigned spectrogram to characterize laser-generated and detected Lamb waves. By applying the reassigned spectrogram to an ultrasonic waveform measured in a flat aluminum plate, it is possible to accurately determine the dispersion relationship for this plate.

Lamb waves, which are dispersive and contain multiple modes, have received extensive attention since the study by Mindlin.¹ Recent experimental work has shown that it is possible to obtain a plate's dispersion relationship by using the two-dimensional Fourier transform (2D-FT) to operate on multiple, equally spaced waveforms.^{2,3} Unfortunately, the need for exact, spatially sampled data restricts the practicality of the 2D-FT for some inspection applications. In contrast, time-frequency representations (TFRs) require only a single signal. Recently, Prosser *et al.*⁴ used the smoothed Wigner-Ville distribution (a TFR) to determine the Lamb modes of numerically simulated waveforms in an aluminum plate. They also consider real experimental data for a composite plate and identify the s_0 and the a_0 Lamb modes for frequencies below 500 kHz. Hayashi *et al.*⁵ determined the thickness and the elastic properties of thin metallic foils (thickness of less than 40 μm) by calculating the group velocity of a single mode (the a_0 up to 3.5 MHz) using the wavelet transform (another TFR) of laser-generated and detected Lamb waves.

The current study shows that the reassigned spectrogram is an extremely accurate TFR capable of distinguishing multiple (seven in this example), closely spaced Lamb modes in the ultrasonic frequency range (up to 10 MHz).

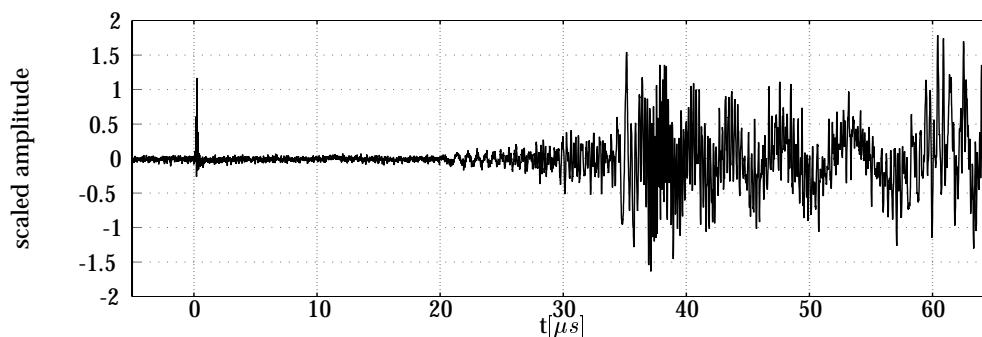


Fig. 1: Time-domain signal measured in 0.93 mm aluminum plate, propagation distance of 11 cm.

2. Transient time-domain signal

The experimental procedure makes high-fidelity (resonance-free) measurements of Lamb waves over a wide frequency range (200 kHz to 10 MHz). Broad-bandwidth Lamb waves are generated with the beam from a Nd:YAG laser (4-6 ns pulse) (see Scruby and Drain⁶ for details on laser ultrasonics). Laser detection of these waves is accomplished with a heterodyne interferometer⁷ that uses the Doppler shift to measure out-of-plane surface velocity (particle velocity) at a point on the specimen's surface. The high-fidelity, broad-bandwidth and noncontact nature of laser ultrasonics are critical elements for the success of this research. The specific plate examined is 0.93 mm thick 3003 aluminum, 203 mm long by 153 mm wide.

Figure 1 shows a (transient) time-domain signal with a propagation distance of 11 cm measured in the 0.93 mm aluminum plate. The Nd:YAG laser fires at $t = 0$ and generates a Lamb wave at the source location (the spot where the Nd:YAG hits the plate). Note that the electromagnetic discharge of the Nd:YAG's firing causes a spurious noise spike at $t = 0$. The signal in Fig. 1 is discretized with a sampling frequency of 100 MHz, low-passed filtered at 10 MHz, and represents an average of one hundred Nd:YAG shots to increase the signal-to-noise ratio.

3. The reassigned spectrogram — background

It is possible to use a TFR to transform this signal (Fig. 1) into the time-frequency domain and then quantitatively characterize the plate's features. This study establishes the effectiveness of using a specific TFR, the reassigned spectrogram, to accomplish this task. Instead of considering the Fourier transform of the entire signal at once, use the STFT to chop a signal into a series of small overlapping pieces. Each of these pieces is windowed and then individually Fourier transformed.⁸ The STFT of a function $s(t)$ is defined as:

$$S(\omega, t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-i\omega\tau} s(\tau) h(\tau - t) d\tau, \quad (1)$$

where $h(t)$ is a window function. The energy density spectrum of a STFT is defined as $E(\omega, t) = |S(\omega, t)|^2$ and called a spectrogram.

Unfortunately, TFRs such as the spectrogram suffer from the Heisenberg uncertainty principle,⁸ making it impossible to simultaneously have perfect resolution in both time and frequency. The standard deviations for time and frequency, σ_t and σ_ω ,

respectively, of the window function for a specific spectrogram are not independent of each other; the Heisenberg uncertainty principle limits a spectrogram's time and frequency resolution by the following inequality:⁹ $\sigma_t^2 \sigma_\omega^2 \geq 0.25$. Note that the window type ($h(t)$) determines the time-frequency spread of a spectrogram.⁹ For example, the product of $\sigma_t^2 \sigma_\omega^2$ is 0.2635 for a spectrogram calculated with a Hanning window. A Gaussian window function satisfies the equality $\sigma_t^2 \sigma_\omega^2 = 0.25$, but the current application aims to alter the shape of the time signal as little as possible while avoiding discontinuities across the boundaries of the windowed signal.⁹ The Hanning window is chosen as a compromise.

The time-frequency resolution of a spectrogram depends only on the window size and type and is independent of frequency. A wide window gives better frequency resolution, but worsens the time resolution, whereas a narrow window improves time resolution but worsens frequency resolution. This is in contrast to a wavelet transform;⁹ the wavelet transform tiles the time-frequency plane in an irregular fashion, resulting in a frequency dependent, time-frequency resolution. The wavelet transform of small frequency values provides good frequency resolution, but the time resolution is bad. On the other hand, the wavelet transform of large frequency values provides poor frequency resolution, but the time resolution is good.

It is possible to improve the time-frequency resolution of a spectrogram with the reassignment method, a technique developed by Auger and Flandrin¹¹ that provides a computationally efficient way to compute the modified moving window method first proposed by Koderer *et al.*¹² for the spectrogram and the scalogram (the energy density spectrum of a wavelet transform). In the reassignment method, "energy" is moved away from its original location, coordinates (t, ω) , to a new location, the re-assigned coordinates $(\hat{t}, \hat{\omega})$, thus greatly reducing the "spread" of a spectrogram. The reassignment method improves the time-frequency resolution of a spectrogram by concentrating its energy at a center of gravity. Note that the reassignment method is not restricted to a specific TFR such as the spectrogram but can be applied to any time-frequency shift invariant distribution of Cohen's class.⁸

Auger and Flandrin¹¹ show that the reassigned coordinates \hat{t} and $\hat{\omega}$ for a spectrogram are:

$$\hat{t} = t - \Re \left(\frac{S_{\mathcal{T}h}(x, t, \omega) \cdot \overline{S_h(x, t, \omega)}}{|S_h(x, t, \omega)|^2} \right) \quad (2)$$

and:

$$\hat{\omega} = \omega - \Im \left(\frac{S_{\mathcal{D}h}(x, t, \omega) \cdot \overline{S_h(x, t, \omega)}}{|S_h(x, t, \omega)|^2} \right) \quad (3)$$

where $S_h(x, t, \omega)$ is the STFT (Eq. 1) of the signal x using a normalized window function $h(t)$; and $S_{\mathcal{T}h}(x, t, \omega)$ and $S_{\mathcal{D}h}(x, t, \omega)$ are the STFT's with $t \cdot h(t)$ and $\frac{dh(t)}{dt}$ as their respective window functions. The application of Eqs. 2 and 3 is computationally straight forward and implemented with a MATLAB program.

4. The reassigned spectrogram — application to Lamb waves

Assessment of the the accuracy of the dispersion curves obtained with the spectrogram and the reassigned spectrogram requires benchmark, analytical results, obtained by solving the Rayleigh-Lamb frequency spectrum.¹ Solution of the Rayleigh-Lamb spectrum provides dispersion curves in the frequency-wavenumber (f, k) domain, whereas

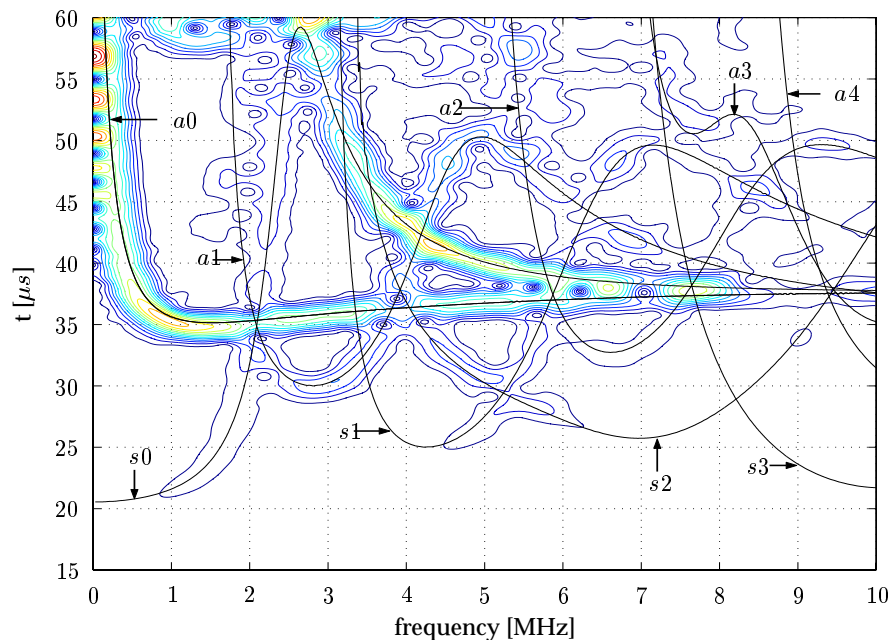


Fig. 2: “Original” spectrogram of the time-domain signal in Fig. 1 obtained with a 384-point Hanning window plus analytical modes (solid lines). Note that Figs. 2 through 4 appear in color in the archived online version of this brief note.

the spectrogram maps a signal into the time-frequency domain. To obtain the analytical dispersion curves in the time-frequency domain, the group velocities for each of the different modes at all relevant frequencies are determined by numerically differentiating f with respect to k .

Fig. 2 shows a contour plot of the square root of a spectrogram of the signal in Fig. 1 for a 384-point long Hanning window together with the analytically obtained dispersion curves (solid lines). The (experimental) s_0 and a_0 modes are clearly visible through the entire frequency bandwidth (to 10 MHz), the a_1 mode appears from 2 MHz to 7 MHz, and traces of the s_1 , s_2 and a_2 modes are evident. Overall, there is very good agreement between the analytical and experimental results, although there is a general lack of time-frequency resolution (clarity) in the experimental results. For example, it is difficult to positively identify the individual modes for frequencies above 5 MHz and times greater than $40 \mu s$. Note that Niethammer¹⁰ calculates spectrograms for a variety of Hanning window lengths for the signal shown in Fig. 1 and determines that the 384-point window provides the best compromise between time and frequency resolution for this multimode, ultrasonic signal.

The reassignment method is used to improve the time-frequency resolution of this “original” spectrogram, providing better clarity and definition of the individual modes. Fig. 3 shows a contour plot of the square root of the reassigned spectrogram obtained by applying the reassignment procedure (Eqs. 2 and 3) to the original spectrogram of Fig. 2. The reassigned spectrogram (Fig. 3) provides a crisper definition of the individual modes (when compared to the original spectrogram), and the reassigned, experimental modes are localized to the analytical curves. However, some lack of definition occurs at the intersection of modes. These “fuzzy” regions illustrate one difficulty with the reassignment method — the strongest mode (the one with the high-

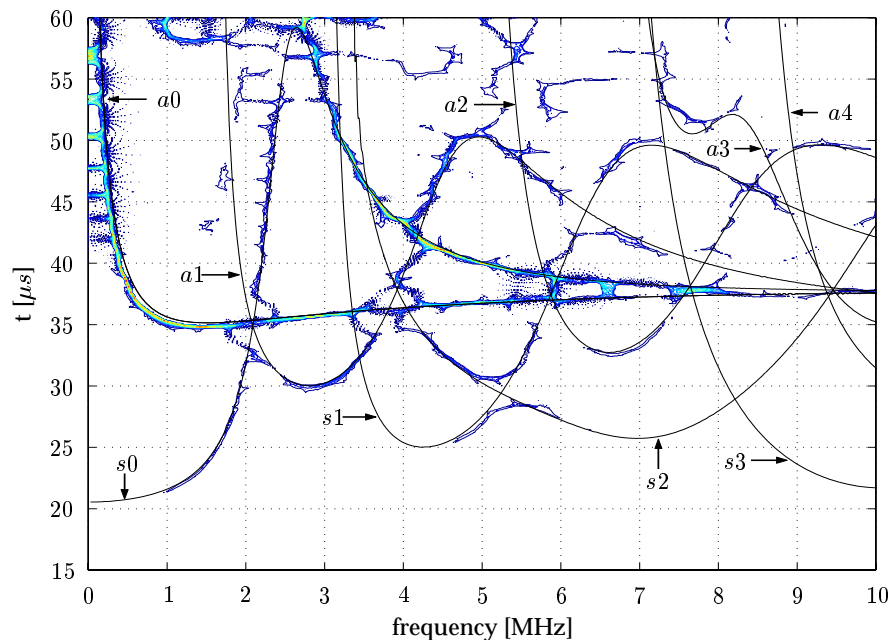


Fig. 3: Reassigned spectrogram obtained by reassignment of the original spectrogram in Fig. 2.

est amplitude in the spectrogram) becomes the mode that attracts the center of gravity during reassignment. As a result, the strongest mode remains a continuous line, but this continuity is at the expense of weaker modes that become separated in the intersection region (e.g., the intersection of the a_0 and s_0 modes around 2 MHz in Fig. 3). Finally, broken lines show up above $50 \mu s$. These are most likely caused by reflections from the boundaries of the plate and can sometimes (especially for short propagation distances to the boundaries) lead to unwanted distortion of the reassigned spectrogram. Overall, there is excellent definition of seven modes (s_0 – s_2 and a_0 – a_3) through a wide frequency range (up to 10 MHz), demonstrating that the reassigned spectrogram is capable of distinguishing multiple, closely spaced Lamb modes in the ultrasonic frequency range.

An additional portion of this research¹⁰ shows that the wavelet transform is ineffective in resolving the multiple Lamb modes of this aluminum plate through such a wide frequency range. Figure 4 shows the square root of an “original” and reassigned scalogram of the same time-domain signal (Fig. 1) calculated with a Gabor wavelet. Although the time resolution at high frequencies is very good, there is not enough frequency resolution to separate the different modes at the high frequencies (e.g., above 2 MHz). Note that the scalogram is effective in resolving the a_0 mode up to 10 MHz — an important feature for some applications. In addition, the proposed reassignment procedure does not significantly improve the time-resolution of the “original” scalogram in this example.

5. Conclusion

This note clearly demonstrates the effectiveness of applying the reassigned spectrogram to determine the dispersion curves of multi-mode Lamb waves in the ultrasonic

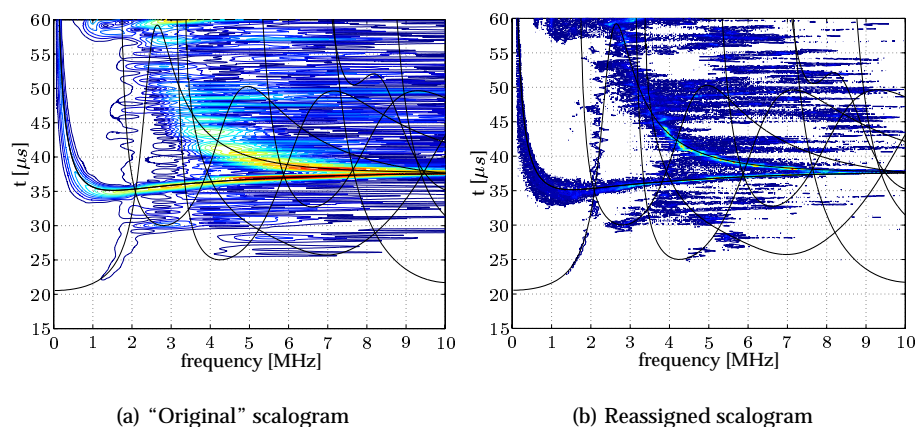


Fig. 4: Scalograms of the time-domain signal in Fig. 1, calculated using a Gabor wavelet.

frequency range, propagating in a flat plate. In general, the "original" spectrogram provides a qualitative representation of the plate's dispersion relationship, whereas the reassignment procedure refines the time-frequency resolution of these dispersion curves. Although the reassigned spectrogram has slight difficulties with mode intersections, this technique is extremely effective in localizing multiple, closely spaced modes in both time and frequency.

Acknowledgment

This work is supported by the Office of Naval Research M-URI Program "Integrated Diagnostics" (Contract number: N00014-95-1-0539). The Deutscher Akademischer Austausch Dienst (DAAD) provided partial support to Marc Niethammer. The authors thank Mr. Christoph Eisenhardt for his contributions.

References

- ¹ R.D. Mindlin, "Waves and vibrations in isotropic elastic plates," in *Structural Mechanics*, edited by J.N. Goodier and N.J. Hoff (Pergamon Press, New York, 1960).
- ² D. Alleyne and P. Cawley, "A two-dimensional Fourier transform method for measurement of propagating multimode signals," *J. Acoust. Soc. Am.*, **89**, 1159–1168 (1991).
- ³ C. Eisenhardt, L.J. Jacobs, and J. Qu, "Application of laser ultrasonics to develop dispersion curves for elastic plates," *J. Appl. Mech.*, **66**, 1043–1045 (1999).
- ⁴ W.H. Prosser, M.D. Seale, and B.T. Smith, "Time-frequency analysis of the dispersion of Lamb modes," *J. Acoust. Soc. Am.*, **105**, 2669–2676 (1999).
- ⁵ Y. Hayashi, S. Ogawa, H. Cho, and M. Takemoto, "Non-contact estimation of thickness and elastic properties of metallic foils by wavelet transform of laser-generated Lamb waves," *NDT & E Int.*, **32**, 21–27 (1999).
- ⁶ C. B. Scruby and L.E. Drain, *Laser Ultrasonics: Techniques and Applications* (Adam Hilger, Bristol, 1990).
- ⁷ D.A. Bruttomesso, L.J. Jacobs, and R.D. Costley, "Development of an interferometer for acoustic emission testing," *J. Eng. Mech.*, **119**, 2303–2316 (1993).
- ⁸ L. Cohen, *Time-Frequency Analysis* (Prentice-Hall, New Jersey, 1995).
- ⁹ S. Mallat, *A Wavelet Tour of Signal Processing* (Academic Press, New York, 1998).
- ¹⁰ M. Niethammer, *Application of Time-Frequency Representations to Characterize Ultrasonic Signals* (M.S. thesis, Georgia Institute of Technology, Atlanta, 1999).
- ¹¹ F. Auger and P. Flandrin, "Improving the readability of time-frequency and time-scale representations by the reassignment method," *IEEE Trans. Signal Processing* **43**, 1068–1089 (1995).
- ¹² K. Kodera, R. Gendrin and C. de Villedary, "Analysis of time-varying signals with small BT values," *IEEE Trans. Acoust., Speech and Signal Processing* **26**, 64–76 (1978).

An integrated Printed Circuit Board (PCB) microphone

Jarmo Hietanen^{*,#} and Nick Zacharov^{*}

^{*}Nokia Research Center, Speech and Audio Systems Laboratory, P.O.Box 100, FIN-33721 Tampere, Finland
Nick.Zacharov@nokia.com

[#]Laboratory of Computational Engineering, Helsinki University of Technology, P.O.Box 9400, FIN-02015 HUT, Finland
Jarmo.Hietanen@nokia.com

Abstract: An electret microphone has been implemented into a printed circuit board. Due to novel construction the printed circuit board serves as the microphone backplate. Furthermore, other features of printed circuit board manufacturing technology are applied to minimize the number of components of the microphone. This results in a microphone that can be manufactured with the assembly of other electronic components during the manufacturing process. As a result a complete manufacturing process has been developed. To analyze the manufacturing process and the acoustical performance of the so-called PCB-microphone, the acoustics have been modeled, and backplate surface roughness, air gap thickness, electret charge, and overall sensitivity have been measured from prototypes.

©2000 Acoustical Society of America

PACS numbers: 43.38.Kb

1. Introduction

In most consumer audio applications, the aim is to optimize audio transducers for optimum sound quality, costs, device size, industrial applicability, and other production factors that yield a realizable device. Acoustic transducers are often the limiting factors in the application.¹ For example, one limiting factor in decreasing the size of a mobile station (i.e., a mobile phone) is the physical size of acoustic transducers. Acoustic transducers are typically discrete, enclosed components, which are connected by pins or separate connection wires to the main electronics. This is due, for example, to the fact that the transducer must be encapsulated and constructed separately to be mechanically durable. Thus the space required by the housing increases the space required by the acoustic transducer in both the plane of the printed circuit board and the height axis. These factors limit the possible size reduction of portable devices in particular. In the manufacturing assembly process, acoustic transducers often cannot be automatically soldered due to high temperature considerations. This leads to increased production costs and complexity.²

To solve the majority of these problems, the integrated electret microphone and printed circuit board (PCB) concept was developed.³ The basic idea is to form the backplate, support structure, and electrical shielding of the microphone with the printed circuit board. The microphone structure has been designed to be assembled with the rest of the commercial product, providing functional benefits such as protection against contact, impact, and electrical shielding for the microphone. Only the diaphragm and the impedance transformer components need to be manufactured and preassembled. The diaphragm is premanufactured with a frame, to ensure the mechanical and acoustic properties of this critical element.

The schematic layout of the PCB-microphone can be seen in Fig. 1a. The electret microphone (1) comprises an electret film (4), which vibrates with sound. The metal layer (2) of the printed circuit board acts as a backplate (5). To separate the film and the backplate, the normal polymer coating of the printed circuit board is used in association with a separator (6). Holes (3) through the printed circuit board allow for equalization of static pressure behind the

diaphragm and sensing of pressure difference across the diaphragm. It is thus possible to create microphone directivities easily with either omni- or bi-directional characteristics. The frame (7) of the microphone film is glued with a conductive adhesive (J) to the printed circuit board contact ring (9).

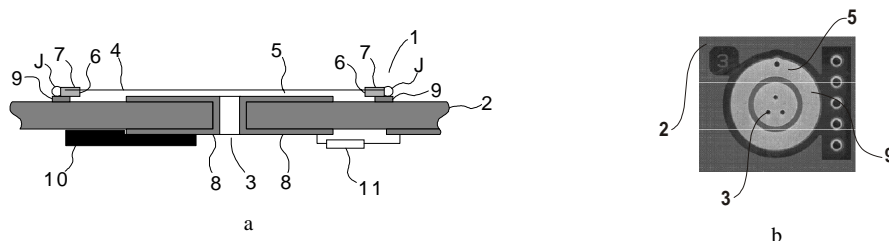


Fig. 1. a) Cross-section profile of the microphone and b) top view of the printed circuit board with three backplate holes and electrical contact pads. 1. microphone, 2. PCB, 3. hole(s), 4. diaphragm, 5. back plate, 6. separator, 7. diaphragm frame, 8. wire(s), 9. contact ring, 10. pre-polarised plastic, 11. electrical component(s) and JFET, J electrical junction (conductive adhesive).

2. Theory

Electrostatic transducers have been studied in detail since the last century. Their operation principle is well known, and detailed analyses can be found in the literature,^{4,5} where the theories have been applied to condenser and electret microphone principles based on conventional^{6,7} or more modern semiconductor^{8,9,10} manufacturing techniques. In these microphones, the transduction of signals is based on a change in the mutual geometry of two microphone plates, such as a film (also called a membrane or diaphragm) and a backplate. The transformation causes a change in the capacitance between the microphone plates. The power effect between the plates depends on the electric potential between the plates and on parameters of the mechanical structure, such as diaphragm mass, tension, air gap thickness, and hole ratio through the backplate. This system is typically modeled by a second order differential equation.

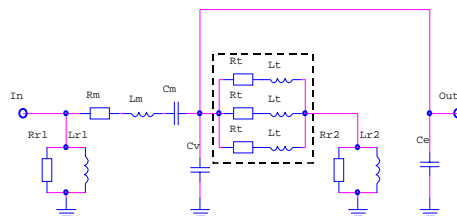


Fig. 2. Electrical equivalent circuit for PCB-microphone.

It is assumed that the amount of stretch in the diaphragm produced by a pressure change is proportional to the restoring force.¹¹ The equation of motion is easily obtained by combining Hooke's law with Newton's second law. Closer examination of the microphone shows that the air between the microphone diaphragm and backplate also provides its own contribution. In a microphone, viscous drag in the fluid is an important source of damping for the motion of the diaphragm. By applying the known complex solution to the second order differential equation, the diaphragm motion can be written

$$Ae^{i\phi}e^{i\omega t}(-\omega^2 m + i\omega R + K) = Fe^{i\omega t}. \quad (1)$$

Here the amplitude A describes the maximum extent of the motion, ω is the angular frequency, and ϕ is the phase compared to applied force, F . m is the mass of the diaphragm, R is the loss in the system, and K is the spring constant.

Alternatively, the system can be studied by creating the electrical equivalent circuit for the PCB-microphone illustrated in Fig. 1 with the well-known lumped parameter method.^{12,13} The complete electrical equivalent circuit is shown in Fig. 2. The signal source is connected to the microphone diaphragm, Z2. The diaphragm is loaded with the radiation load of a piston, Z1. The diaphragm is connected to the cavity of air gap, Z3. From the air gap, parallel tubes conduct to the back cavity of the microphone, Z4. The number of parallel tubes depends on the hole configuration. In this study no back cavity existed behind the PCB. Instead, the tubes radiated into an infinite space, Z5. Finally, a capacitor, Z6, is connected from measurement output of the PCB-microphone to the ground to attenuate the high frequencies by creating a lower-pass filter. This electrical network can also be expressed as an impedance function, Z, as follows

$$Z = \frac{1}{\left\{ \frac{1}{Z_1} + \frac{1}{Z_2 + 1/[1/Z_3 + 1/(Z_4 + Z_5) + 1/Z_6]]} \right\}} \quad (2)$$

The estimated component values in Table 1 are calculated from standard textbook equations.^{12,13}

Table 1. Component values used in the simulation.

Element	Impedance	Electrical component values		
		Resistance	Inductance	Compliance
Membrane	Z2	Rm = 10.0 Ω	Lm = 380.0 H	Cm = 120.0 fF
Volume	Z3	-	-	Cv = 7.1 fF
Parallel tubes	Z4	Rt = 53.0 MΩ	Lt = 4.2 kH	-
Radiation load 1	Z1	Rr1 = 8.0 MΩ	Lr1 = 76.0 H	-
Radiation load 2	Z5	Rr2 = 780.0 MΩ	Lr2 = 750.0 H	-
Electrical load	Z6	-	-	Ce = 20.0 pF

Z4 is a parallel circuit of tubes each having unique values. SI-units employed.

3. Manufacturing process

To produce the PCB-microphones, a HiSAC assembly cell was used. The cell is intended for the flexible assembly of electronics products. The robot is a manipulator having an arm of four degrees of freedom with an assembly site. Different tools such as a dispenser and a pressure gripper can be changed according to requirements. The components for assembly are situated in feeders.

Commercially available components were employed for prototyping the PCB-microphones, including the diaphragm and supporting frame, junction field effect transistor (JFET), and a protection diode. These components were assembled on the PCB, which had a suitable surface pattern. In Fig. 1b the PCB for one microphone backplate is presented. The assembly of the JFET and protection diode on the printed circuit board is a standard operation, thus it is not presented here. Suffice it to mention that the surface-mounted electrical components were first assembled on the side of the PCB opposite to the diaphragm. The assembly of the microphone diaphragm is described in detail.

First a pre-adhesive is dispensed to four points on the printed circuit board on top of a soldering restraint mask. The pre-adhesive drops are located partly under the diaphragm. The microphone diaphragm assembly is placed on the printed circuit board. Then the pre-adhesive is hardened in a conventional oven at a temperature of 100 Celsius. This temperature limit is critical because the microphone diaphragm is a prepolarized electret. To ensure that the joint between the frame of the diaphragm and the metal layer of the printed circuit board is conductive, a conductive adhesive is dispensed in this corner. The conductive adhesive is hardened in an oven.

The hardening times for adhesives depend on the materials selected. In this case, 10 minutes was used for the pre-adhesive and up to 120 minutes for the conductive adhesives. If UV-hardening adhesives are applied, then the presented time spans would be significantly shorter. The temperature treatment can temporarily change the mechanical properties of the microphone diaphragm and thus increase the risk of adherence to the backplate. However, this problem can be easily solved with suitable cooling procedures. Dust can be quite harmful, if it has access to the air gap of the microphone. Optical inspection of the produced PCB-microphone indicated that in 2.0% of the samples the diaphragm was adhered to the backplate. In 9.8% of samples there were dust particles in air gap preventing the microphones operation.

14

4. Measurements

Because the air gap thickness of the electret microphone is an essential factor for microphone sensitivity, a set of profile measurements was performed on manufactured PCBs. Surface formation and roughness were analysed using a Sloan Dektak II device. It was found that the surface roughness of gold-coated printed circuit boards is $\pm 1 \mu\text{m}$. The observed curvature of the measured backplates indicate that the outer ring of the pattern can be $4 \mu\text{m}$ lower than the backplate. These tolerances are appropriate for microphone applications.

When the soldering restraint mask is printed with silk-screen techniques on the top of the printed circuit board, the separation between the microphone diaphragm and the backplate, i.e., the air gap, is produced as shown in Fig. 1. Two different layer thicknesses of separator were used: 22 and 54 μm . These values were measured with the inductance technique. Henceforth these are referred to as A-type and B-type microphones, respectively. The separation of the A-type was performed with a single silk-screen printing and the B-type by repeating twice the A-type silk-screen printing. It was known by experiment that double silk printing does not produce double thickness. The measurement over the surface indicated that the actual separation for the A-type was from 20 to 46 μm and for the B-type from 36 to 58 μm varying from one printed circuit board to another. The diaphragm and the backplate were planar to within 10 μm over 10.000 μm , which should create no adverse effects. Thus, these microphone plates can be treated as plane capacitors.

The hardening process of the adhesives in the oven can be quite problematic for the electret diaphragm. The trapped charge of the electret after the hardening process was studied using the thermally stimulated current measurement (TSCM).¹⁵ In this measurement, the temperature was ramped 3 Celsius per minute from 35 up to 200 Celsius. Although the signal-to-noise ratio of the performed TSC measurement is low due to the very small size of the sample membranes, it was found out that the internal charge in thermally and nonthermally pre-treated diaphragms was approximately in the same range. This was valid when the temperature was not increased above the normal storage temperature of consumer products. When the temperature was increased above this point the internal charge starts to decay permanently. Based on these measurements it can be concluded that the employed thermal treatment of the electret diaphragm does not affect the performance of the microphone when the treatment is carefully performed. However, due to the excessive storage temperature during this manufacturing process, the achieved sensitivity of the microphones was diminished by 3 dB compared to the none heated microphone diaphragm. In Table 2, thermally stimulated current measurements are shown.

The sensitivity of the PCB-microphones was measured at 1 kHz using an acoustical pressure signal of 1 Pa. The maximum length sequence (MLS) measurement method^{16,17} was used. The median sensitivity of A-type microphones was -51 dB and of B-types was -54 dB, when the reference sensitivity was 1 V. This 3 dB difference is due to different microphone plate separations. When A- and B-type microphone sensitivities are compared to the reference microphone with an average sensitivity of -41 dB and an air gap of 75 μm , it is obvious that

sensitivities are from 10 to 13 dB less than the sensitivity of the reference microphone. Thus, the sensitivity of the PCB-microphone is more dependent upon the integrated circuit selection than due to mechanical tolerances or the temperature treatment. When clearly disfunctional microphones with sensitivities of less than -70 dB were omitted, the standard deviation of the A-type microphones was 4.6 dB (with 228 samples from 240 samples) and of the B-type 3.5 dB (192/192). The other sources of deviation should be the same for both microphone types. These include diaphragm thickness and tension, JFET parameters, and polarization voltage of the diaphragm. The sensitivity distributions of the B-type microphones are shown in Fig. 3 with a superimposed normal distribution curve. The simulated versus measured frequency responses are shown in Fig. 4. The simulation results from optimising the presented electrical equivalent circuit against the measured frequency response by using the Nelder-Mead method in the Aplac[®] program. In measurements there was an acoustical short circuit between the diaphragm and the mentioned holes, which diminished the sensitivity at low frequencies.

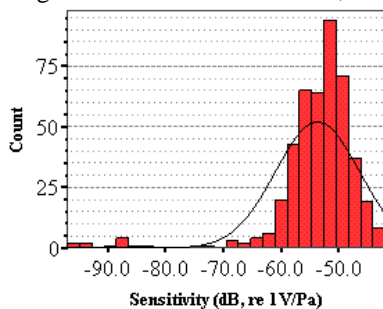


Fig. 3. The sensitivity distribution of B-type microphones with superimposed normal distribution. The used reference is 1V/Pa.

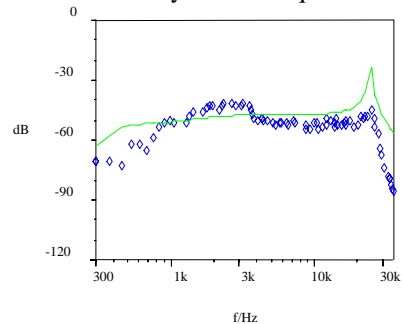


Fig. 4. The frequency response of the A-type microphone. Simulated data (solid line), measured data (diamonds).

Table 2. Thermally stimulated current measurements for none heated and heated microphone diaphragms.

Temperature	Sample 1	Sample 2	Sample 3	Sample 4
(Celsius)	No heat	Heating: 10 min.	Heating: 10+10 min.	Heating: 10+120 min.
80	5.0 pA/cm ²	2.5 pA/cm ²	2.5 pA/cm ²	2.5 pA/cm ²
100	2.5 pA/cm ²	2.0 pA/cm ²	1.0 pA/cm ²	2.0 pA/cm ²
120	1.5 pA/cm ²	1.0 pA/cm ²	0.5 pA/cm ²	1.0 pA/cm ²

Heated microphone diaphragms were warmed up to 100 Celsius during manufacturing process. First heating time is for pre-adhesive and the second heating time for adhesive.

5. Conclusion and remarks

The PCB-microphone provides considerable advantages over the prior art solutions. The microphone can be manufactured with the assembly of the PCB. Part of the structure of the microphone can be implemented already at the stage of manufacturing the PCB. Thus, in the manufacture of the microphone, it is possible to utilize a standard production line for the assembly of the integrated PCB-microphone into the end product. Implementation of the microphone according to the concept presented meets the requirements of manufacturing. A movable separator between the diaphragm and PCB is not used, and a soldering process is also avoided.

The microphone can be manufactured in a very compact form, particularly in terms of thickness. Furthermore, the microphone's outer casing and contacts have been eliminated thus leading to fewer components. This results in a simpler, more compact and cost-effective design compared to conventional microphone structures. A further benefit is the possibility for in-line assembly that eliminates the need for the microphone to be post soldered to the PCB, thus improving reliability and service life. The microphone is well protected from corrosion,

due to the smaller number of solder joints and the fact that most of the wiring can be implemented in the conductor layers of the printed circuit board.

However, before the PCB-microphone concept can be applied to mass production, the quality issues have to be implemented. First, a dust free assembly site would considerably increase the quality of production. Secondly, the height variation of the soldering restraint mask separator should be minimized. Attention should also be paid to productivity, which could be crystallized with the words *production time*. In practice, short production times could be attained with customized assembly tools and choice of the appropriate adhesives such as UV-hardening types.

To conclude, the PCB-microphone has been developed and realized. Based on this research, the acoustical behavior of the PCB-microphone can be modeled and evaluated with a fair degree of accuracy, and the manufacturing process developed. Although the presented PCB-microphone structure operates in the pressure gradient mode, it is obvious that the pressure mode can be created with a different microphone design. The attained sensitivity results are very promising. The deviations in sensitivity of the studied microphones are within expected tolerance, and the suggested action points can reduce this deviation.

Acknowledgments

The authors would like to express sincere thanks to several employees of Nokia Corporation for many useful discussions and practical arrangements. Also K. Aaltonen, J. Lehmuskoski, and R. Tuokko from Tampere University of Technology, R. Lappalainen from Helsinki University, M. Paaanen from VTT, and Aspocomp Corporation are thanked for cooperation.

References and links

- ¹ J. Hietanen, "Integration of small transducers in commercial products," 16th ICA / 135th ASA - Seattle, June 20th-26th, 917-918 (1998).
- ² Y. Yasuno, "Electret condenser mikes follow general trends for thin designs," JEE, 92-96 (1993).
- ³ J. Hietanen and N. Zacharov, "Method for implementation of an acoustic transducer, and an acoustic transducer," Nokia Mobile Phones Ltd, Finland. Patent application US-09/028,753 (1998).
- ⁴ E.C. Wente, "A condenser transmitter as a uniformly sensitive instrument for absolute measurement of sound intensity," Phys. Rev. **10**, 39-63 (1917).
- ⁵ J. Merhaut, "A contribution to the theory of electroacoustic transducers based on electrostatic principle," Acustica **19**, 283-292 (1967/68).
- ⁶ G.M. Sessler, "Electrostatic microphones with electret foil," J. Acoust. Soc. Amer. **35**, 1354-1357 (1963).
- ⁷ A.J. Zuckewar, "Theoretical response of condenser microphones," J. Acoust. Soc. Amer. **64**, 1278-1285 (1978).
- ⁸ G.M. Sessler, "Acoustic sensors," Sensors and Actuators A, **25-27**, 323-330 (1991).
- ⁹ R. Schellin, "Integrierte piezosensitive und kapazitive Ein-Chip-Silizium-Mikrofone," Ph.D. thesis, Tech. Univ. Darmstadt, Shaker Verlag, Germany, ISBN 3-8265-3333-X (1997).
- ¹⁰ M.A. Pedersen, "A polymer condenser microphone realised on silicon containing preprocessed integrated circuits," Ph.D. thesis, University of Twente, Febodruk, The Netherlands, ISBN 90-365-0994-7 (1997).
- ¹¹ D.E. Hall, *Basic acoustics*, (Krieger Publishing Co., Florida, USA 1993).
- ¹² B.B. Bauer, "Equivalent circuit analysis of mechano-acoustic structures," Transactions of the IRE. AU-2, 112-120, (1954).
- ¹³ J. Borwick, *Loudspeaker and headphone handbook*, (Butterworth & Co. Ltd., London, UK 1988).
- ¹⁴ K. Aaltonen, J. Lehmuskoski, J. Hietanen and R. Tuokko, "Precision Assembly and Joining of Mechanical Structures on Printed Circuit Board," World Manufacturing Congress 1999, September 27th-30th (1999).
- ¹⁵ G.M. Sessler, "Electrets," Topics in Applied Physics, **33**, (Springer Verlag, Berlin, Heidelberg, 1980).
- ¹⁶ M.R. Schroeder, "Integrated-impulse method measuring sound decay without using impulses," J. Acoust. Soc. Amer. **66**(2), 497-500 (1979).
- ¹⁷ D.D. Rife and J. Vanderkooy, "Transfer-Function Measurement with Maximum-Length Sequences," J. Audio Eng. Soc., **37**, 419-443 (1989).

Selective attention to the parameters of a physically informed sonic model

Stephen Lakatos

Department of Psychology, Washington State University, Vancouver, Washington 98686
lakatos@vancouver.wsu.edu

Perry C. Cook

Departments of Computer Science and Music, Princeton University, Princeton, New Jersey 08544
prc@cs.princeton.edu

Gary P. Scavone

Center for Computer Research in Music and Acoustics (CCRMA), Stanford University, Stanford, California, 94305
gary@ccrma.stanford.edu

Abstract: Two experiments tested listeners' ability to attend selectively to the properties of a physical model comprising collisions between multiple independent sound-producing objects. A probe signal paradigm measured attention to two properties – resonant frequency and number of colliding objects. Listeners completed a baseline task measuring absolute sensitivity at each stimulus against a background noise. Subsequently, stimuli served as both cues and targets; cue validity was probabilistic. When cue and target were generated by the same object (Experiment 1), greater detectability occurred with valid cues for both resonant frequency and object number, implying the presence of attentional mechanisms for these properties. When cue and target were generated by different objects (Experiment 2), selective attention persisted for object number but not for resonant frequency.

© 2000 Acoustical Society of America

PACS numbers: 43.66.J, 43.60.G, 43.75W

1. Introduction

There is growing evidence that the auditory system can parse and represent the physical properties of sound-generating sources. For example, Warren and Verbrugge (1984) discovered that listeners can distinguish between "breaking" and "bouncing" sounds of glass bottles by attending to their higher-order temporal properties. Lakatos, McAdams, and Caussé (1997) used a crossmodal matching task to measure listeners' ability to detect differences in the width-height ratios of steel and wooden bars and found that listeners are able to distinguish such dimensions by parsing the principal vibrational modes of the bars.

If physical properties of sound sources can be presented auditorially, can we attend selectively to them? This is a difficult question to address with natural sources, since it is often cumbersome to isolate individual physical attributes. Recently developed tools that can permit such manipulations, however, are physical models of sound sources implemented digitally. A number of such models is now available, ranging from exhaustive syntheses based on finite difference solutions of sound sources to computationally efficient, heuristics-based techniques that model the most perceptually salient acoustic properties of a source. To measure selective attention to source properties, we used an algorithm designed by Cook (1997) that models random collisions between particles using parametric stochastic synthesis.

Studies of auditory attention using variants of the auditory probe signal paradigm (Greenberg and Larkin, 1968) have consistently demonstrated selective attention to frequency, spatial location, and intensity (see Scharf, 1998). Recent neuropsychological findings support the notion of sensory gating mechanisms involved in attending to these attributes (e.g., Alcaini *et al.*, 1994). In the current work, we used the probe signal paradigm to examine selective attention not to frequency or spatial location, but to the acoustic

properties of a physical model of colliding objects. Using a two-interval forced choice method, we first determined signal levels that produced a uniform detectability across all stimuli for individual subjects. We then measured detectability of the same stimuli under conditions in which one signal was expected (attended) and the other was not (unattended). Attention was manipulated by cueing subjects before each trial with a weak suprathreshold signal of the kind that was likely to be presented. Greater detectability for attended signals would indicate an effect of selective attention for acoustic properties of the cue.

2. Physically Informed Sonic Modeling (PhISM)

The goal of physically informed sonic modeling (PhISM) is to couple physical simulations to efficient synthesis techniques (Cook, 1997). PhISEM (physically informed stochastic event modeling) is a PhISM algorithm based on particle models. In PhISEM, models of particles in containers were solved numerically using the basic Newtonian equations of motion. Simulations with varying numbers of particles and damping (loss of energy when particles collide with each other or the container) were run and statistics were collected about the likelihood of a sound-producing collision, the overall decay in sound energy, etc. Sound is produced only by particles hitting the container shell, because collisions between particles do not couple efficiently to the radiated sound.

The resulting PhISEM synthesis algorithm reduces the behavior of particle systems to a statistical process in which parameters relate directly to the parameters collected in the direct simulations. System energy, which represents the total kinetic energy in the system, decays exponentially. This exponential decay is rapid for systems with high damping. There is a Poisson probability of sound-producing collisions with a high waiting time for few objects and a low waiting time for many objects. This model approaches the ideal for larger numbers of particles. Sound-producing events are modeled as a short exponentially decaying of white noise, and the system resonances are modeled using biquad resonance filters. Even though the original models studied were particles within a sphere (beans within the gourd of a virtual maraca), PhISEM extends well to other systems with multiple independent sound-producing objects. From a psychoacoustic standpoint, the PhISEM model is appealing because it permits control over several physical parameters including the number of colliding objects, the damping properties, or the resonant frequency of the gourd (or chimes) themselves. We selected two parameters - the number of objects and the resonant frequency - to determine whether listeners could attend to them.

The bamboo wind chime and guiro models are different in two important ways. In the guiro, the resonance is fixed, modeling the gourd, whereas in the wind chimes, the resonance parameter is an average center resonance of a distribution of chime frequencies. Each time there is a collision in the chimes, a new frequency is allocated randomly ± 20 percent around the target center resonance. The number of objects parameter expresses the (statistical) number of bamboo cylinders in the wind chimes. The guiro number of objects parameter applies to how many serrations are caught by the stick as the guiro is scraped.

3. Experiment 1

3.1 Method

Subjects were ten undergraduates between the ages of 18 and 40 recruited from Washington State University and compensated for their efforts. None reported any hearing problems.

Four stimuli were generated with the PhISEM model of the bamboo chimes by crossing two levels of resonant frequency - "low" (1.6, 1.7 kHz spectral centroid) and "high" (4.2 kHz centroid) with two levels of object number - "few" (4-6 objects) and "many" (28-32 objects). We selected these values to generate sounds that were clearly able to be discriminated by listeners. A constant shake energy was used to generate all four stimuli. Figures 1a, b, c, and d plot the spectra for the four bamboo chime sounds averaged across the full duration of the sounds. Spectral centroid values are calculated with a formula used by

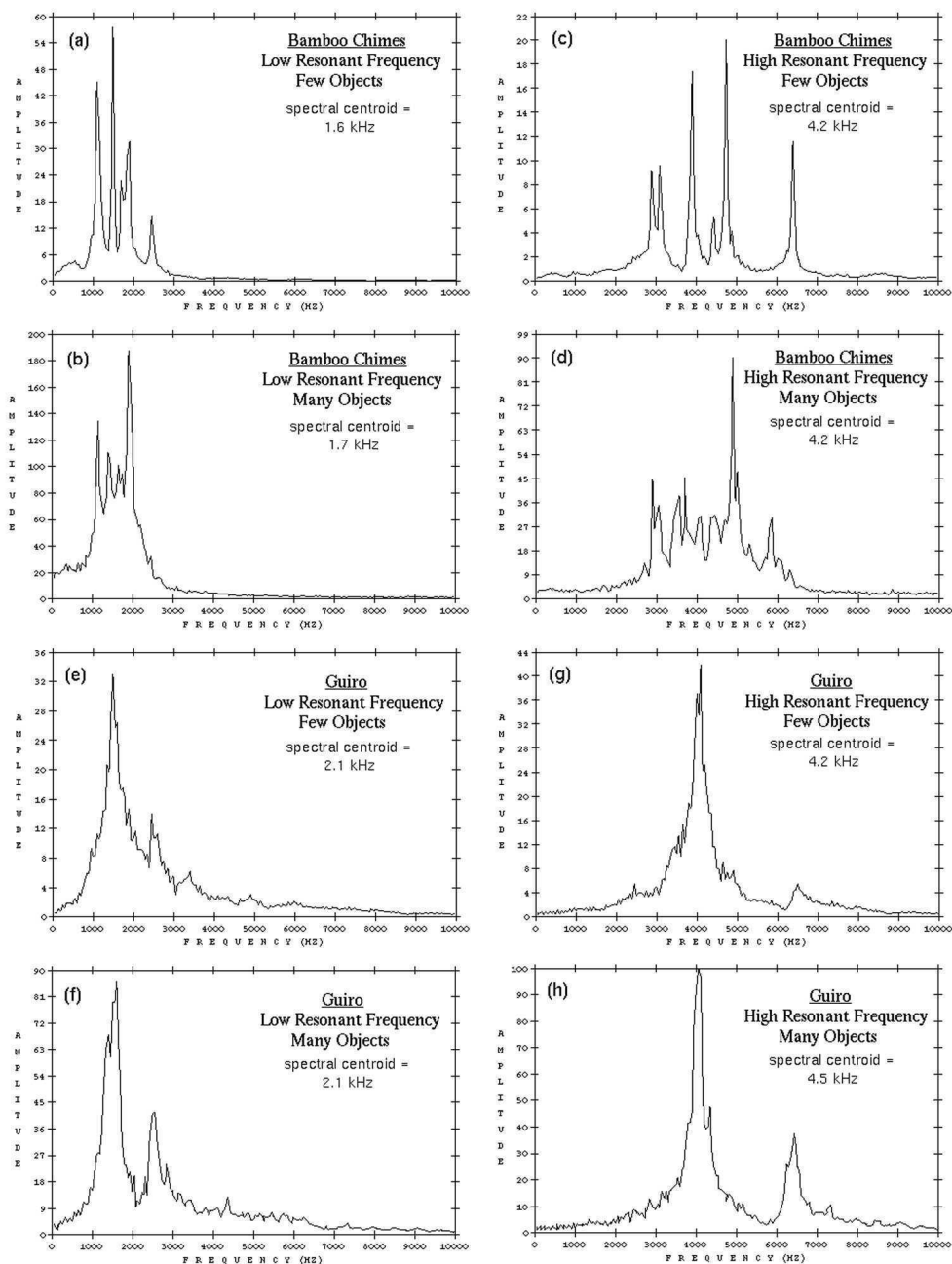


Fig 1. Spectra for the four bamboo chime stimuli in Experiments 1 and 2, and the four guiro stimuli in Experiment 2. Each set of four stimuli was generated by crossing two levels of resonant frequency - "low" (1.6-1.7 kHz centroid for bamboo chimes, 2.1 kHz for guiro) and "high" (4.2 kHz for bamboo chimes; 4.2-4.5 kHz for guiro) - with two levels of object or particle number - "few" (4-6 objects) and "many" (28-32 objects).

Beauchamp (1995). There was little spectral overlap between sounds from sources with low and high resonant frequencies. Those with similar frequencies share several resonance modes, although their peaks shift somewhat depending on the number of colliding objects.

The probe signal paradigm had two parts each lasting about one hour. In the first part, each subject completed a baseline task that measured absolute sensitivity at each of the four sounds against a background noise using an adaptive two-alternative, forced-choice (2AFC) task to ensure that thresholds were independent of criterion and all stimuli were equally detectable. Descending tracks were obtained from each subject using a 2-down, 1-up rule. Signal level was either increased by 1 dB after a single incorrect response or decreased by 1 dB after two consecutive correct responses. Thresholds corresponding to a probability of $0.5^{1/2}$ or 70.7% correct in 2AFC were obtained by averaging across 12 reversal points.

In the second part, separate probe signal tests were run for frequency and object number. For blocks of trials measuring selective attention to frequency, the frequencies of cue-target pairs were randomly selected from among four possible combinations: low-low, high-high, low-high, or high-low. For such blocks, object number was held constant within blocks at either "few" (4-6 objects) or "many" (23-32 objects). Similarly, cue-target pairs for blocks of trials for object number were randomly selected from among the following four sets of values: few-few, many-many, few-many, and many-few. Frequency was held constant within blocks at either a low (1.6, 1.7 kHz) or high (4.2 kHz) value. Measurements obtained from the first part of the experiment were used to calibrate target stimulus levels individually for each subject. Each trial began with a 625-ms cue at a level 15 dB above the subject's threshold for that sound. That cue was followed after 1000-ms by two observation intervals separated by 500 ms, one of which contained a 625-ms target signal (probe) at the level of the subject's threshold. The cue and target were derived from the active "shaking" of the sound object and omitted the exponential decay that followed cessation of the shaking (a 50-ms artificial decay was superimposed at the end of the signal), to prevent subjects from attending simply to differences in decay rates. The subject then responded by indicating which interval contained the signal. Over the session, there was a 75% likelihood that the cue and target shared the same physical property (25% likelihood that the properties differed), and therefore cue validity was probabilistic rather than certain. Greater detectability with valid cues would imply the presence of attentional mechanisms associated with these object properties.

Stimuli were presented with the SigGen/PsychoSig psychoacoustic testing software packages operating in conjunction with Tucker-Davis Technology hardware and running on a Pentium microcomputer. Subjects listened to stimuli over Sennheiser HD265 headphones.

3.2 Results

To compare detection performance across trials containing valid and invalid cues, for each subject we computed the proportions of correct responses for the two types of cues at each resonant frequency and object number. For the analyses of variance, we applied an arcsine transformation to the data to equalize variances. Figure 2 shows the untransformed data in Experiment 1 expressed as the percentage of correct responses for attended and unattended physical properties, shown separately for resonant frequency and object number [Bars for attended stimuli each comprise 720 judgments, whereas those for unattended stimuli comprise 240 judgments.] Detection performance was higher for attended (66.5%) vs. unattended (47.9%) targets generated by bamboo chimes of different resonant frequency [$F(1,9)=22.5$, $p=.001$], indicating that subjects were able to use object number in the detection task. There was no significant difference in detection performance for low (54.5%) vs. high (60.0%) resonant frequency cues pooled across attended and unattended targets [$F(1,9)=2.47$, $p=.15$].

Similarly, when object number varied, detection performance was significantly higher for attended (67.5%) vs. unattended (61.2%) targets [$F(1,9)=6.67$, $p=.03$]. Curiously, detection was better for cues comprising many objects (67.2%) vs. cues with few (61.5%), although the difference misses statistical significance [$F(1,9)=3.8$, $p=.08$]. The reason for this latter performance difference is unclear, although we suspect that because the exponential decay function for sounds generated by many objects was less steep, a relatively longer segment of the signal may have been available at or near listeners' threshold for such sounds.

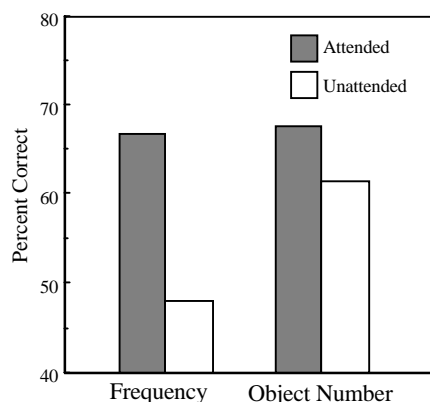


Fig. 2. Listeners' detection performance in Experiment 1.

Results from both the resonant frequency and object number conditions indicate that listeners were able to attend selectively to these properties. The ability to attend to resonant frequency is perhaps not surprising given the considerable evidence for selective attention to pure tone frequency. However, selective attention to the number of colliding objects points to attentional mechanisms more complex than those found for pure tone frequencies, perhaps arising at central levels. In this light, we attempted to increase the complexity of the detection task in Experiment 2 by determining if selective attention to resonant frequency and object number would persist even if cue and target sounds arose from *different* percussive instrument models, which nonetheless shared these two underlying physical parameters.

4. Experiment 2

4.1 Method

The ten subjects from Experiment 1 were also invited to participate in Experiment 2. All agreed to return and were compensated for their participation..

Stimuli were the four bamboo chime sounds used in Experiment 1 and four guiro sounds (see Figures 1e, f, g, and h) generated by crossing two levels of resonant frequency - "low" (2.1 kHz) and "high" (4.2-4.5 kHz) - with two levels of object number - "few" (4-6 objects) and "many" (28-32 objects). Although the spectral centroids of the low frequency guiro sounds were higher than those of the low frequency bamboo chimes - a function of differences in the physical models appropriate to simulating these two percussive instruments - we were confident that they would nonetheless serve as effective cues to resonant frequency. Thresholds for the four guiro sounds were obtained for each subject with the tracking method of Experiment 1. The probe signal paradigm was also retained, except that cue and target on each trial came from different instruments: cue (guiro)-> target (bamboo chimes), or cue (bamboo chimes)->target (guiro). Software/hardware were retained from Experiment 1.

4.2 Results

The untransformed percentage of correct responses for attended and unattended physical properties is shown separately for resonant frequency and object number in Figure 3. With different percussive instruments for cue and target, listeners showed no significant difference in detection performance for resonant frequency for attended (58.2%) vs. unattended (55.0%) targets [$F(1,9)=.01$, $p=.92$], contrary to our initial expectations. Perhaps information about resonant frequency cannot be used to direct attentional mechanisms to sounds with different sources and/or highly distinct timbres. Detection of low (56.4%) vs. high (56.8%) resonant frequency cues pooled across attended and unattended targets was not significantly different [$F(1,9) = .233$, $p=.64$]. Listeners were better able to detect attended (61.1%) vs. unattended (51.4%) targets when cued to the number of colliding objects [$F(1,9)=5.19$, $p=.05$], suggesting

that the perceptual salience of this parameter transcended the particular physical model (i.e., bamboo chimes or guiro) that it controlled. Detection was no better for cues comprising many objects (54.7%) vs. those comprising few objects (59.7%), [$F(1,9)=.41$, $p=.54$].

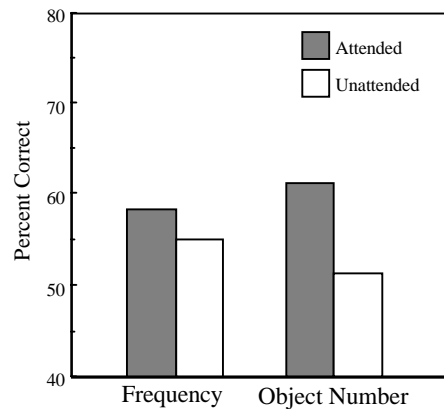


Fig. 3. Listeners' detection performance in Experiment 2.

5. Conclusion

The results reported here point to a fairly remarkable ability of the auditory system to monitor individual acoustical properties of sound sources. Perhaps most impressive was that subjects were still able to attend selectively to the number of objects in Experiment 2, even when the cue and target signals were generated by different objects. Conversely, the absence of selective attention to resonant frequency in this context may point to limitations to frequency-based attention across widely disparate timbres. It remains unclear precisely what features of the proximal waveform may be the most salient "markers" of these distal acoustic properties. However, given the quasi-random nature of the collisions simulated by the PhISM models used here and the ability of listeners to use information about properties in one instrument to attend to those in another, it is clear that the proximal spectral and temporal cues must be highly complex. In future work, we intend to isolate classes of such cues in a number of different physical models to derive a predictive, ecological model of timbre that links the acoustic properties of sounds and their sources to their perceptual correlates.

Acknowledgments

This work was supported by the Air Force Office of Scientific Research (grant #F49620-99-1-0293). We thank Candice Lindsay and Faiza Noor for their help in running participants.

References and links

- Alcaini, M., Giard, M-H., Eschali r, J-F., and Pernier, J. (1994). "Selective auditory attention effects in tonotopically organized cortical areas: A topographic ERP study," *Human Brain Mapping*, **2**, 159-169.
- Beauchamp, J. W., and Horner, A. (1995) "Wavetable interpolation synthesis based on time-variant spectral analysis of musical sounds," Audio Engineering Society Preprint No. 3960.
- Cook, P. R. (1997). "Physically inspired sonic modeling (PhISM): Synthesis of percussive sounds," *Computer Music Journal*, **21**, 38-49.
- Greenberg, G., and Larkin, W. (1968). "Frequency-response characteristic of auditory observers detecting signals of a single frequency in noise: The probe-signal method," *Journal of the Acoustical Society of America*, **44**, 1513-1523.
- Lakatos, S., McAdams, S., and Caus  , R. (1997). "Perception of auditory source characteristics: Simple geometric form," *Perception & Psychophysics*, **59**, 1180-1190.
- Scharf, B. (1998). "Auditory attention: The psychoacoustical approach," in *Attention*, edited by H. Pashler et al. (Psychology Press, Hove, U.K.).
- Warren, W., and Verbrugge, R. (1984). "Auditory perception of breaking and bouncing events: A case study in ecological acoustics," *Journal of Experimental Psychology: Human Perception & Performance*, **10**, 704-712.

The editorial below, insofar as opinions are expressed, gives the views of the Editor-in-Chief at the time of submission, and they should not be construed as unchangeable. An ongoing discussion on the future of acoustics, the Society, and the Journal is desired, and readers and members are invited to submit their own views on any topic of general interest for publication in the Forum section of the Journal.

Literate writing and collegial citing

Allan D. Pierce

Acoustical Society of America, Office of the Editor-in-Chief, P.O. Box 323, East Sandwich, Massachusetts 02537

(Received 2 March 2000)

In the present editorial, the Editor-in-Chief of the *Journal* explains the processes by which manuscripts are handled when they are submitted for publication. Various categories of problem papers are described, and it is emphasized that the outcome of the review process is not always predictable. Metrics for measuring paper quality and journal quality are reviewed and discussed. Arguments are given to the effect that the quality of a paper and its chances of being selected for publication will be considerably improved if the authors adopt a philosophy of literate writing and collegial citing. The detailed meaning of these phrases is discussed with accompanying examples, including the case of a paper by J. J. Waterston, the publication of which was delayed for 47 years until the paper was rediscovered by Rayleigh. © 2000 Acoustical Society of America.

[S0001-4966(00)05005-0]

PACS numbers: 43.05.Gv, 43.10.Gi [ADP]

INTRODUCTION

Although the quality of the papers published in this journal is, by most accounts, regarded as high, it would be even higher if all of the papers exemplified literate writing and collegial citing. Many of the papers do indeed have these qualities, but there are a sufficient number that lack them to prompt this editorial. What is here meant by the two terms, literate writing and collegiate citing, is a priori ambiguous, so an explanation must first be given of the sense in which they are used here.

The history of this journal and of the professional society that sponsors it dictates that its communications be in the English language and, moreover, in the American version of that language. It is a rich language which has proven to be considerably adaptable for the concise expression of complicated ideas. It is not an especially precise language, however, and many individual words have widely divergent meanings. Such is the case with *literate* and *collegial*. Here the term *literate* is understood to mean¹ *showing or marked by an acquaintance with the fundamentals or background of a particular field*. The term *collegial* is understood to mean *of or relating to a group of individuals belonging to the same profession or having similar objectives*. The field of interest here is acoustics, the profession is acoustics or the aggregate of related professions, and the objectives are those stated in the mission statement of the Acoustical Society—to increase and diffuse the knowledge of acoustics and to promote its practical applications.

Thus, in what follows, should the writer refer to someone as an *illiterate author*, such a person is not being identified as someone unable to read and write, but as one whose writing displays either no acquaintance with the literature of acoustics in general or no acquaintance with the prior litera-

ture of the subject treated in that author's paper.

The purposes of this editorial are as follows: (1) to explain why literate writing and collegial citing are essential to a high-quality journal article, (2) to persuade editors and reviewers that the absence of literate writing and collegial citing should be valid and important criticisms of a manuscript, (3) to persuade prospective authors to write their papers so that they do not encounter such criticisms, (4) to suggest to disappointed authors that one of the reasons, although possibly not communicated to them by the editor or the reviewers, for their papers not being published was that they evinced such criticisms, and, of course, (5) to achieve an improvement in the quality of the *Journal*.

I. THE SELECTION PROCESS

Before dealing with the principal subject matter of this editorial, it seems appropriate to review the process by which papers are selected for publication. The *Journal of the Acoustical Society of America* (JASA) currently has 30 associate editors who receive manuscripts directly from the authors. The right to submit articles is extended to everyone, and with a few exceptions all of the submitted papers are unsolicited. If no authors submitted papers to the *Journal*, it would cease to exist. The sustaining of the *Journal* as a high-quality and sizable publication requires (1) some efforts to encourage potential authors to do good research and to submit quality papers based on that research to this journal and (2) a careful selection as to which of the submitted articles should be published.

The associate editors are volunteers, and they serve without any financial compensation for the extensive amount of time that they have to devote to their editorial tasks. Over the 71 year history of the *Journal* an impressive list of dis-

tinguished acoustics researchers and professionals have served as associate editors. The list of current associate editors can be found on the back cover of this issue or in the "Information for Authors" section of the CD ROM. Principal tasks of the Editor-in-Chief are to select candidates for associate editorship, to persuade them that serving as associate editors is something they should do, and to propose such candidates to the Executive Council for appointment.

There is a strong analogy between the Editorial Board (the set of associate editors) and an academic department in a university. Just as professors have the responsibility for assigning grades in the courses they teach, so do the associate editors have the responsibility for deciding which manuscripts of those they receive are to be published. Just as professors have an academic freedom of deciding just how to teach their courses and of the detailed content of their courses, so do associate editors have the freedom of controlling the detailed process as to how they arrive at their publication decisions. Although literate writing and collegial citing may be regarded as important factors by the Editor-in-Chief, the individual associate editors have the right to regard them as minor or even inconsequential in their decisions as to whether papers are to be selected for publication.

Invariably, the associate editors use reviewers to advise them on the decision to publish. Reviewers are also unpaid volunteers and, moreover, they must remain anonymous; the only recognition they receive is to have their names included in a very lengthy list² of past reviewers that is published in the *Journal* once a year. The reviewers typically receive manuscripts unannounced (and often at very inconvenient times) from the editors with cover letters imploring their assistance. The reviewers are eminent people and busy people; they are also human beings, possibly with a variety of strong opinions and a modicum of irrepressible professional prejudices. Their knowledge of the applicable literature, although extensive, is not exhaustive.

There are certain by-products of the administrative structure and process described above that prospective authors should recognize at the outset:

- (1) Prompt handling and prompt reaching of decisions is in no way guaranteed. An author has no right to expect or demand such promptness. The process is, of course, not intentionally slow, but it is constrained because of its extensive reliance on unpaid volunteers. The overriding consideration is the quality of the *Journal*.
- (2) No submitted paper is guaranteed eventual acceptance. Although statistics are difficult to distill, it appears at this time that only about half of the submitted papers end up eventually being published. The process is not just a formality; many authors will not be happy with the outcome.
- (3) There is a substantial degree of *luck* involved in whether a paper becomes published. An author may, for example, submit two papers, one of which the author considers to be *great guns*,³ the other to be mediocre. The mediocre paper may be published and the other may not be. The present writer hesitates to use the term *mistake* in this context; the process involves human beings and there is a fair amount of statistical fluctuations in its outcome.

The process should not be judged by how it handled any one paper, but by its overall results. Such may be occasionally disturbing to individual authors, but it is a price the scientific community has to pay for using the peer review system. To paraphrase Winston Churchill's famous statement⁴ concerning democracy—*No one pretends that [the JASA system] is perfect or all-wise. Indeed, it has been said that [it] is the worst [system], except all those other [systems] that have been tried from time to time.* (Nevertheless, the possibility that the outcomes may be too capricious is worrisome to the present writer—many authors will not wish to submit a paper to a journal with a sustained reputation for capricious handling of manifestly high-quality manuscripts. One objective of the present editorial is to suggest to authors a methodology by which the capriciousness can be minimized.)

- (4) The editors are under no obligation to explain in detail why a paper is not selected for publication. The process is not intended for the continuing education of researchers and prospective journal-article writers; it exists primarily for the purpose of deciding which papers are to be published. One example of when publication is not warranted is when the editor, after an extensive search, is unable to find and recruit a reviewer with sufficient background and competence (and whose advice and opinion the editor trusts) to give an adequate review of the paper. (The present writer, like many of the associate editors, is reluctant to use the term "reject.") A paper that is not published may have considerable merit, and it may end up being published in another journal that is of equal or higher eminence than JASA. The associate editor has to make a decision, and the prolongation of that decision may be deemed inappropriate. There is no a priori reason to give the author the benefit of the doubt. (There is a certain analogy here with denying tenure to a professor at a university; many colossal mistakes have been made over the years, but the universities typically have to make a decision, often in the presence of political and economic forces, and there are often strong reasons for not stalling the decision. The tenure candidate is rarely given the benefit of the doubt, but that does not mean the end of the world for that candidate. The university, in guaranteeing life-time employment to a candidate for whom it has insufficient basis for a confident decision, risks far more than does the candidate, who would have an additional year to find alternate employment. Similarly, an author must realize that it is the Acoustical Society of America which bears the greater risk when a questionable paper is published.)

II. LITERATE WRITING AND ALLIES

The need for literate writing is well-accepted by workers in the humanities, but it is less appreciated in some of the areas that pertain to acoustics. (Such is, for example, especially so for the engineering sciences.) Lofty and eloquent arguments in favor of literate writing can be found in Mary-Claire van Leunen's *A Handbook for Scholars*,⁵ in Barzun

and Graff's *The Modern Researcher*,⁶ and in Mortimer Adler's *How to Read a Book*.⁷ Van Leunen, for example, gives the following sentences:

Scholarly writing is distinguished from all other kinds by its punctilious acknowledgment of sources. This acknowledgment is not just an empty form. ... Citation can also strengthen your rhetoric. When you must stand alone in an opinion, so be it. But when you have allies, call them to your side by citing them.

For the present author, the most telling argument in favor of literate writing is what might be regarded as an elaboration of the latter three sentences in the above quotation. When one reads a journal article, one normally desires strong assurance that the author is well acquainted with the subject and, moreover, that the ideas being proffered have not been formulated in a vacuum, without full understanding of the related ideas of one's contemporaries and predecessors. If done with sufficient skill, literate writing will go a long way in providing such assurance.

A classic case where a stronger familiarity with the literature and literate writing that *called one's allies to one's side* might have made a big difference is the case⁸ of John Jacob Waterston (1811–83). The *Journal* involved is not JASA, but the general circumstances could just as easily have occurred within modern times with JASA being the desired publication. The story also illustrates the point that not every paper “rejected” by a journal is either not new, not significant, or not correct. Waterston's paper was definitely new, it was significant, and, as eventually turned out to be the case, it was substantially correct. That Waterston's paper submitted in 1845 was not published at the time is of considerable embarrassment to the Royal Society of London, but Waterston himself has to share part of the blame.

Waterston's paper was partly literate in the sense that it did cite a substantial literature, although the citing was incomplete and capricious. The basic idea was speculative, and for this reason a reviewer stated that “this paper is nothing but nonsense.” The manuscript lay untouched in the Archives of the Society (presumably some musty room in the building that housed the Royal Society's administrative offices) until circa 1892. Rayleigh learned of its existence when he came across a little-known 1858 paper by Waterston which alluded to it. When Rayleigh went to the Archives to look at the manuscript, he confirmed that it was the first paper to correctly conjecture on what we now call the equipartition of energy (energy $kT/2$ per translational degree of freedom, regardless of the nature of the molecule). Rayleigh, being the great scientist that he was, recognized the paper for its intrinsic worth, and had the paper published in the *Philosophical Transactions of the Royal Society* in 1892, about 47 years after the original submission. Rayleigh wrote a short introduction⁹ which appeared just before the 1892 printing. In that introduction, Rayleigh gave the following relevant sentences:

One circumstance which may have told unfavourably upon the reception of Waterston's paper is that he mentions no predecessors. Had he put forward his investigation as a development of the theory of D. Ber-

noulli, a referee might have hesitated to call it nonsense. It is probable that Waterston was unacquainted with Bernoulli's work, and doubtful whether at that time he knew that Herapath had to some extent foreshadowed similar views.

In short, Waterston may have lost because he was either (1) unfamiliar with all of the relevant literature, (2) had failed to appreciate how that literature tied in with his own ideas, or (3) failed to cite and intelligently discuss that literature. Waterston's paper might have been published, although probably only after a requested revision, if Waterston had had the 1845 counterpart of a Lord Rayleigh as a reviewer, but he wasn't so lucky. (Rayleigh was born in 1842.)

III. CREDIBILITY AND PAPER Q98

To further explain the pragmatic benefits of literate writing in the context of publication in JASA, the present writer here discusses a hypothetical modern paper that might have been submitted to JASA. This paper is for brevity referred to as paper Q98. The nomenclature is such that individual papers are labeled by their level of quality: Q01, Q02, ..., etc., with Q100 being a perfect paper. Paper Q98 is slightly less than perfect. It is not a speculative paper as was Waterston's, but it has other problems.

The hypothetical paper Q98 has two authors; neither was previously known to the editor or the reviewers, and neither had published before in JASA or any of the other better-known acoustics journals. The subject matter of the paper is applicable to architectural acoustics; the text is highly mathematical, and the level of the mathematics and the elegance with which the mathematics is written are impressive. Nine references are cited in the paper and appear in the bibliography at the end. Seven of the references are to books, the other two are to journal articles. Among the books, two are older general textbooks on acoustics, one is a vintage textbook on architectural acoustics, one is a handbook devoted to mathematical functions, one is a monograph on spectral analysis, one is an older monograph on wave propagation in general, and the other three are to vintage books on mathematical physics. The two cited papers appeared in JASA over 20 years ago. There are no acknowledgments at the end of the paper and consequently no evidence of institutional or external support of the reported research.

The subject matter of Q98 deals with a standard partial differential equation that appears in acoustics and in many other fields. Standard boundary conditions are imposed; the basic feature distinguishing the problem from what one would find in standard texts is that the spatial region within which the partial differential equation applies does not have a simple shape. Perhaps with a tacit recognition that the analysis applies to subjects other than acoustics, the authors refer to the symbol c as the wave speed, rather than as the speed of sound. The one cited JASA paper whose title suggests some substantial relation to the subject matter of Q98 is a relatively pedestrian paper resting on very simple mathematics, having substantial graphical display, and having an extensive discussion of practical applications. In paper Q98,

there is actually very little explicit discussion of the two cited JASA articles; the emphasis is strongly on the mathematical development.

Paper Q98 does not correspond directly to any area in which the associate editor who received it had done research, so two reviewers are selected whose research interests have some relevance to the subject matter.

The first question of concern to the reviewers is whether the mathematical development was correct. They note that the authors present some numerical results for simple cases which agreed favorably with numerical results obtained by another method, so the likelihood of the mathematics being in error seems slight. They scan the mathematical steps, looking for a clear-cut instance of mathematical incompetence, and find none.

The next question addressed is whether the overall idea is truly new. Who can say for sure on such a matter? The reviewers do not go to the local library and devote extensive time to an exhaustive literature search. Even if they were to do so, the local library would probably prove to be inadequate. To really be sure, they would have to read the abstracts and dig into the contents of a great number of mathematical publications, many of which would have to be secured on interlibrary loan. A mere scanning of titles would have been insufficient. Instead, the reviewers think of all the related works that they recall seeing, at one time or another, and ask themselves whether they had ever seen anything quite like what was in paper Q98. The answer is no.

The question of the significance of the work is another matter to be addressed, but here—who can really say for sure what is significant? The mathematics is somewhat intricate and the succession of steps does require a nontrivial amount of thought; the problem is beyond what one would assign as a homework problem in a graduate course.

Thus one has a paper which is most probably correct, arguably new, and arguably significant. Should it be published? The present writer, were he the editor, would say no for the principal reason that the paper does not exhibit literate writing. The editor and the reviewers had no *a priori* reason at the outset for believing that the authors were literate in acoustics or in the subject matter of the paper. The references were all 20 years old or older, and the selection seemed somewhat haphazard. There was no evidence that the authors had done much of a literature search or that they had assimilated an understanding of related acoustical literature in the course of writing the paper. The present writer would have been uncomfortable in giving the authors the benefit of the doubt that the contribution was new. Given that the proposed boundary value problem is of a standard type which occurs in many branches of mathematical physics, and given the voluminous literature on partial differential equations, it seems inconceivable that some competent mathematically oriented scientist should not have tackled a closely related problem. Although it is possible that no one tackled a problem identical to that addressed by the authors, a truly literate work would have mentioned papers that addressed similar problems and would have discussed how the content of the cited papers differed from that of the paper under consideration.

An author might counter that they had indeed searched the literature exhaustively and that they found no such paper worth mentioning. Journal space is at a premium and they felt that no papers should be cited unless they are truly relevant to the work being presented. The present writer might acquiesce to such an argument if the total number of citations were much larger (say, 20 or more), with the bulk of the citations being to papers published within the past ten years. Otherwise, the assertion that nothing they had found was worth citing would be viewed as a distortion of the truth or as an excuse to avoid work that the authors did not enjoy doing.

A final criticism of the paper is that the significance of the work was not persuasively argued in the text. Literate writing may have accomplished this, but such was lacking. If the authors indeed found nothing in the past 20 years pertaining to the subject that was worth citing, then why should anyone in the next 20 years, except possibly the authors themselves, find their paper to be worth citing.

Should a revision be encouraged? The writer would here again say no. The content of the first version of the manuscript strongly suggests that the work is not of sufficient significance to warrant publication. A well-written unbiased literate revision with an accurate discussion of the existing literature would possibly only confirm this. The authors' precipitous submission of the manuscript without a careful survey of the relevant recent literature and without a well-written discussion of how their work fits into the context of the literature has severely prejudiced their case, and the credibility that any revision subsequently submitted is without bias in its reporting of the related literature will be lacking. If literate writing is required to establish the credibility of the authors' understanding of the field and of their claims that the work is both new and significant, then the literate writing must be present in the initial submission, not in a revision that the authors were coerced into writing.

IV. CITATION METRICS AND THERMOMETERS

The assertion, implicitly stated toward the end of the preceding section, that the expected number of future citations of a publication should be taken into account in the assessment of the significance of the work, is intrinsically controversial and warrants some discussion. A characteristic of modern times is that those who must make economic decisions desire quantitative indicators of quality that are easily measured. For scholarly journals, principal indicators¹⁰ used by librarians are the following:

- (1) *impact factor*—The number of citations in the current year to articles published in a specific journal in the immediately preceding two-year period divided by the total number of articles published in the same journal in the corresponding two-year period. For example, suppose¹¹ that in 1999, one finds, among all the articles in a very large collection of journals, that a certain number N_1 of the citations in these "citing articles" are to articles which appeared in JASA during the years 1997 and 1998. Also, in the same two-year period (1997 and 1998)

JASA published N2 articles. The reported impact factor would be N1/N2. This number would be reported as the 1999 impact factor for JASA.

- (2) *five-year impact factor*—The number of citations in the current year to articles published in a specific journal in the immediately preceding five-year period divided by the total number of articles published in the same journal in the corresponding five-year period. The definition is analogous to the impact factor described above, only the average is carried out over a five-year period. In the example given above in the definition of the (two-year) impact factor, one would redefine N1 to be the number of 1999 citations to articles which appeared in JASA during 1994, 1995, 1996, 1997, and 1998. The number N2 would be redefined to be the total number of articles which appeared in JASA during the same five-year period.
- (3) *cited half-life*—The number of years, going back from the current year, that account for 50% of the total citations received by the cited journal in the current year. For example, suppose that in 1999 JASA received a total number of NT citations from all the articles published in all the journals in that year, of which a number N99 were to articles published in JASA in 1999, a number N98 were to articles published in 1998, etc. Suppose in addition that the sum of N99, N98, N97, N96, and N95 is (0.47)NT, while the sum of N99, N98, N97, N96, N95, and N94 is (0.54)(NT). Then the reported cited half-life for JASA in 1999 would be 5 years.

The data on which these calculated numbers are based is readily available (although not *freely* available) to librarians, so there is ample opportunity for creative librarians with access to large computers to come up with metrics that are custom-tailored to their institutions. One can envision, for example,¹² the following:

- (4) *cost-per-citation metric*—One starts with the database listing all of the articles published during, say, 1999 by all of the faculty in the library's university. From this database, one can build a more extensive database listing all of the articles that were cited in those faculty publications. Of these cited articles, let us say that NU2 were to articles published in JASA during 1997 and 1998, and that NU5 were to articles published in JASA during 1994, 1995, 1996, 1997, and 1998. The yearly subscription price to JASA is JDOLLARS. The estimated future cost of each faculty citation to JASA can be estimated to be either (2)(JDOLLARS)/(NU2) or (5)(JDOLLARS)/(NU5), where it might be supposed that both calculations yield approximately the same number. A cost-conscious librarian with a fixed budget could then choose the journals "most appropriate" to the institution based on one or both of these metrics. Journals would be prioritized by the smallness of their cost-per-citation metric. If JASA should, for example, be number 23 in this ranking, but the total subscription costs of the journals ranked 1 through 22 exceed the library's annual budget, then the library would not subscribe to JASA.

The JASA editors could play an analogous "game" and use the same database to calculate an *effective contributory impact factor* for any given paper published in JASA. Such a factor would take into account the three ways a paper can affect, positively or negatively, JASA's overall impact factor: (a) the publication of the article adds to the number of papers that are in JASA, (b) the paper may cite previous JASA publications, and (c) the paper itself may be cited in future publications. If the JASA impact factor is I_J and the effective contributory impact factor for any given paper is I_C , then the definition of a suitable I_C must be such that the average of all the I_C 's is I_J . A mathematical analysis leading to an appropriate definition is somewhat intricate, but could presumably be worked out in a short time by most of the readers of this editorial. In the interest of brevity, the analysis is omitted here. The result, which in retrospect should be in accord with one's intuition, is the following:

- (5) *effective contributory impact factor*—Suppose a given paper, say Q50, is published in 1997. Paper Q50 cites a number NREF of references that were published in JASA in 1995 and 1996. In the two years, 1998 and 1999, one finds that, in all of the papers published in all of the journals, that Q50 is cited a total of NCITED times. Then the appropriate value I_C for paper Q50 is

$$I_C = \frac{1}{2}(\text{NREF}) + \frac{1}{2}(\text{NCITED}).$$

If a given paper's I_C is less than I_J , then that paper can be regarded as having a negative influence on the *Journal's* impact factor; if it is higher, then it has a positive influence.

A creative editor may seek to estimate what the I_C of a submitted manuscript would be should that manuscript be published; the numbers NREF and I_J are known at the outset, it being a reasonable assumption that I_J does not change much over a short period of time. The remaining number, NCITED, can be estimated from the first author's track record. In this manner, one arrives at the following definition:

- (6) *projected effective contributory impact factor*—Suppose a paper Q51 is submitted in 2000 and has NREF references to papers published in JASA in 1998 and 1999. The database shows that the first author has published, say, three papers in the ten year period ending with 1997. For these papers, the average number of citations by others in publications that appeared in the two immediately following years is (NCITED)_{TR}, where the subscript "TR" abbreviates track record. Then the projected effective contributory impact factor is

$$I_{C,\text{proj}} = \frac{1}{2}(\text{NREF}) + \frac{1}{2}(\text{NCITED})_{\text{TR}}.$$

If the author has no track record, then the number might be calculated with NCITED_{TR} set to zero.

A reader may justifiably criticize any reliance on this number, one reason being that a manuscript that an author *submits* is not necessarily representative of the author's papers that eventually end up in print. The present writer would

counter that the definition is here made with “tongue in cheek,” and one can make whatever use of it one wishes. One intriguing observation is that an author need only cite a number greater than $2I_J$ of references which were published in JASA during the past two years to achieve an increase in JASA’s impact factor. As discussed further below, an author who does this is one who is practicing *collegial citing*.

To individuals who read and use JASA, as contrasted with librarians and possibly with editors, the metrics described above are somewhat irrelevant. If an article pertains strongly to one’s research, then one should read that article, regardless of where it is published and regardless of how many times it has been cited. On the other hand, it is only human that authors would like for their work to be read and appreciated by others. It is difficult for an author to gauge how much his or her work is being read and to what depth the readers are reading a paper. The only practical measurement, however imperfect, that exists at present is how often (and in what manner) the work is cited in subsequent literature. Most scholars would strongly decry the idea that a paper that is cited twice as often as another paper is substantially better or more significant than the other paper. They would, however, sense that there is something wrong with a paper that is never cited over a ten-year period.

The idea of taking number of citations as a metric of paper quality or of journal quality is analogous to using body temperature as a measure of one’s being ill. If someone runs a temperature of 105°F (or 41°C), then all would agree that that person is ill. The person is not necessarily more ill than a person who runs a temperature of 102°F , and one would not seek to cure the illness by plunging the person into a cold water bath and waiting for the body temperature to drop to 98.6°F . Nevertheless, the fact of the illness cannot be dismissed—regardless of how one interprets the detailed significance of the numerical value of the body temperature.

That a relatively low impact factor may be of concern to some editors is exemplified by a recent editorial¹³ in the *Journal of Applied Mechanics* (JAM), which is published by the American Society of Mechanical Engineers (ASME). There the editor laments that the impact factor for JAM is lower than that of some peer journals and lower than that of two of the other journals published by the ASME. The editor continues with the statement: “*It would thus appear that there are quite a few papers appearing in the Journal which fail to be cited in the near term and it is clear that, if such a trend continues, the Journal will suffer.*” (With considerable trepidation, especially since JAM is a journal he greatly respects, the present writer suggests that the opening sentence in the introduction of the current editorial would apply equally as well to JAM.)

The dominant question remaining is whether JASA should knowingly publish papers that are unlikely to ever be cited, or which at best will be seldom cited. If the editor and the reviewers are convinced that a paper is truly “great guns,” then that paper should be published, no matter what. Perhaps someday a Lord Rayleigh will come along, discover the paper, bring it to the attention of the scientific community, and it will “blossom forth.” At that time, the Acoustical Society can congratulate itself on having had, in contrast

to the Royal Society of 1845, editors and reviewers with great perception and foresight. On the other hand, if it is certain that the work, even if by some fluke it should be extensively cited, could never be of anything but minor consequence, the present writer would argue that the paper should not be published.

V. THE EGOIST AND PAPER Q99

An *egoist* is the antithesis of a colleague, and in many cases the author of a manuscript comes across as an egoist. To illustrate the point and to persuade potential authors to write as colleagues rather than egoists, the writer here describes another hypothetical paper, this one labeled as Q99, as it is possibly of slightly better quality than Q98, but still not perfect.

Paper Q99 is authored by a person who has been writing papers for some time on a somewhat specialized topic, some of which have appeared in JASA. The current paper is, to some extent, a continuation of that research. The manuscript has 12 references, of which six are to the author’s previous papers. Of the remaining six, two are to journal articles authored more than 20 years ago, the remaining four are to generic textbooks.

The editor who receives the manuscript has a difficulty in identifying an appropriate reviewer. Ideally, such a reviewer should have some familiarity with the author’s previous papers, but the list of citations gives no clue as to who such a person might be; possibly no such person exists. The editor goes to *The Citation Index of the SCI*¹⁴ to find who has been citing the author’s work in the past and finds that there are very few citations—other than those given in subsequent publications written by the same author. None of the names of the citers are familiar to the editor, so the editor consequently sends the paper to potential reviewer A, who has a reputation for knowing just about all there is to know about this general area of acoustics, but who has no knowledge in depth about the specific topic of the submitted paper. Reviewer A declines, stating that he or she is really very busy. This process—of the editor asking, and of the reviewer declining—goes through several iterations, until eventually a good citizen is found who agrees to review the paper.

The good citizen reviewer has not read any of the author’s previous papers. Moreover, being human, the reviewer is not willing to diligently read all of those papers in preparation for a thorough review of paper Q99. The reviewer may not even look at those previous papers, possibly because the paper’s author does not write with exceptional clarity and possibly because some of the cited papers are difficult to retrieve. The disposition of the paper at this point is capricious. One possibility is that the reviewer simply goes through the manuscript and looks for obvious errors and makes notes as to suggestions that would improve the paper. The novelty of the paper is taken as a given; the author’s previous papers were all taken as novel; this is different from any previous paper by the author; and it seems certain that no one but the author would have addressed the present problem. That the paper is significant seems evident, as all the previous papers in this sequence were adjudged by other reviewers to have been significant, so this one must be also.

The good citizen reviewer sends back a long list of suggested cosmetic improvements; this list is transmitted to the author; the author submits a revision with the suggested improvements taken into account; the paper is accepted and published.

An alternate scenario is that the reviewer recommends the paper not be published because the case for the work being significant is too weak. Although this is a continuing work that has resulted in a number of previous publications by the author, there is no indication at all that anyone has been reading those papers or is carrying on related research. If there were any such person, then why didn't the author cite them? Another disturbing feature is that the author is blasting on ahead without looking around in the scientific community to see if anyone is doing work which might impact the present author's work. Perhaps it is time to call a halt to this chain of noncollegial publishing. No one other than the author would miss the next few installments.

It is not clear that the reviewer or the editor could ever convince the author that he or she has been operating as an egoist rather than as a colleague, but that is the basic problem. If the author had written all the papers in the sequence with a concerted effort to discuss the relationship of the current work with work that was being carried out by others, then the cited persons might have taken notice of the author's research. There might have been a dialog in the literature, with a synergism of work carried out by different groups. The small price that the author would have had to pay is that he or she would have to read some papers written by persons other than himself or herself; these papers would have to have been understood in some detail and then assimilated in the writing of the subsequent papers—and they would have to be cited. In brief, the author would have to assiduously cultivate the art of *collegial citing*.

A cynical reviewer might harbor the suspicion that, in actuality, the author was incapable of doing the reading and assimilation that was required to produce the literate writing and collegial citing that the papers were so strongly lacking. Perhaps at some time in the distant past, some thesis advisor had carefully laid out the relevant background and pointed the author in a certain direction. Momentum, persistence, and a certain luck in the assignment of reviewers for the author's submitted manuscripts had resulted in a healthy list of refereed publications. If one's ego is sufficiently great, or if one does not care whether anyone reads one's papers, or if there is no pressure to secure external funding for one's research, then one blithely carries on.

The present writer's view is that, were every paper in the *Journal* to have been written by an egoist (rather than a collegial) author, then the *Journal* would be in serious trouble. Possibly, most egoist authors are capable of reform; they may only need a loud "wake-up call." If so, then the discussion in this editorial might help.

VI. THE ALLEGED INTERLOPER AND PAPER Q99.5

The use of the word *interloper* is here intended to be provocative, but the adjective *alleged* is intended to soften the provocation. The standard dictionary definition of an interloper, as being one who thrusts himself or herself into any

position or affair, which others consider as pertaining solely to themselves, is much stronger than the sense which the present writer intends. Unfortunately, the English language provides no single word that succinctly conveys the image of someone who enters briefly into a group endeavor for no apparent reason and who has no intention of joining that group.

The term *interloper* is here intended to imply a hypothetical person, who may be entirely nonexistent, and who submits a paper to JASA with the following attitude:

I never read JASA, and I certainly have no intention of ploughing through its pages to find something worth citing. You may have a nice Society and have nice semi-annual meetings, but I have no intention of entering into the affairs of your Society or ever going to any of your meetings. Probably the other authors of papers in JASA would welcome me as a colleague, but I really don't care whether they do or not; I already have a fine set of colleagues, and they are all I need. Typically, I publish all of my papers in other journals, but just this once I am condescending to submit a paper to JASA. Consider yourself fortunate that I have done so. There are of course no references to JASA in this paper, but I can't imagine there would be anything previously published in JASA that would be relevant to what I have done. You may have some difficulty in finding a suitable reviewer among the membership of your organization, but if you go outside that group, you should be able to find such a person, although this paper is so much obviously better than what you usually publish that you should not need much of a review. The other authors of papers in JASA have my permission to cite this paper as much as they wish, but don't expect me to reciprocate.

The present writer agonized considerably as to whether the above paragraph should be included in this editorial. One risks being considered paranoid, for, quite possibly, no author of a submitted manuscript has such a blatantly arrogant attitude. However, some authors do incur the risk of being perceived by editors, reviewers, and JASA readers as *alleged interlopers* and they can avoid this risk by incorporating *collegial citing* in their manuscripts at the outset. Of course, if they really are interlopers, then they won't want to do this.

Paper Q99.5 is received by an editor from an author of whom the editor has no prior knowledge. The paper appears to be truly concerned with acoustics, and it is quite possible that there have been papers in JASA at one time or another that may have been related to the topic of the paper. The topic is nevertheless not a mainstream of contemporary acoustics, so the title or author of a relevant JASA paper does not come immediately to the editor's mind. The present paper has no references to JASA at all; the plurality of the references are to papers published in one particular journal, here referred to as *Journal X*. There are also some references to papers published in *Journal Y*, and some to papers published in *Journal Z*. Many of the references are to papers previously published by the author, although the number of

these is not so overwhelming that one would tag the author as an egoist.

One question the editor might ask himself or herself is why is the author submitting this paper to JASA—why not to *Journal X* instead? The author's cover letter gives no clue; it may even be the case that the format of the manuscript suggests that the author has made no special attempt to follow the instructions that are spelled out in JASA's *Information for Contributors*.¹⁵ The editor has a vague suspicion that the paper was first submitted to *Journal X* and was rejected, but the cover letter does not say this was so, and the *Journal* currently does not have a policy that requires authors to disclose such information. Also, in all handling of such papers up until the present, the editor typically disregards the fact that no JASA papers are cited.

Thus the first task at hand is to find an appropriate reviewer. If the editor cannot think of anyone that he or she knows who is an ideal match for the subject matter of the paper, the next recourse is to examine the reference list and the manner in which the references are cited to discover some clue as to whom to ask to review the paper. What becomes evident from this examination is that the author is associated with a *collegium* that is different from any with which the editor is associated. (A *collegium* is "an association of individuals of the same class or rank formed to promote their common interest in some business pursuit or enterprise.") The editor does not personally know any of the cited authors, although a few of the names are vaguely familiar. Picking the right reviewer is now analogous to the party game of "pin the tail on the donkey." The editor picks someone (reviewer X) and hopes for the best.

What bothers the present writer about the above hypothetical scenario is that, once reviewer X is selected, JASA is operating fully as a surrogate for *Journal X*. Reviewer X has no special interest in the good of JASA or of the ASA, and the expectation is that he or she will review the paper just as if it had been submitted to *Journal X*. The reviewer may have no special acquaintance with what has previously been published in JASA and will not be bothered at all by the fact that no JASA articles are cited. Even if it were so that the paper had previously been submitted to *Journal X* and turned down, there is only a small chance that the current reviewer is the same person as the previous reviewer. Given the capricious nature with which the reviewer has been selected, the editor may have picked the least discerning of all the possible reviewers that an editor of *Journal X* might have picked—thus there is a reasonable chance that the paper will be selected for publication in JASA even though it would never have been selected for publication in *Journal X*.

How could JASA have possibly handled Paper Q99.5 any differently? This is a difficult question that warrants input from a variety of individuals. The present writer's tentative method of handling such a case, were he the associate editor receiving the manuscript, would be the following. First, the paper would never be considered as submitted unless or until the author had fully complied with what is spelled out in the *Information for Contributors*. Second, the editor would act as an initial or screening reviewer and send a "review" to the author with the following criticisms or

questions. One question would be that of why there are no JASA articles cited in the paper. Given that JASA is the world's largest and oldest journal devoted to acoustics, it is highly surprising that a paper on acoustics should not have cited any JASA articles. The present writer's estimate is that, over its lifetime, JASA has published of the order of 25 000 research papers. Does the author really intend to imply that none of these papers have any relevance to the submitted paper? If such is genuinely the case, then why is the author submitting the paper to JASA? The writer would further ask the author: who does he or she hope will read the paper and make use of its results in future research? The writer will moreover ask the author, in the answer to this question, to include some specific examples of persons who either publish occasionally in JASA, present papers at ASA meetings, or who either subscribe or frequently read articles in JASA. If no such person is identified, then the author would be asked why he or she would expect any such person to regard the work as significant. The author would be told that full consideration for publication in JASA will not take place until the editor receives a literate article that shows that the author is familiar with the relevant related work that has previously been published in JASA. Even if the author continues to assert that there is no such work, the editor will insist that this be demonstrated by literate writing with citations to whatever is most closely related, even though the relationship be slight. The author will also be asked to practice some collegial citing in the writing of the paper, so the author brings out a case for why any of the JASA collegium might be able to make use of the results of the author's research.

Some readers may object to the above handling and express the feeling that it is a bit harsh. The writer has some sympathy with this feeling, and is willing to consider articulate discussions of alternative handlings. Nevertheless, the reader should realize that there is nothing illegal in the proposed scenario. Moreover, the outcome has a possibility of improving the quality of the *Journal*. Although the citation metrics discussed in Section IV may seem somewhat crude (and even crass), they do have some relevance in gauging the quality of the *Journal*. Paper Q99.5, if not substantially rewritten to incorporate appropriate literate writing and collegial citing, is expected to have a negative influence on the JASA impact factor. It achieves this in a double fashion: first, it cites no JASA papers, and, second, it stimulates no future citations from the collegium of JASA authors. If the author complies with what the writer suggests (or insists upon) above, then both of the terms, NREF and NCITED, in the effective contributory impact factor will increase. Furthermore, the *alleged interloper* ceases to be such and becomes a *colleague*.

VII. CONCLUDING REMARKS

The dominant theme of the present article is that, in the archival reporting of scientific research, a straightforward factual account of what one has done is simply not enough. The scientific literature is, almost by definition, accumulative; and this is especially so for most of acoustics. No paper

stands alone and no writer stands alone. There is a rich heritage from the past, and there are others who are thinking about similar problems.

Scientific societies such as the Acoustical Society of America were formed to bring people together who had a common interest in the progress of a certain branch of science. While some may perhaps think that *The Journal of the Acoustical Society of America* is something entirely separate from the Society that sponsors it, that is not the case. JASA is the Society's chief instrument for achieving the purpose of facilitating communication among researchers in acoustics.

Communication is a two-way process; one listens and one speaks; one reads and one writes. When one speaks, one should do so with a full cognizance of what the other conversationalist has just said. Similarly, when one writes, one should do so with a full cognizance of what has been written by others and with a full cognizance of the interests of others who one would want to read one's papers.

The present editorial argues that citing applicable prior work, giving credit where credit is due, and citing original sources of ideas and procedures used in the research is also simply not enough. The author has to explain, and often in some detail, just how the present paper fits into the grand scheme of things. The scientific literature of today is overwhelming and it is inevitable that it will become even more so in the future. A journal such as JASA plays an important role in the management of such literature. In its editorial process, it seeks a careful selection of what is being written on acoustics; it provides not a representative selection, but a quality selection of the current acoustics literature. With well-written articles, each of which exemplifies literate writing and collegial citing, the hope is that the readers of JASA will have access to a manageable source for following the dominant trends in acoustics research. The *Journal*, like the authors who publish within its pages, wants to be unique. It does not want to be regarded as just one of a vast proliferation of places where authors can store the accounting of their research results.

The phrases, *literate writing* and *collegial citing*, that figure prominently in the present editorial, are not just convenient catch phrases; they encapsulate a philosophy of archival research paper writing. Prolific writers who adopt this philosophy may perhaps find that their output is slowed down. Nevertheless, the prediction is that their impact will be considerably increased. The literate writing in the papers will increase the credibility that the authors know what they are writing about and that what is written is worth reading and worth contemplating. The outreach to other workers in the field through *collegial citing* will attract specific interest from those whose future work could benefit from the results reported in the author's papers. The author will find that his or her work is being cited and cited often, and the frustrations that one sometimes has, that one's work is neither read nor appreciated, will begin to disappear.

Writers of research articles should recognize, if they have not already done so, that good archival writing, going beyond the usage of good grammar and beyond the skillful selection of appropriate phrases, can be fun. It may not be as

much fun as doing the research itself, but there is a great potential satisfaction in the execution of a scholarly account of that research which places it fully and securely among the best literature of one's field. The realization and the relation of the fact that one is not alone is important to anyone; writing one's papers so that one tells others that they also are not alone is even more important. Just as Robinson Crusoe was elated when he found the footsteps in the sand, so will your colleagues be elated when they learn that there is someone out there, someone whom they respect, who has read and assimilated their work, and who moreover appreciates its significance and its place in the panorama of acoustics research.

¹The definitions given are paraphrased from various definitions in *Webster's Third New International Dictionary of the English Language, Unabridged*, Philip Babcock Gove, Editor-in-Chief (Merriam-Webster, Springfield, Massachusetts, 1961).

²See, for example, D. W. Martin, "Appreciation to the 1997 reviewers of manuscripts submitted to the *Journal*," *J. Acoust. Soc. Am.* **104**, 2–7 (1998).

³The (quaint, unabashedly British and Victorian) term "great guns" is taken from a remark by J. C. Maxwell in a letter to a colleague, in which he stated that he had a paper in the works which he regarded as great guns. Given what Maxwell accomplished, it is here regarded as the highest accolade that an author might give to the author's own work.

⁴W. Churchill, from a speech made in the British House of Commons, 11 November 1947; cited in *The Concise Oxford Dictionary of Quotations* (Oxford University Press, 1981), 2nd ed., p. 71.

⁵M.-C. van Leunen, *A Handbook for Scholars* (Alfred A. Knopf, New York, 1978), esp. pp. 9–10.

⁶J. Barzun and H. F. Graff, *The Modern Researcher* (Harcourt, Brace, and Jovanovich, San Diego, 1985), 4th ed. (the first edition appeared in 1957), esp. pp. 357–376.

⁷M. J. Adler, *How to Read a Book* (Simon and Schuster, 1940), esp. pp. 127–129, 276–282.

⁸S. G. Brush, *The Kind of Motion We Call Heat* (North-Holland, Amsterdam, 1976), Book 1, Chap. 3, pp. 134–149.

⁹J. W. Strutt (Lord Rayleigh), "On the physics of media that are composed of free and perfectly elastic molecules in a state of motion," paper 191 in *Scientific Papers by Lord Rayleigh* (Dover Publications, New York, 1964), Vol. 3, pp. 558–561. This is a reprinting of an introduction to a memoir by Waterston with the same name: *Philos. Trans. R. Soc. London, Ser. A* **A183**, 1–5 (1892).

¹⁰Institute for Scientific Information (ISI), "ISI hypertext terminology and concept glossary," appearing on the World Wide Web at URL site <http://www.isinet.com/help/glossary.html> (version as read on 28 February 2000).

¹¹E. Garfield, "The impact factor," *Current Contents* **25**, 3–7 (20 June 1994). Reprinted on the World Wide Web at URL site <http://www.isinet.com/hot/essays/7.html> (version as read on 28 February 2000).

¹²E. Garfield, "The application of citation indexing to journals management," *Current Contents* **33**, 3–5 (15 August 1994). Reprinted on the World Wide Web at URL site <http://www.isinet.com/hot/essays/9.html> (version as read on 28 February 2000). The hypothetical metric described in the text is inspired by Garfield's account of some activities carried out by Joshua Lederberg of Rockefeller University.

¹³L. Wheeler, "Special announcement from the technical editor," *Trans. ASME, J. Appl. Mech.* **66**, 1054 (1999).

¹⁴The *Citation Index of the SCI* is an alphabetical list by [first] author of all the references (cited items) found in footnotes and bibliographies of journals carried in the SCI [Science Citation Index]. The print version is published annually by the Institute of Scientific Information (Philadelphia) and typically extends over a large number of volumes (13 in 1998). The on-line version can be found on the World Wide Web at URL site <http://www.isinet.com/products/citation/wos.html>

¹⁵"Information for Contributors to The Journal of the Acoustical Society of America (JASA)," published in the front matter of the first issue of each volume of the *Journal*. The most recent appearance was *J. Acoust. Soc. Am.* **107**(1), ix–xiii (2000).

SOUNDINGS

Section Editor: Richard Stern

This front section of the *Journal* includes acoustical news, views, reviews, and general tutorial or selected research articles chosen for wide acoustical interest and written for broad acoustical readership.

ACOUSTICAL NEWS—USA

Elaine Moran

Acoustical Society of America, Suite 1N01, 2 Huntington Quadrangle, Melville, NY 11747-4502

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.

The 138th meeting of the Acoustical Society of America held in Columbus, Ohio

The 138th meeting of the Acoustical Society of America was held 1–5 November 1999 at the Hyatt Regency Hotel in Columbus, Ohio. This was the first time the Society has met in this city.

The meeting drew a total of 1156 registrants. There were 156 non-members and 205 students in attendance.

Attesting to the international ties of our organization, 91 of the registrants (that is, about 8%) were from outside North America, the United States, Canada, and Mexico, which accounted for 1022, 41, and 2, respectively. There were 24 registrants from Japan, 11 from the United Kingdom, 8 from France, 6 each from Germany and Italy, 5 each from Australia and People's Republic of China, 4 from Spain, 3 each from The Netherlands and Singapore, 2 each from Belgium and India, and 1 each from Brazil, Chile, Denmark, Israel, Malaysia, Romania, Russia, Solvena, Sweden, Taiwan R.O.C., Ukraine, and Venezuela.

A total of 814 papers organized into 93 sessions covered the areas of interest of all 12 Technical Committees and 1 Technical Group. The meeting also included 20 different meetings dealing with standards, indicating a healthy level of activity in this area of our endeavors.

The local meeting committee arranged the following technical tours: Borror Laboratory of Bioacoustics which houses one of the largest collections of recorded animal sounds in the world; Owens Corning Testing Systems; The Ohio State University Music Cognition Laboratories; Center for Automotive Research; Metatec Manufacturing Plant, a leading manufacturer of CDs and DVDs; and the recently restored Southern Theatre.

A pipe organ demonstration and recital was held at the The First Congregational Church in downtown Columbus in honor of the late ASA Editor-in-Chief Daniel W. Martin.

The tutorial lecture series was continued at the Columbus meeting. Uwe J. Hansen of Indiana State University and James M. Pyne of The Ohio State University presented "An Evening with the Art and Science of Music: The Families of Musical Instruments" to an audience of about 140. The

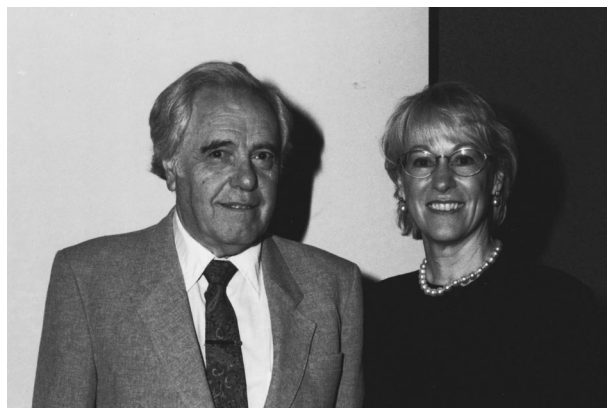


FIG. 2. Ronald Verrillo, recipient of the Silver Medal in Biomedical Ultrasound/Bioresponse to Vibration with ASA President Patricia Kuhl.

presentation included performances by The Ohio State University String Quintet, The Ohio State University Brass Ensemble, The Ohio State University Chamber Orchestra and special keyboard and percussion demonstrations.

An equipment exhibit was held Monday evening through Wednesday afternoon including exhibits on computer-based instrumentation, sound-level meters, sound intensity systems, signal processing systems, devices for noise and vibration control and acoustical materials.

Other special events included a Gallery of Acoustics, a concert performance by The Ohio State University Marching Band, receptions for students in various areas of acoustics, and a dinner for past Presidents of the Society to discuss current and future directions of the Society. The Fellows Lun-



FIG. 1. Ilene J. Busch-Vishniac, recipient of the 1998 ASA Science Writing Award for Professionals in Acoustics.

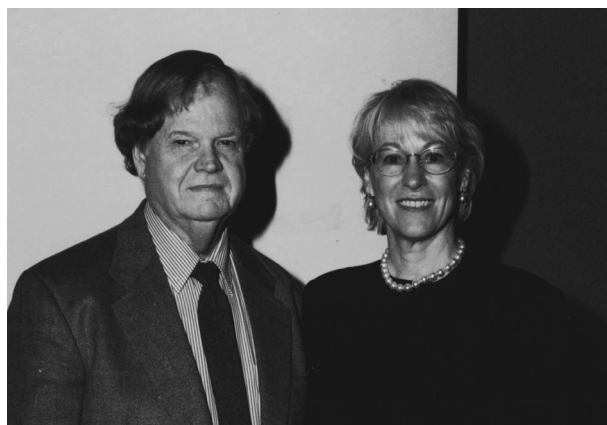


FIG. 3. Larry Royster, recipient of the Silver Medal in Noise with ASA President Patricia Kuhl.



FIG. 4. David Feit, recipient of the Trent-Crede Medal with ASA President Patricia Kuhl.

cheon was held and included a presentation by Dr. Lonnie Thompson of The Ohio State University titled "Global Climate Change, Past and Present."

The plenary session included the presentation of awards, announcement of newly elected Fellows of the Society, and recognition of the meeting organizers. The president delivered a "State of the ASA" address in which she reviewed the current economic status of the ASA, recent membership statistics, the results of the ASA Web-based survey of the membership, and invited comments from all members on the future of ASA. Four Society awards were presented and the election of 16 new Fellows was announced (see Figs. 1–5). The Science Writing Award for Professionals in Acoustics was presented to Ilene J. Busch-Vishniac for her article in the July 1998 issue of *Physics Today* titled "Trends in Electromechanical Transduction." The Silver Medal in Biomedical Ultrasound/Bioresponse to Vibration was presented to Ronald R. Verrillo of Syracuse University "for contributions to the psychophysics and physiology of vibrotactile sensitivity." The Silver Medal in Noise was presented to Larry H. Royster of North Carolina State University "for contributions to worldwide hearing conservation." The Trent-Crede Medal was presented to David Feit of the David Taylor Naval Research and Development Center "for contributions to high frequency noise reduction from submerged structures and to the vibration of fuzzy structures."

Election of the following persons to Fellow grade was announced: James W. Beauchamp, Ann E. Bowles, Arlene E. Carney, René Carré, Christopher W. Clark, Harold Conklin, Robert E. Green, Jr., Robert D. Hellweg, Jr., R. Lawrence Kirkegaard, Peter B. Nagy, Kenneth P. Roy, Nancy S. Timmerman, Sadayuki Ueha, and Shigeru Yoshikawa (see Fig. 5).



FIG. 5. Newly-elected Fellows of the Acoustical Society of America receive their certificates from ASA Vice President Mauro Pierucci and President Patricia Kuhl. First row: Mauro Pierucci, Kenneth P. Roy, Peter B. Nagy, Robert D. Hellweg, Arlene E. Carney, René Carré, Patricia Kuhl. Second row: Anne E. Bowles, Robert E. Green, Jr., Richard H. Talaske, Christopher W. Clark, James M. Beauchamp, Nancy Timmerman.

The President expressed the Society's thanks to the Local Committee for the excellent execution of the meeting, which clearly required meticulous planning. She introduced the Chair of the Meeting, Mardi C. Hastings, who acknowledged the contributions of the members of her committee including: Larry Feth, Technical Program Chair; Keith Johnson, Audio-Visual; Dick Godfrey, Exhibits; Jim Pyne, Special Events; Doug Pape, Paper Copying Service; Rob Fox, Technical Tours; Nandini Iyer, Administrative Assistance and Accompanying Persons Program; and Corrie Derenburger, Signs.

The President also extended thanks to the members of the Technical Program Organizing Committee: Larry Feth, Technical Program Chair; Kevin D. Heaney and James F. Lynch, Acoustical Oceanography; Larry Pater, Animal Bioacoustics; Angelo J. Campanella, Architectural Acoustics and Noise; Inder R. J. Makin, Biomedical Ultrasound/Bioresponse to Vibration; Uwe J. Hansen, Education in Acoustics; Ahmet Selamet, Engineering Acoustics; James P. Cottingham, Musical Acoustics; Robert Keolian and James M. Sabatier, Physical Acoustics; Lisa Stover, Psychological and Physiological Acoustics; James V. Candy, Signal Processing in Acoustics; Marios Fourakis and Mark Pitt, Speech Communication; Courney B. Burroughs and Jerry H. Ginsberg, Structural Acoustics and Vibration; and Timothy Duda, Underwater Acoustics.

PATRICIA K. KUHL
President 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.

2000	
17–19 May	ASNE Day 2000, Arlington, VA [Andrea Zari, Meetings Department, American Society of Naval Engineers, 1452 Duke St., Alexandria, VA 22314-3458; Tel.: 703-836-6727; Fax: 703-836-7491; E-mail: azari@navalengineers.org].
30 May–3 June	139th meeting of the Acoustical Society of America, Atlanta, GA [Acoustical Society of America, Suite 1NO1, 2 Huntington Quadrangle, Melville, NY 11747-4502; Tel.: 516-576-2360; Fax: 516-576-2377; E-mail: asa@aip.org; WWW: asa.aip.org].
8–10 June	Synaptic Function in Hearing and Balance, Johns Hopkins Center for Hearing and Balance, Baltimore, MD [Paul Fuchs, Tel.: 410-955-6311; E-mail: sfhb@bme.jhu.edu; WWW: www.bme.jhu.edu/labs/chb/symposium].
22–24 June	2000 Binaural Hearing, Hearing Loss & Hearing Aids, Iowa City, IA [Richard Tyler, Tel.: 391-356-2471; E-mail: tyler@uiowa.edu; WWW: www.medicine.uiowa.edu/otolaryngology/news/news].
13–16 July	ClarinetFest 2000, Norman, OK [Dr. Keith Koons, Music Dept., Univ. of Central Florida, P.O. Box 161354, Orlando, FL 32816-1354; Tel.: 407-823-5116; E-mail: kkoons@pegasus.cc.ucf.edu].
27–28 July	First International AutoSEA Users Conference, San Diego, CA [Vibro-Acoustic Sciences, Attn: Ali Behnam, 12555 High Bluff Dr., Suite 310, San Diego, CA 92130, Tel.: 858-350-0057; Fax: 858-350-8328, E-mail: info@vasci.com, WWW: www.vasci.com/company/events/asuc/].
21–23 September	Eighth Annual Conference on the Management of the Tinnitus Patient, Iowa City, IA [Richard Tyler, Tel.: 391-356-24271; E-mail: tyler@uiowa.edu; WWW: www.medicine.uiowa.edu/otolaryngology/news/news].
4–8 December	Joint Meeting: 140th meeting of the Acoustical Society of America/NoiseCon 2000, Newport Beach, CA [Acoustical Society of America, Suite 1NO1, 2 Huntington Quadrangle, Melville, NY 11747-4502; Tel.: 516-576-2360; Fax: 516-576-2377; E-mail: asa@aip.org; WWW: asa.aip.org].

2001

30 April–3 May 2001 SAE Noise & Vibration Conference & Exposition, Traverse City, MI [Patti Kreh, SAE Int'l., 755 W. Big Beaver Rd., Suite 1600, Troy, MI 48084; Tel.: 248-273-2474; Fax: 248-273-2494; E-mail: pkreh@sae.org].
Deadline for submitting abstracts: 14 July 2000.

4–8 June 141st Meeting of the Acoustical Society of America, Chicago, IL [Acoustical Society of America, Suite

7–10 October

1NO1, 2 Huntington Quadrangle, Melville, NY 11747-4502; Tel.: 516-576-2360; Fax: 516-576-2377; E-mail: asa@aip.org; WWW: asa.aip.org].

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

REVIEWS OF ACOUSTICAL PATENTS

Lloyd Rice

11222 Flatiron Drive, Lafayette, Colorado 80026

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

Reviewers for this issue:

GEORGE L. AUGSPURGER, *Perception, Incorporated, Box 39536, Los Angeles, California 90039*

DAVID PREVES, *Songbird Medical, Inc., 5 Cedar Brook Drive, Cranbury, New Jersey 08512*

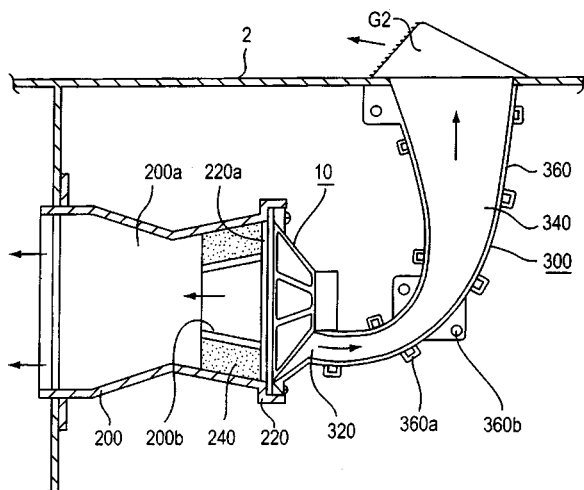
5,822,443

43.38.Ja SPEAKER SYSTEM FOR A TELEVISION

Jae-Nam Kim, assignor to Samsung Electronics Company, Limited

13 October 1998 (Class 381/388); filed in Republic of Korea 22 April 1996

Loudspeaker 10 is located in the back portion of a TV cabinet, where there is unused space available. Sound is conducted to the front and top of the cabinet through "amplifying horns" 200a and 340. "Sound absorbing



member" 240 is intended to minimize standing waves in the forward horn and thus, "... obtain a clear, high quality sound."—GLA

5,854,435

43.38.Ja NATURAL RESONANT SPEAKER

Young-ahn Kim, Teajon, Republic of Korea

29 December 1998 (Class 84/291); filed in Republic of Korea 13 October 1995

The basic concept of this invention crops up every five years or so. In this instance, a speaker enclosure is intended to mimic resonant-plate musical instruments such as the bass viol, both in appearance and performance. The front plate serves as an extension of, and is driven by, the cone of an otherwise conventional loudspeaker.—GLA

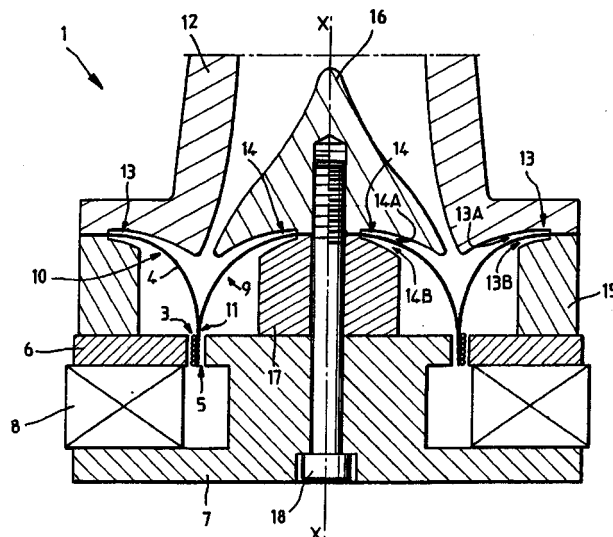
5,875,252

43.38.Ja LOUDSPEAKER FOR HIGH FREQUENCIES

Philippe Lesage, assignor to P. H. L. Audio

23 February 1999 (Class 381/156); filed in France 16 June 1995

At first glance this appears to be a cross section of the original JBL "Ring Radiator," introduced in the 1950s. In this case however, trough-shaped diaphragm 4 is much deeper and, "... is made from a flexible and



elastic material so as to exhibit a defined deformation during the motion of the electromagnetic coil."—GLA

5,910,991

43.38.Lc METHOD AND APPARATUS FOR A SPEAKER FOR A PERSONAL COMPUTER FOR SELECTIVE USE AS A CONVENTIONAL SPEAKER OR AS A SUB-WOOFER

Douglas M. Farrar, assignor to Apple Computer, Incorporated
8 June 1999 (Class 381/59); filed 2 August 1996

Most personal computers contain an inexpensive little loudspeaker to reproduce blips, bleeps, and speech. The patent document states that, "It is possible, and now deemed desirable, to provide a very high quality monophonic speaker contained in the computer enclosure, a speaker capable of reproducing lower bass frequency (sic) than many external speaker sets. In this situation, it is advantageous to use the internal high quality speaker as a 'subwoofer,' while allowing the external speakers to behave as 'satellite'

speakers.” Or, perhaps the user wants to plug in headphones and disable the internal speaker entirely.

Most audio engineers would simply put a selector switch on the face of the cabinet, but remember, this is a computer cabinet. An analog multiplexer, comparator, summing amplifier, and a few hundred lines of code will do the job nicely.—GLA

5,894,119

43.38.Md DEVICE FOR ENABLING LISTENERS TO PREVIEW MUSIC RECORDINGS

Bruce Tognazzini, assignor to Sun Microsystems, Incorporated
13 April 1999 (Class 235/375); filed 1 July 1996

A browser in a music store carries a small, pen-shaped device that functions as a bar code reader, radio transceiver, and audio amplifier. To preview a compact disk, the customer simply scans its bar code. A signal is transmitted to a digital storage unit and corresponding program material is transmitted back to the customer, who listens through headphones. The invention is well thought out and decidedly clever.—GLA

5,867,582

43.38.Si HEADPHONE

Atsushi Nagayoshi, assignor to Matsushita Electric Industrial Company, Limited
2 February 1999 (Class 381/370); filed in Japan 22 February 1994

A headphone housing contains not only the headphone receiver but an inertia transducer to generate secondary vibrations felt through the housing. The effect is said to be pleasant.—GLA

5,822,440

43.38.Tj ENHANCED CONCERT AUDIO PROCESS UTILIZING A SYNCHRONIZED HEADGEAR SYSTEM

Randy Oltman et al., assignors to The Headgear Company
13 October 1998 (Class 381/82); filed 7 April 1997

Rather than relying exclusively on (enormous) stage loudspeaker systems, a pop music concert might provide far listeners with personal audio receivers synchronized to the main sound system. This much is prior art. If a listener stays at one location, synchronization is easy. The invention is an improved method for automatically switching to a properly synchronized transmission channel, no matter where the listener wanders.—GLA

5,802,180

43.38.Vk METHOD AND APPARATUS FOR EFFICIENT PRESENTATION OF HIGH-QUALITY THREE-DIMENSIONAL AUDIO INCLUDING AMBIENT EFFECTS

Jonathon Stuart Abel and Scott Haines Foster, assignors to Aureal Semiconductor Incorporated
1 September 1998 (Class 381/17); filed 17 January 1997

The invention combines fixed filters with variable-gain amplifiers to convey accurate impressions of three-dimensional sound fields. The patent document is fairly long and quite technical, but it is clearly written in English rather than legalese. Those interested in the field are advised to obtain a copy.—GLA

5,812,674

43.38.Vk METHOD TO SIMULATE THE ACOUSTICAL QUALITY OF A ROOM AND ASSOCIATED AUDIO-DIGITAL PROCESSOR

Jean Marc Jot et al., assignors to France Telecom
22 September 1998 (Class 381/17); filed in France 25 August 1995

The method described here incorporates three primary characteristics with a number of secondary attributes plus artificial reverberation in a kind of signal processing “room” module. The goal is to simulate the acoustical quality of a virtual sound source as well as its location. “This method can be used to modify sound signals coming from a real source, or to create sound effects on recording media.”—GLA

5,910,990

43.38.Vk APPARATUS AND METHOD FOR AUTOMATIC EQUALIZATION OF PERSONAL MULTI-CHANNEL AUDIO SYSTEM

Dae-Young Jang, assignor to Electronics and Telecommunications Research Institute; Korea Telecom
8 June 1999 (Class 381/1); filed in Republic of Korea 20 November 1996

Numerous prior patents are concerned with electronic correction of listening environments. Digital equalization schemes usually start out by measuring transfer functions between individual loudspeakers and the preferred listening location, and then using various proprietary algorithms to improve the situation. This invention adds another level of sophistication by including synthesized phantom channels.—GLA

5,935,166

43.66.Ts IMPLANTABLE HEARING ASSISTANCE DEVICE WITH REMOTE ELECTRONICS UNIT

Joel Kennedy, assignor to St. Croix Medical, Incorporated
10 August 1999 (Class 623/10); filed 25 November 1996

An electronics unit for use with a middle ear implant or a cochlear implant is implanted subcutaneously in a remote location from the ear such as the pectoral region, allowing use of a larger battery and simplifying its replacement. The input signal is sent via wired or wireless links from an input transducer to the remote electronics unit; similarly the output signal from the remote electronics unit is sent to an output transducer in the middle ear or cochlea.—DAP

5,944,672

43.66.Ts DIGITAL HEARING IMPAIRMENT SIMULATION METHOD AND HEARING AID EVALUATION METHOD USING THE SAME

Dong-wook Kim and Young-cheol Park, assignors to Samsung Electronics Company, Limited
31 August 1999 (Class 600/559); filed 15 April 1998

A simulated hearing sensitivity loss is calculated for 20 critical bands and stored digitally in a hearing loss table. Audio inputs are analyzed digitally and an input level-dependent gain is calculated for each critical band to compensate for the simulated hearing loss and to provide an output signal at

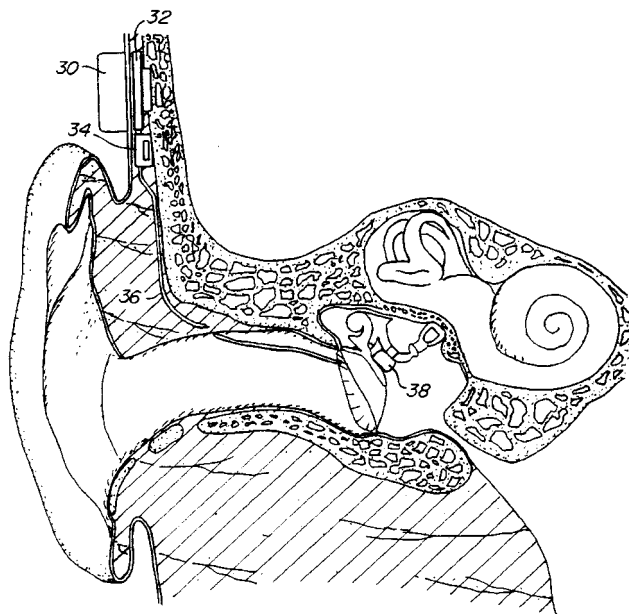
a most comfortable listening level. The invention is said to allow a normal hearing person to hear as a hearing impaired person does and to evaluate the performance of the simulated hearing aid.—DAP

5,949,895

43.66.Ts DISPOSABLE AUDIO PROCESSOR FOR USE WITH IMPLANTED HEARING DEVICES

Geoffrey R. Ball *et al.*, assignors to Symphonix Devices, Incorporated
7 September 1999 (Class 381/326); filed 7 September 1995

A low cost, low power disposable audio processor for use with a middle ear implantable hearing aid is held against the skin by magnets or adhesive. Included may be a microphone, a battery, a coil and a monolithic integrated circuit. For simplification, rather than having complicated adjust-



ments, the processor performance is tailored to a hearing loss category, resulting in several different processors. Processed signals are modulated and transmitted via electromagnetic field through the skin to an implanted receiver and output transducer.—DAP

5,890,111

43.70.Dn ENHANCEMENT OF ESOPHAGEAL SPEECH BY INJECTION NOISE REJECTION

Hector Raul Javkin *et al.*, assignors to Technology Research Association of Medical Welfare Apparatus
30 March 1999 (Class 704/226); filed 24 December 1996

A person who has had the larynx removed, due to cancer or another reason, has few options in being able to speak. One of the possibilities involves injecting air into the esophagus, which then vibrates in a way which can be heard as speech. However, the injection of air causes a disturbing sound, disrupting the speech. This device combines the results of two speech analysis pathways, then performs vector classification with a hidden Markov model classifier, to discriminate between speech and air noises. The signal is cut off during the air injection pulses.—DLR

5,897,614

43.72.Ar METHOD AND APPARATUS FOR SIBILANT CLASSIFICATION IN A SPEECH RECOGNITION SYSTEM

Frank Albert McKiel, Jr., assignor to International Business Machines Corporation
27 April 1999 (Class 704/208); filed 20 December 1996

The device is a special-purpose sibilant recognizer sub-system for use in a speaker-independent speech recognizer. The sibilant analyzer receives a voiced/unvoiced decision from other parts of the system. In either case, amplitude measurement, spectrum analysis, and format tracking are done. According to the patent, a final classification can then be made from the formant characteristics. An example is given in which "sh" is distinguished from "ch." No reference is made to affrication.—DLR

5,889,857

43.72.Gy ACOUSTICAL ECHO CANCELLER WITH SUB-BAND FILTERING

Jerome Boudy and Francois Capman, assignors to Matra Communication
30 March 1999 (Class 379/410); filed in France 30 December 1994

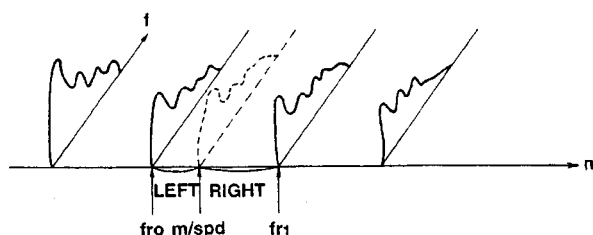
This echo canceller, intended for use in a hand-held device, such as a cellular phone, divides the signal spectrum into multiple bands. An echo signal estimate is subtracted from the outgoing voice signal independently in each band. The method differs from prior echo cancellers in that it uses a QR decomposition to estimate the echo signal in each band, rather than the more common Fourier transform. This reduces computation time and aliasing effects due to overlap between adjacent bands.—DLR

5,926,788

43.72.Gy METHOD AND APPARATUS FOR REPRODUCING SPEECH SIGNALS AND METHOD FOR TRANSMITTING SAME

Masayuki Nishiguchi, assignor to Sony Corporation
20 July 1999 (Class 704/265); filed in Japan 20 June 1995

This speech coder uses line spectral pairs (LSP) parameters for the spectral representation in order to allow maximum flexibility of speed control during resynthesis. During resynthesis at reduced playback speeds,



a method is provided for interpolating spectral coefficients in order to insert additional frames. The interpolation mechanism provides the added benefit of allowing adjustment of the transmission bitrate.—DLR

5,890,115

43.72.Ja SPEECH SYNTHESIZER USING WAVETABLE SYNTHESIS

Terry Lynn Cole, assignor to Advanced Micro Devices, Incorporated
30 March 1999 (Class 704/258); filed 7 March 1997

The patent describes a method for generating data tables for a wave-table music synthesizer, resulting in synthetic speech as the synthesizer out-

put signal. A reference speech signal is sampled and divided into segments representing each of the various phonemes. Various wavetable controls, including tremelo and vibrato, are then used to introduce variations during playback, such as pitch, loudness, and duration. These variations are said to allow the synthesis of multiple voices, emotions, and other speaker effects.—DLR

5,890,113

43.72.Ne SPEECH ADAPTATION SYSTEM AND SPEECH RECOGNIZER

Keizaburo Takagi, assignor to NEC Corporation
30 March 1999 (Class 704/231); filed in Japan 13 December 1995

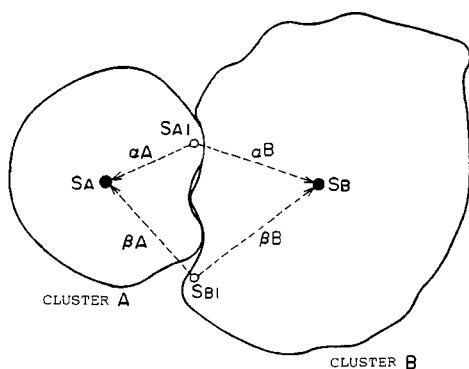
This adaptive speech recognizer performs two types of signal adaptation before making a final classification for recognition. Based on a preliminary acoustical analysis, the first adaptation estimates an environmental background contribution. After this estimate is removed, a trial comparison with the stored reference patterns provides the information needed to perform a second adaptation to estimate and remove speaker characteristics. A final comparison with the reference patterns then provides a final word classification.—DLR

5,890,114

43.72.Ne METHOD AND APPARATUS FOR TRAINING HIDDEN MARKOV MODEL

Jie Yi, assignor to Oki Electric Industry Company, Limited
30 March 1999 (Class 704/256); filed in Japan 23 July 1996

The patent describes a two-step method of training a hidden Markov model (HMM) reference for use in a speech recognizer. A preliminary speech analysis is first performed using known acoustical analysis techniques with a limited amount of training speech data. HMMs are computed for each observed phoneme. Centroids are then computed by clustering the



SA CENTROID STATE OF CLUSTER A
SB CENTROID STATE OF CLUSTER B
SAI STATE BELONGING TO CLUSTER A
SBI STATE BELONGING TO CLUSTER B

-----> STATE SHARING RELATIONSHIP

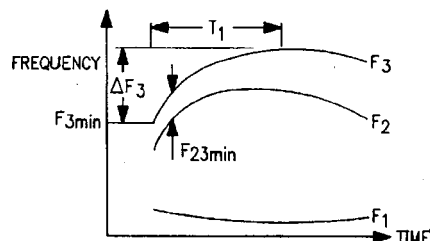
states of the HMMs of similar phonemes. A second speech analysis is then performed and the resulting phoneme vectors used to select the best centroid. The performance is said to be comparable to systems using diphone or triphone models.—DLR

5,893,058

43.72.Ne SPEECH RECOGNITION METHOD AND APPARATUS FOR RECOGNIZING PHONEMES USING A PLURALITY OF SPEECH ANALYZING AND RECOGNIZING METHODS FOR EACH KIND OF PHONEME

Tetsuo Kosaka, assignor to Canon Kabushiki Kaisha
6 April 1999 (Class 704/254); filed in Japan 24 January 1989

This speech recognition system uses a variety of classification techniques to perform phoneme class assignments. A preliminary classification is done to select a method to be used for a more detailed classification. Available methods include a neural network for vowel phonemes, dynamic



programming for semivowels, and a Bayesian decision tree method for other phonemes. Depending on the results of the specialized classifier, an acoustical vector may be rerouted back through the classification process.—DLR

5,895,447

43.72.Ne SPEECH RECOGNITION USING THRESHOLDED SPEAKER CLASS MODEL SELECTION OR MODEL ADAPTATION

Abraham Poovakunnel Ittycheriah and Stephane Herman Maes, assignors to International Business Machines Corporation
20 April 1999 (Class 704/231); filed 2 February 1996

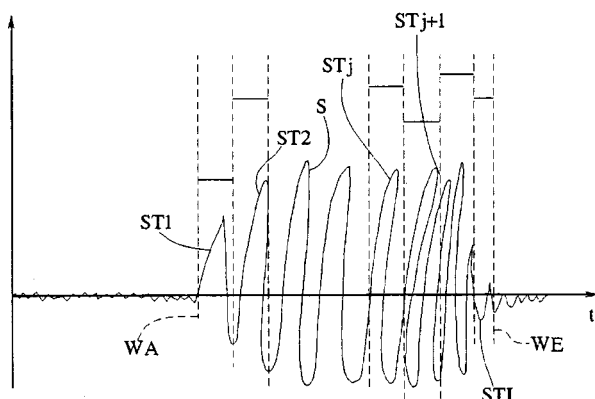
This speech recognition system constructs models of groups of speakers by doing a cluster analysis of the acoustical feature vectors. As a speaker class is identified from the clustering, a class codebook is constructed which may be used in a speaker-independent manner by speakers within the class. Speech from any speaker within the identified class may be used to train a class model.—DLR

5,899,971

43.72.Ne COMPUTER UNIT FOR SPEECH RECOGNITION AND METHOD FOR COMPUTER-SUPPORTED IMAGING OF A DIGITALIZED VOICE SIGNAL ONTO PHONEMES

Luc De Vos, assignor to Siemens Aktiengesellschaft
4 May 1999 (Class 704/231); filed in Germany 19 March 1996

The device described in this patent represents an attempt to gain the pattern storage efficiencies of phoneme-oriented methods used in speaker-independent recognizers, while pursuing a speaker-dependent training strategy. The idea is that words spoken by a new system enrollee are automati-



cally segmented into the appropriate phoneme sequences and then stored as phonemic reference patterns, rather than the whole-word patterns typical of speaker-dependent systems. Several well-known analysis strategies are cited in proposing how the task might be done.—DLR

5,899,973

43.72.Ne METHOD AND APPARATUS FOR ADAPTING THE LANGUAGE MODEL'S SIZE IN A SPEECH RECOGNITION SYSTEM

Upali Bandara *et al.*, assignors to International Business Machines Corporation
4 May 1999 (Class 704/256); filed 4 November 1995

The speech recognizer described in this patent uses an n-gram type of language model, rather than a syntactic constraint model. The problem cited is that trigram phoneme models require large amounts of memory. In order to reduce that memory usage, a typical strategy is to eliminate some of the trigrams. The patent asserts that in prior known systems, some of the rarely used trigrams are discarded. In this system, rare trigrams are retained, while commonly used ones are discarded, filling the gap with bi-gram representations, which are more easily trained due to the greater abundance of the corresponding speech sounds.—DLR

5,899,976

43.72.Ne METHOD AND SYSTEM FOR BUFFERING RECOGNIZED WORDS DURING SPEECH RECOGNITION

Michael J. Rozak, assignor to Microsoft Corporation
4 May 1999 (Class 704/270); filed 31 October 1996

The patent describes a number of strategies for performing corrections to text items entered by dictation into a recognizer-equipped text processor. Both the recognizer and the dictation buffering systems are accessible by the word processor so that, for example, error word lists may easily be displayed. The assumption is that the user will find it convenient to use mouse actions and keystrokes to perform any functions needed in the process of correcting recognition or dictation errors.—DLR

5,903,864

43.72.Ne SPEECH RECOGNITION

Gregory J. Gadbois and Stijn A. Van Even, assignors to Dragon Systems
11 May 1999 (Class 704/251); filed 30 August 1995

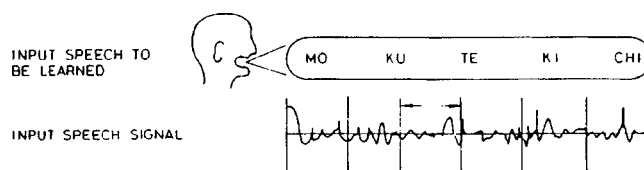
The patent describes a strategy for recognizing a pair of related utterances. In the particular case cited, the first utterance consists of a zip code and state name read from a shipping label during package sorting operations. The second utterance consists of the city name. Using a zip code database, the two utterances are alternately reprocessed as needed to find the most likely interpretation of the pair of utterances.—DLR

5,903,865

43.72.Ne METHOD OF PREPARING SPEECH MODEL AND SPEECH RECOGNITION APPARATUS USING THIS METHOD

Shunsuke Ishimitsu and Ikuo Fujita, assignors to Pioneer Electronic Corporation
11 May 1999 (Class 704/256); filed in Japan 14 September 1995

This patent describes a method for rapidly training a hidden Markov (HMM) speech model for use in a speech recognition system. When the operator signals that recognition of an input speech item has failed, the system creates new HMM sequences for the failed sequence of input speech



features. By copying from existing similar HMMs and merely updating the probability values, the new model can be built from as few as one or two utterances of the new item.—DLR

5,905,773

43.72.Ne APPARATUS AND METHOD FOR REDUCING SPEECH RECOGNITION VOCABULARY PERPLEXITY AND DYNAMICALLY SELECTING ACOUSTIC MODELS

Chi Wong, assignor to Northern Telecom Limited
18 May 1999 (Class 379/88.03); filed 28 March 1996

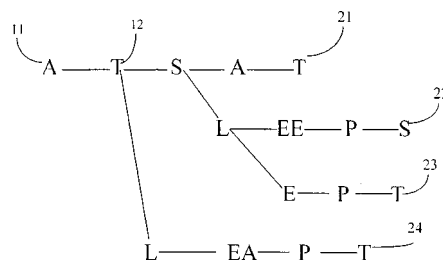
Speech recognition technology has still not progressed to the point of providing service comparable to a human telephone operator in a situation involving hundreds or thousands of phone subscribers. The system presented here improves the success of the recognizer by limiting the number of names through which the system must search. This is done by using additional location information to select a subset of subscribers, such as the set of clients of a particular private switching network.—DLR

5,905,971

43.72.Ne AUTOMATIC SPEECH RECOGNITION

Simon Alexander Hovell, assignor to British Telecommunications public limited company
18 May 1999 (Class 704/257); filed in the United Kingdom 3 May 1996

Most speech recognition systems produce some sort of network of possible sequences of phonetic elements, which are compared with the input speech features as recognition proceeds. This patent describes a system in which the network elements are initially kept in a symbolic form, allowing



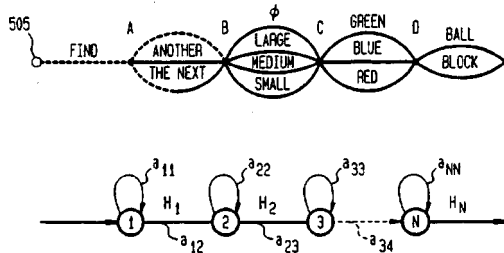
the greatest generality of the represented phonetic sequences. Only when a match is attempted for a particular input sequence is the symbolic element replaced by a triphone phonetic element, which captures phonetic characteristics of the neighboring elements.—DLR

5,907,634

43.72.Ne LARGE VOCABULARY CONNECTED SPEECH RECOGNITION SYSTEM AND METHOD OF LANGUAGE REPRESENTATION USING EVOLUTIONAL GRAMMAR TO REPRESENT CONTEXT FREE GRAMMARS

Michael Kenneth Brown and Stephen Charles Glinski, assignors to AT&T Corporation
25 May 1999 (Class 382/226); filed 21 January 1994

This speech recognition system reduces the amount of memory required for storing the grammar and phonetic matching models by dynamically constructing only the needed portions of the phonetic match network and discarding portions as soon as possible. The phonetic models are typical



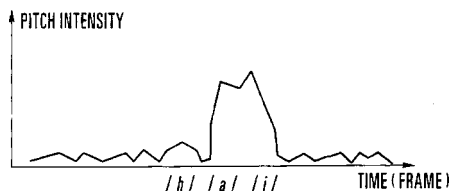
mixture density hidden Markov models and the grammar seems to be a context-free tree structure from which local word sequence constraints are extracted.—DLR

5,907,826

43.72.Ne SPEAKER-INDEPENDENT SPEECH RECOGNITION USING VOWEL/ CONSONANT SEGMENTATION BASED ON PITCH INTENSITY VALUES

Keizaburo Takagi, assignor to NEC Corporation
25 May 1999 (Class 704/251); filed in Japan 28 October 1996

A speech recognizer is described in which a pitch analysis plays a major role in segmentation of the input feature sequence. Going beyond the typical voiced/unvoiced decision as a binary value, in this patented device,



the amplitude of the pitch component is treated with at least as much importance as any spectral feature.—DLR

5,909,662

43.72.Ne SPEECH PROCESSING CODER, DECODER AND COMMAND RECOGNIZER

Yasushi Yamazaki *et al.*, assignors to Fujitsu Limited
1 June 1999 (Class 704/221); filed in Japan 11 August 1995

Speech processors optimized for any one of coding, recognition, or synthesis typically use different internal speech representations than are used for the other purposes. The device described here sacrifices a little and tweaks a little and provides all of the functions using a single representation. A code-excited cepstrum analysis uses a stochastic codebook. Excitation

codes double as phonetic elements in the dictionary storage of words. Coding and resynthesis are thus fairly typical. There is little further elaboration of recognition strategies.—DLR

5,909,665

43.72.Ne SPEECH RECOGNITION SYSTEM

Yasuko Kato, assignor to NEC Corporation
1 June 1999 (Class 704/241); filed in Japan 30 May 1996

The patent describes a method of reducing the computational load of computing the distances between phonetic features of the input speech and the stored reference patterns. The method disclosed is the well-known technique of using a preliminary distance computation to select a subset of the standard patterns. Full distance computations are then done for only those reference patterns in the subset. A sentence in the Summary section makes the curious statement that the "invention" is "unique and obvious."—DLR

5,909,666

43.72.Ne SPEECH RECOGNITION SYSTEM WHICH CREATES ACOUSTIC MODELS BY CONCATENATING ACOUSTIC MODELS OF INDIVIDUAL WORDS

Joel M. Gould *et al.*, assignors to Dragon Systems, Incorporated
1 June 1999 (Class 704/251); filed 13 November 1992

The speech recognition system described here allows the user to add new phrases to the recognition vocabulary by treating entire phrases as "compound" words, with models built up from existing word phonetic models. The idea would seem to apply only to a recognizer such as a "template matching" or speaker-dependent device, which treats words as fundamental units rather than being built up in turn from phonemes. However, the patent does discuss the possibility that words may be represented by smaller units.—DLR

5,895,448

43.72.Pf METHODS AND APPARATUS FOR GENERATING AND USING SPEAKER INDEPENDENT GARBAGE MODELS FOR SPEAKER DEPENDENT SPEECH RECOGNITION PURPOSE

George J. Vysotsky and Vijay R. Raman, assignors to Nynex Science and Technology, Incorporated
20 April 1999 (Class 704/251); filed 29 February 1996

Various speech recognition systems have been described in which multiple recognizers operate simultaneously, searching for alternate analyses of an input word or phrase. In some cases, the alternate analyzers have included both speaker-dependent and speaker-independent systems. In the patented system, a word which is considered to be an allowable recognition target in a speaker-independent sub-system may also be used and considered as a garbage word (out of vocabulary) by a speaker-dependent sub-system. This is said to reduce the chance of contradictory recognition results from two different sub-systems.—DLR

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

**Dimitri Kanevsky and Stephane Herman Maes, assignors to
International Business Machines Corporation
27 April 1999 (Class 704/246); filed 11 June 1997**

This system for allowing user access to secure services or facilities makes exclusive use of voice communication for the evaluation. The system uses speaker voice characteristics as well as privately known information to determine whether access should be permitted. For example, the system may ask the entry applicant to recite certain details which are available from the database and presumably known only by the person the applicant claims to be.—DLR

**43.72.Pf SPEAKER IDENTIFICATION WITH USER-
SELECTED PASSWORD PHRASES**

**Sarangarajan Parthasarathy and Aaron Edward Rosenberg,
assignors to AT&T Corporation
15 June 1999 (Class 704/256); filed 22 August 1997**

This speaker identification system performs three sequential steps to improve the reliability of identification. An applicant first speaks a user-specific password. A speaker-independent phonetic recognition task first searches a vocabulary of passwords to determine the spoken phonetic sequence. Each of N best matches are then looked up in a user database to determine N potential speaker IDs. The input is then reexamined using a speaker-dependent recognizer for each of the N potential speakers. The scores from each of the first two steps must then meet validation criteria in a final verification step.—DLR

Synthetic array measurements of acoustical waves propagating into a water-saturated sandy bottom for a smoothed and a roughened interface

Harry J. Simpson^{a)} and Brian H. Houston

Physical Acoustics Branch Code 7136, Naval Research Laboratory, Washington, D.C. 20375

(Received 10 March 1998; revised 30 November 1999; accepted 15 December 1999)

Measurements of acoustical waves propagating in an unconsolidated water-saturated porous medium that involved the use of a two-dimensional synthetic array technique are presented. From these measurements, wave-front propagation is evaluated using time domain, wave-number frequency and two-dimensional spatial analysis techniques. First, a synthetic array measurement is employed to study sound penetration for a “smoothed” interface condition. This smoothed interface measurement is then compared to an identical synthetic array measurement with a “roughened” interface. A comparison of the smoothed and roughened interface measurements shows enhanced bottom penetration at shallow grazing angles as a result of the roughened interface. At shallow grazing angles, this scattered pressure from the roughened interface has the appearance of a wave front in the water-saturated porous medium that has a virtual wave speed of ~ 1200 m/s. Slower compressional waves in the sandy bottom are not observed. [S0001-4966(00)04604-X]

PACS numbers: 43.20.Jr, 43.30.Hw [DLB]

INTRODUCTION

Understanding the physics of sound propagation into a sandy bottom of a shallow water environment is driven by a need to improve techniques for finding and classifying objects that are buried. For sandy bottoms, the wavespeeds tend to be slightly greater than that of the water.¹⁻³ Therefore, from a simple ray model there exists a critical angle below which the acoustical waves are totally reflected and no sound propagates into the sandy bottom. However, recent measurements by Chotiros⁴ have indicated that sound penetration is possible for shallow grazing angles. To explain this enhanced bottom penetration, Chotiros has modeled the bottom as a Biot medium. With the Biot model, the traditional fast compressional wave does have a critical angle, but there exists a second, slower, compressional wave that does not have a critical angle since its wavespeed is less than that of the water.

Thorsos *et al.*⁵ and Moe⁶ have described enhanced bottom penetration using a mechanism of scattering from relatively small amplitudes of roughness [~ 4 mm rms (root-mean-square) at 20 kHz] at the water-sand interface. From this interface roughness model, enhanced bottom penetration at shallow grazing incidence can be described with the more traditional fast compressional wave speed that follows a different ray path. This second wave field is created when the incident wave field scatters from roughened areas of the water-sand interface into the sandy bottom. Williams *et al.*⁷ modeled this roughened interface bistatic scattering into the sand with a Kirchhoff approximation near the refracted ray direction and with an interface and volume inhomogeneity perturbation theory for the off refracted ray directions.

In the work reported here, measurements of acoustical waves in an unconsolidated water-saturated porous medium

have been conducted in a pristine laboratory environment at the Naval Research Laboratory that employed two-dimensional synthetic array techniques. Unlike many of the previous measurements taken in open ocean conditions, this “shallow water” laboratory provides an environment where the experimental parameters are controlled and the measurements are repeatable and stable. To better simulate the generic shallow water environment, a sand was chosen that falls in the mid-range of sands found in coastal conditions.⁸⁻¹⁰ However, to eliminate the introduction of biologics and to ensure control over material properties and grain size, a manufactured sand with a 240 μm mean diameter and a standard deviation of 71 μm was used for the sandy bottom.

Acoustical sound speed measurements were performed using a nontraditional buried hydrophone technique. Rather than building an array of phase matched hydrophones and inserting them into the bottom, the method chosen synthesized a buried hydrophone array by using only a single receiver. This was done by burying an omnidirectional hydrophone (B&K 8103) to its deepest point, 90 cm for these measurements. The source was subsequently moved to discrete locations parallel to the interface, which incrementally mapped out the field in the sand at that receiver depth. Next, the hydrophone was moved vertically up one step, 1.5 cm for these measurements, and the source was again moved through the same horizontal positions. By repeating this process, the wave fields in the sand were mapped out directly. However, great care must be put into a robotics system to insure repeatable positioning throughout the measurement. The robotics system employed in these measurements is repeatable to 50 μm and has been leveled and aligned to the sandy bottom.

Using this synthetic array technique, it was possible to very accurately map out the pressure in the sandy bottom and

^{a)}Electronic mail: simpson@maddog.nrl.navy.mil

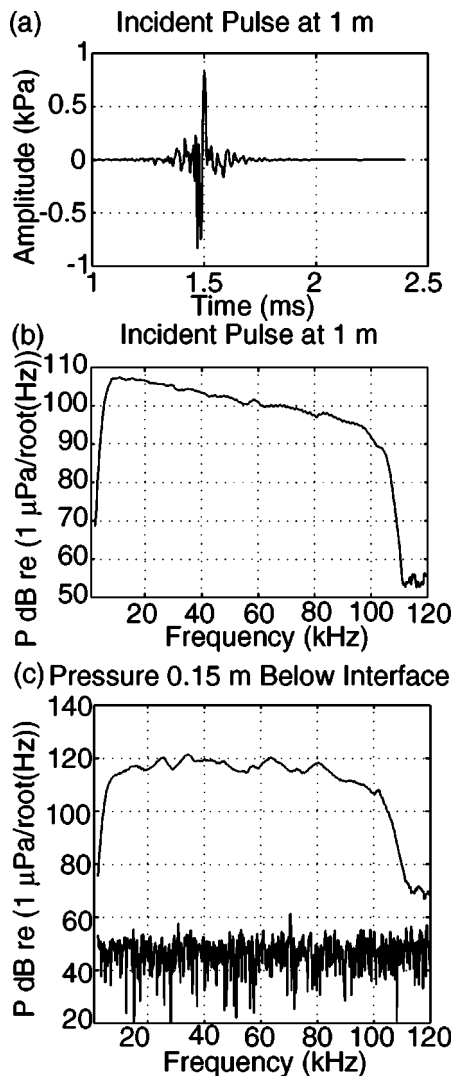


FIG. 3. (a) is the source-receiver temporal response at a 1.0 m separation distance using a shaped impulse for the excitation. (b) is the source-receiver frequency response at a 1 m separation distance. (c) is the measured frequency response of the buried hydrophone 15 cm below the water-sand interface. Also shown in (c) is the noise floor for the measurement, which shows the measurement has a typical signal-to-noise of 70 dB across the measurement bandwidth.

15 cm below the sand and the source directly above, the signal and noise of the measurement is shown in Fig. 3(c). For this typical measurement, the SNR was nominally 70 dB over the bandwidth of the measurement.

The sand used in these measurements was purchased from a commercial manufacturing company, Unimin, and consisted of 300 tons of sifted and washed sand. The sand was produced via a crushing process and then sized using sieves. The manufacture uses a name designation of “5030,” which indicates that 30% of the sand is larger than the #50 sieve, which contains a 300 μm mesh. Figure 4 is a plot of the sieve analysis showing the distribution of the sand installed in the laboratory. Statistical analysis gives a mean grain diameter of 240 μm and a standard deviation of 71 μm . Unimin Corporation provided these measurements of the grain size along with the grain bulk lattice density and the chemical analysis of the sand, shown in Table I. In ad-

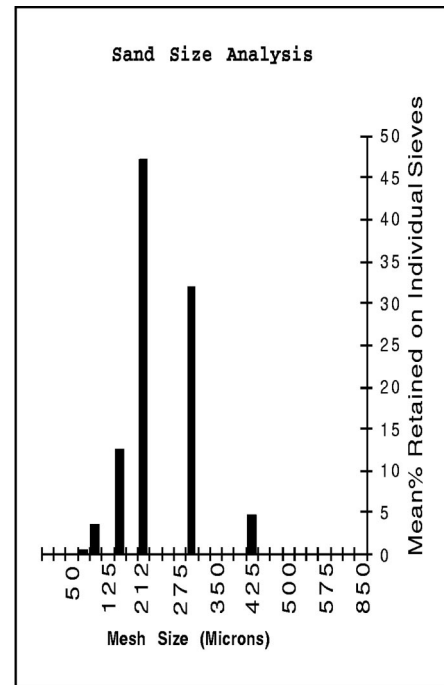


FIG. 4. The sieve analysis of the manufactured sand installed in the shallow-water measurement laboratory at the U.S. Naval Research Laboratory. The mean diameter of the sand is 240 μm . The sieve analysis is provided by the manufacture, Unimin, for this batch of sand. The vertical axis is the percent of sand retained on a specific sieve and the horizontal axis is the size of the sieve mesh.

dition, the authors conducted measurements to determined porosity and permeability, shown in Table I.

II. MEASUREMENTS

Using the Cartesian robotics system to position the spherical source and a vertical stage to position a hydrophone buried in the sand, unique measurements of the sound propagating in the sand were made. The hydrophone was initially buried 90 cm below the water-sand interface. It was inserted using a long 5 cm diam stainless steel pole that was pushed into the sandy bottom, the pole was then removed and the sand-water interface was either smoothed or roughened. A water jet was not necessary for these tank experiments. The source was then moved in discrete steps from directly above the hydrophone to 3.3 m away by 1.5 cm steps. After each sweep of the source, the hydrophone was pulled up 1.5 cm though the sand and then another track of the F42A source was measured for the new hydrophone position.

In this way, an array of buried hydrophones was synthesized using a single hydrophone. The use of a synthesized buried hydrophone array eliminates several of the problems associated with multiple hydrophone arrays, this includes

TABLE I. Measured physical parameters of the sand.

porosity	37.0%
grain density	2650 kg/m ³
permeability	3.98×10^{-11} m ²
chemistry	99.639% SiO ₂

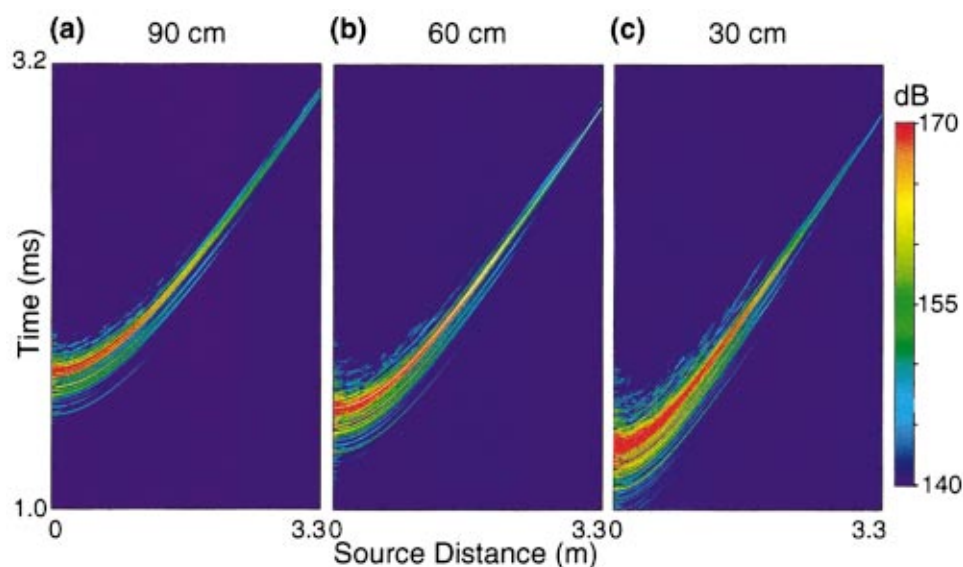


FIG. 5. The temporal pressure response of the buried hydrophone for 90, 60, and 30 cm below the water–sand interface as the source is moved horizontally from above the hydrophone to 3.3 m away. In (b), an overlay of the time delay associated with the fast wave is shown. The time delay is calculated assuming a 1482 m/s wave speed in the water and a 1680 m/s wave speed in the water-saturated sand. Simple refraction at the interface is assumed. The color scale represents the magnitude of the acoustic pressure in dB *re* 1 μ Pa.

problems associated with amplitude and phase matching, absolute positioning errors, and cross coupling between elements. With the synthesized array it becomes possible to investigate the wave fronts propagating within the sandy bottom directly and lends itself to analytical techniques such as wavefront analysis, wave-number frequency analysis and two-dimensional spatial analysis.

In addition to the above, a second preliminary study of the effect of interface roughness on sound penetration into the bottom were made. Both the smooth interface and the rough interface studies are presented in this section, and the results of the investigation are discussed in the following section.

A. Smoothed water–sand interface measurements

The measurement of the sound propagating through the smoothed sandy bottom was taken during a single computer controlled acquisition that lasted approximately five days. Presentation of the data in this paper will initially include the more conventional temporal analysis that is presented in the literature.^{4,11} Secondly, the wavespeeds of the vertical wave fronts through the medium are decomposed using wave-number frequency techniques. Finally, the two-dimensional synthetic array data is analyzed using two-dimensional spatial techniques.

1. Temporal response for specific hydrophone locations

The temporal response plots for the hydrophone buried at 90, 60, and 30 cm below the water–sand interface are shown in Fig. 5(a)–(c), respectively. The absolute value of the pressure in dB *re* 1 μ Pa, $20 \log_{10}|\text{Pressure}|$, is plotted. In Fig. 5(a), the peak of the incident pulse is seen at ~ 1.7 ms for the hydrophone 90 cm below the water–sand interface and the source directly above the hydrophone. As the source is moved from above the hydrophone to 3.3 m away, the pulse takes a progressively longer path to arrive at the hydrophone due to the increasing distance. The data has been

temporally windowed to remove the reverberation from the boundaries of the measurement facility, the water–air interface, the side-walls and the back-wall.

For the hydrophone at a 60 cm depth, Fig. 5(b), the incident pulse is seen at ~ 1.5 ms or the source directly above, and again at later times as the source is moved horizontally to 3.3 m away. The overlay in Fig. 5(b) (solid white line) is the time delay predicted for a ray path with the sound propagating through the water, across the water–sand interface, and then through the water-saturated sand. This ray path assumes simple refraction at the water–sand interface and also assumes a fluid wave speed in the water and sand of 1482 and 1680 m/s, respectively. An 805 μ s delay is included in the ray path calculation to account for the time between the start of the measurement and the peak of the excitation pulse.

2. Line array measurements

The next step in the smooth interface measurement data analysis is to study the frequency response of the synthetic array. To do this, the temporal data records are Fourier transformed and normalized using the 1 m free field acoustical response shown in Fig. 3(b). This resulting transfer function approach is used throughout the rest of the analysis. The first array to be investigated is the 60 measurement points directly below the source transducer. The vertical array occupies buried locations from 90 to 1.5 cm below the water–sand interface (with steps of 1.5 cm). A spatial Fourier transform of this array of data yields the dispersion curves of the waves propagating vertically into the sand, directly below the source, shown in Fig. 6. There is only one dominant wave observed in this measurement, which is propagating at 1680 m/s. Overlays for the 1680 m/s wave speed of the water-saturated sand (solid line), the 1482 m/s wave speed of pure water (dot–dashed line), and the 1200 m/s wave speed predicted by Chotiros⁴ (dashed line) are shown in Fig. 6. There is no indication of a slower, 1200 m/s wave propagating through the water-saturated porous medium.

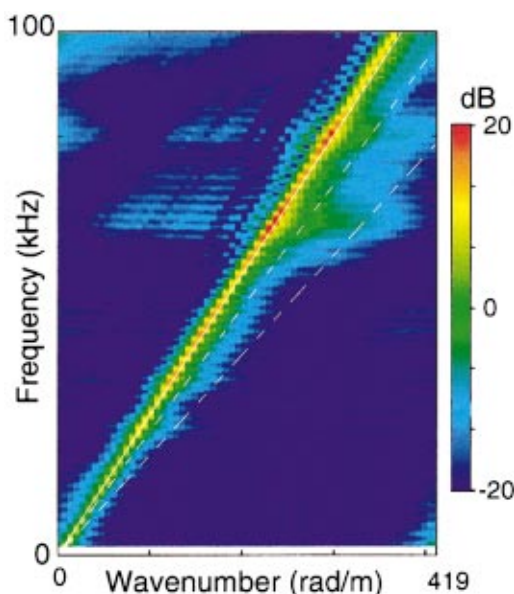


FIG. 6. The dispersion curve for a synthetic vertical array directly below the source. The horizontal axis is the wave number in radians per meter and the vertical axis is frequency in kHz. The predominant feature is the nearly straight line running diagonally up through the figure. Overlaid on this dispersion curve is a solid white line that represents a 1680 m/s nondispersive wave speed. For reference, the wave speed of water, 1482 m/s, is also shown below the fast wave with the dot-dashed line. In addition, a slower 1200 m/s wavespeed is shown on the figure, the dashed line.

It is also possible to extract the attenuation of the wave field as it propagates through the sand using these synthetic array measurements. Seven lines of propagation through the bottom were identified and an averaged attenuation in the bottom is calculated. For each path through the sand, the difference of the pressure between subsequent array points is averaged along the length of the line for each frequency bin of the measurement. The attenuation as a function of frequency for the seven ray paths through the sand were made and then an average attenuation from these seven lines is calculated and shown in Fig. 7. Radial spreading of the wave front is accounted for in the calculation of the attenuation. Also, the broadband standard deviation for the seven lines is 1.9 dB/m. The attenuation shown in Fig. 7 compares favorably with the attenuations shown by Chotiros,⁴ with our results laying in the middle of the attenuations curves and data shown in Fig. 6 of Ref. 4.

3. Two-dimensional synthetic array measurements

The synthetic two-dimensional array allows direct observation of the wave fronts propagating in the fluid saturated porous medium. Figure 8(a) shows the wave field propagating in the sand at 5 kHz, where the horizontal axis is the horizontal source distance starting at 0 m, for the source above the hydrophone, and moving to 3.3 m away. The vertical axis is the hydrophone depth starting at 90 cm below the sand at the bottom of Fig. 8(a) and moving to 1.5 cm below the sand at the top of Fig. 8(a). Figure 8(a) is a plot of the real part of the pressure field, which shows the phase of the wave fronts propagating through the sand [$\text{sign}(P) \times 20 \log|P|$]. Close inspection of the wave fronts shows spherical spreading, as expected from the spherical source.

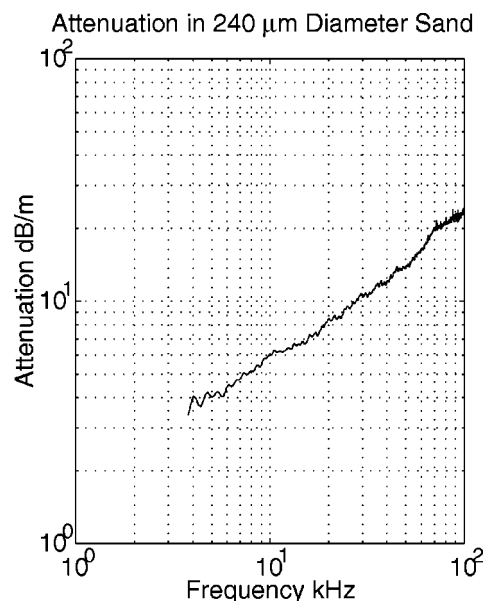


FIG. 7. The measured averaged attenuation from several synthetic line arrays. The attenuation here is calculated by averaging the difference between each adjacent point in a specific synthetic line array. The loss due to spherical spreading along each synthetic line array has been removed from the calculated attenuation. The attenuation along the several synthetic line array are then averaged.

However, near the interface it is clear that the wave fronts are moving with a slower wave speed (shorter wavelength) as compared to the wave fronts a few centimeters deeper into the sandy bottom. For a simple two-fluid model,^{12,13} there exists some finite width region of the interface between the water and the sand where the sound speed transitions from that in the water, 1482 m/s, to that in the bulk water-saturated sand, 1680 m/s. This is due to the boundary condition requirement that the horizontal wave-numbers match at the sand-water interface.

With the water and sand sound speeds of 1482 and 1680 m/s, respectively, there exists a critical angle at $\theta=28.1^\circ$ measured up from the interface. (Note that the horizontal and vertical axes are not shown with the same scale.) For the source ~ 94 cm horizontally from the hydrophone and 50 cm above the interface, the incident wave at and past the hydrophone entry point is totally reflected from the interface. However, since a spherical source was used, some energy still propagates to the hydrophone through the sandy bottom via a simple refractive ray path. At the lower frequencies, the attenuation in the sandy bottom is lower and results in a significant amount of energy reaching the hydrophone even at a horizontal source distance of 3.3 m.

Figure 8(b) shows the real part of the pressure field for 50 kHz. Again, the spherical wave front is visible. In addition, this higher-frequency component has a higher attenuation and thus does not propagate as far through the sand when compared to the energy at the lower frequencies. Again the source reaches the critical angle about 94 cm away from the hydrophone, past which the incident pressure wave is totally reflected from the water-sand interface. The right hand region of Fig. 8(b), where the source is further than 2 m from the hydrophone, has a pressure field that is below the 60 dB cutoff of the color scale, except near the interface. At

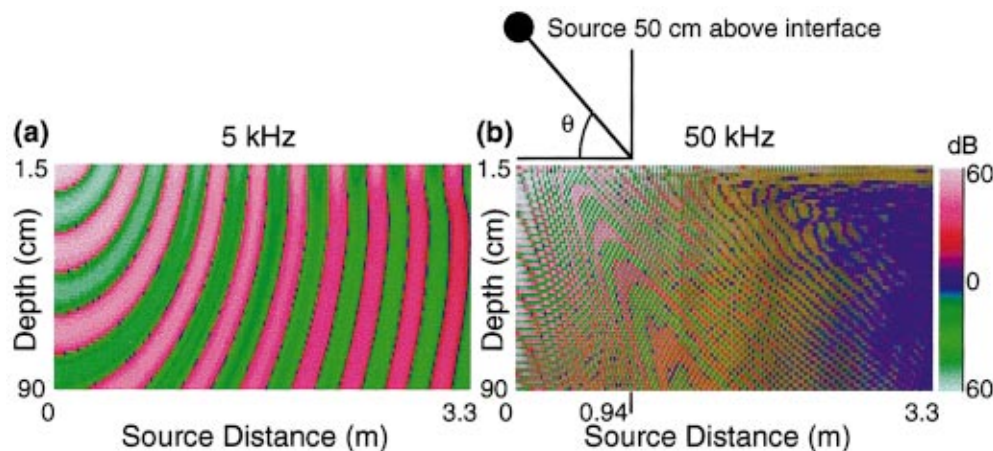


FIG. 8. The real part of the measured pressure field at 5 and 50 kHz for a smoothed interface. (a) Shows a relative 60 dB scale of the real part of pressure field measured by the synthetic array at 5 kHz. The sign of the pressure is preserved and put back into the data after the logarithmic processing to provide phase information. The horizontal and vertical axis are not scaled the same in this figure. (b) Is the same as (a) except it includes the measurement data in the 50 kHz bin.

the greater source distances, greater than 2 m, the interface wave field (or evanescent wave field) is clearly evident, and little penetration of the incident field occurs below the first few centimeters of the water–sand interface. In the two-dimensional space plots of Fig. 8, there is evidence of a single dominant wave speed. From the vertical array dispersion curves, this wave speed has been identified as a 1680 m/s wave.

B. Roughened water–sand interface measurements

To understand the effect of the water–sand interface roughness on the sound penetrating into the sand, the previously smoothed water–sand interface was intentionally roughened and an identical two-dimensional synthetic array measurement was repeated. The interface was roughened with a 5 cm diam pole that was drawn across the interface in a “snake” like pattern. A square patch of ~ 2 m by ~ 2 m

was roughened, centered at the hydrophone insertion point. A visual estimate of the rms interface roughness was between 1 and 2 cm. Subsequent ultrasonic measurements confirmed that the roughening technique used for these measurements produced a 1 cm rms interface roughness.

1. Temporal response for specific hydrophone locations

As with the smoothed interface measurements, the roughened interface measurement included 221 horizontal source positions and 60 vertical hydrophone positions buried in the sand. Figure 9 shows the temporal pressure response of the hydrophone at specific depths, 90, 60, and 30 cm, as the source is moved from directly above the hydrophone to 3.3 m horizontally away. The three hydrophone depths in Fig. 9 are the same as those for the smoothed interface data

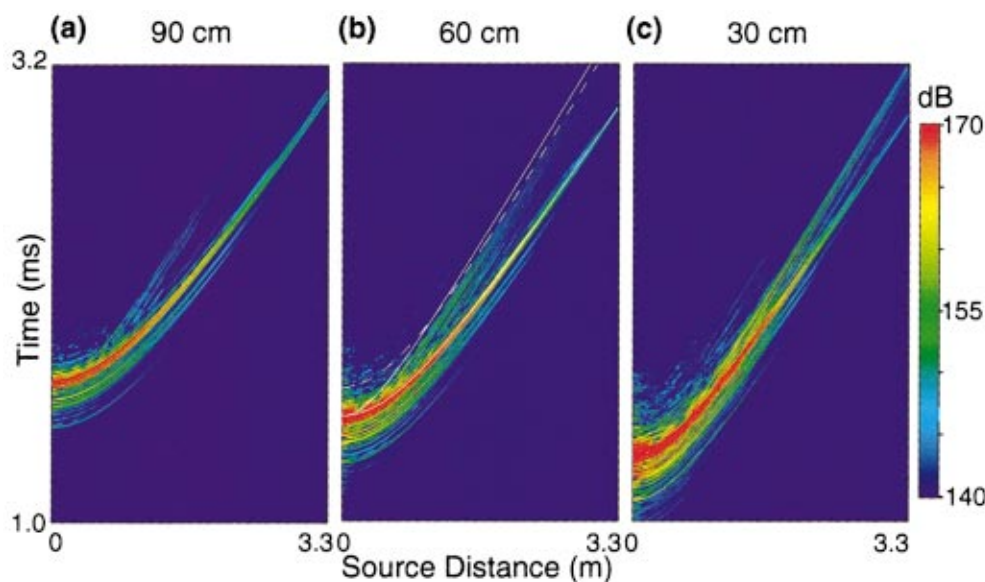


FIG. 9. The temporal pressure response of the buried hydrophone for 90, 60, and 30 cm below the roughened water–sand interface as the source is moved horizontally from above the hydrophone to 3.3 m away. The earliest overlay, indicated by the solid line, is the time delay associated with a ray path through the water at 1482 m/s, simple refraction at the water–sand interface and a bottom wave speed of 1680 m/s. The dashed line represents the time delay for a ray path through the water at 1482 m/s, simple refraction at the interface and a bottom wave speed of 1200 m/s. The other solid line in Fig. 10(b) is the time delay associated with a ray path through the water to the insertion point, which is located vertically above the hydrophone, then scattered into the sand at the water–sand interface and then propagates vertically downward to the hydrophone. The color scale represents the magnitude of the acoustic pressure in dB *re* 1 μ Pa.

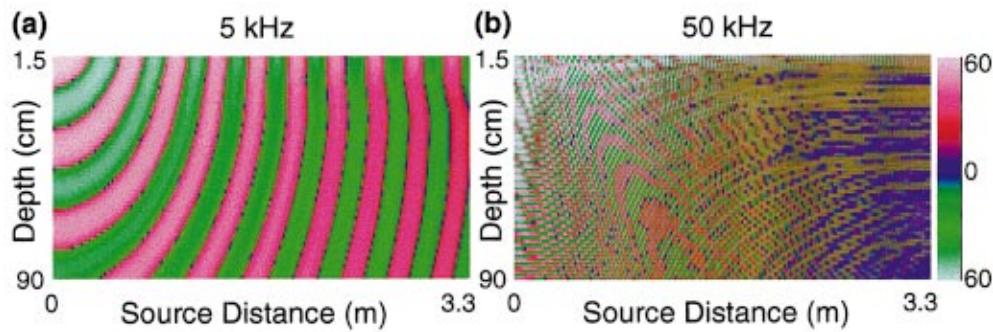


FIG. 10. The real part of the measured pressure field at 5 and 50 kHz for a roughened interface. (a) Shows a relative 60 dB scale of the real part of pressure field measured by the synthetic array at 5 kHz. The sign of the pressure is preserved and put back into the data after the logarithmic processing to provide phase information. The horizontal and vertical axis are not scaled the same in this figure. (b) Is the same as (a) except it includes the measurement data in the 50 kHz bin.

presented in Fig. 5. In the discussion section a direct comparison of the two measurements will be made.

As with the smoothed interface data, the temporal response for the roughened interface is dominated by the incident fast wave traveling at 1680 m/s, especially for source angles above the critical angle. The fast wave is the earliest arrival to the hydrophone and again is modeled with a simple ray path through the water and then the sand. The time delay associated with this ray path is indicated by the earliest overlay in Fig. 9(b) (solid white line) and assumes a water speed of 1482 m/s and a water-saturated sand wave speed of 1680 m/s. For comparison, the time delay associated with a water speed of 1482 m/s and a slower compressional ("second Biot"⁴) wave speed of 1200 m/s in the sandy bottom is also shown (dashed line). A third, later time overlay (solid white line) is for the time delay associated with a ray path of 1482 m/s in water to the sand directly above the hydrophone and then a vertical ray path through the sand at 1680 m/s.

These temporal responses will be discussed in more detail in the discussion section. However, it is clear that there are later time arrivals in the temporal response of the rough-

ened interface measurements that are not present for the smoothed interface measurements. These differences can only be attributed to the change in the topography of the water-sand interface.

2. Two-dimensional synthetic array measurements

Using the same data processing techniques discussed for the smoothed interface measurements, similar transfer functions for the roughened interface measurements were produced. Wave-front analysis for 5 kHz, shown in Fig. 10(a), again shows the radial spreading of the wave front as it propagates through the sand. At 5 kHz these measurements show no enhancement of energy into the bottom due to the roughened interface. However, for 50 kHz, shown in Fig. 10(b), there is a noticeable enhancement of sound penetration at shallow grazing incidence as compared to the smoothed interface measurement shown in Fig. 8(b). This enhancement is typical of frequencies above 20 kHz for these measurements.

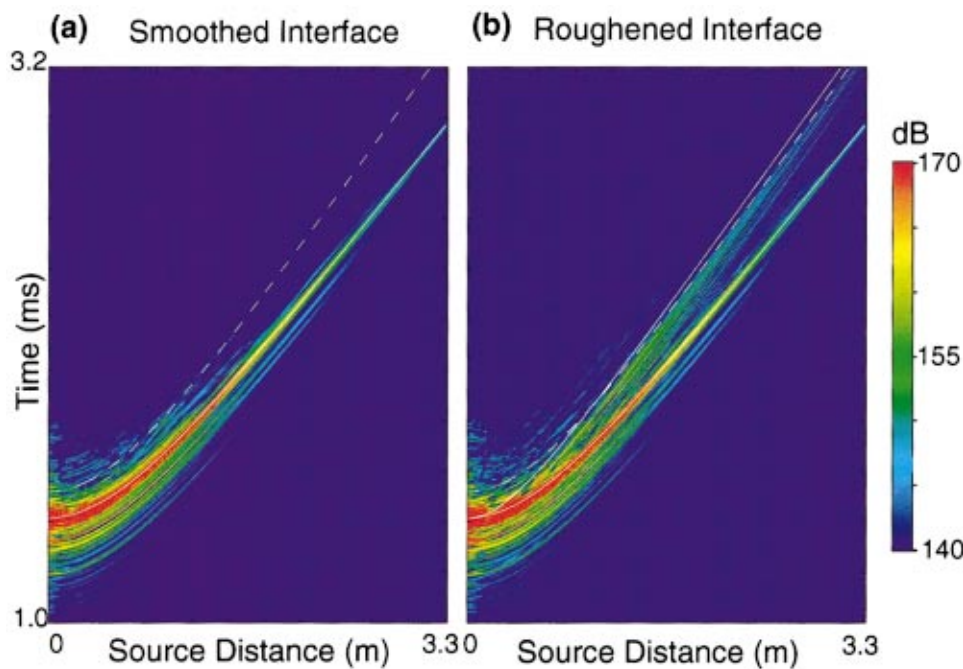


FIG. 11. Comparison of the temporal response for the hydrophone at 45 cm below the water-sand interface. Both data sets include two overlays of the time delays associated with ray paths assuming simple refraction at the interface, 1482 m/s water wave speed and a 1680 m/s and a 1200 m/s sand wave speed are shown. The 1680 m/s time delay is represented with the solid line and the 1200 m/s ray path is shown with the dashed line in both plots. Additionally in the roughened interface data of (b), a third line is shown for a ray path that travels in the water to the insertion point of the hydrophone and is then scattered straight down through the sand to the hydrophone. The water wave speed is 1482 m/s and the sand wave speed is assumed to be 1680 m/s.

III. DISCUSSION

Figure 11 shows a comparison between the temporal responses of the sound penetrating through the smoothed and roughened interfaces for the case of a hydrophone buried 45 cm below the interface. Here, the source is moved horizontally from directly above the hydrophone to 3.3 m away. Again the prominent feature in the temporal response is the fast wave, 1680 m/s. For the smoothed interface measurement, there is little evidence of a later, slower, time arrival. However, for the roughened interface measurement, in addition to the fast wave of 1680 m/s, there is a later time arrival that becomes more distinguishable at the shallower grazing angles. In the roughened interface measurement of Fig. 11(b), the earliest arrival is for an assumed ray path of a 1680 m/s sand wave speed and a 1482 m/s water wave speed, assuming simple refraction at the interface. This time delay as a function of source position is shown by the earliest overlay in Fig. 11(a) and (b) (solid white line). The later time arrivals associated with a roughened interface ray path that goes from the source to directly above the hydrophone at 1482 m/s and then is scattered into the sand, traveling vertically downward at 1680 m/s is shown in Fig. 11(b) (solid white line). This is one of many ray paths that are possible due to interface roughness scattering as proposed by Thorsos *et al.*,⁵ where a range of delay times can be expected due to scattering from the roughened interface. The last time arrival shown in both Fig. 11(a) and (b) (dashed white line) is for a ray path that assumes a sand wave speed of 1200 m/s and simple refraction at the interface, as proposed by Chotiros,⁴ based on a Biot model of the sandy bottom. For the roughened measurement, it is clear that the later time arrival is best described by an assumed ray path in the water to a point a few centimeters from the hydrophone insertion point, simple scattering from the roughened interface and then propagating nearly vertically to the hydrophone at 1680 m/s. In some regions of the diagram this can have the appearance of a wave traveling at the 1200 m/s. A more detailed discussion of how this comes about can be found in Refs. 5 and 6. The measurement data shown in Fig. 11(b) has enough temporal resolution to show that the later time arrival due to the interface roughness is collinear with the fast wave time trace at near normal incidence, and then separates from the 1680 m/s refractive ray path time trace with increasing source distance. For a slower wave of 1200 m/s, predicted by a Biot medium model,⁴ there would always be a time delay between the 1680 m/s and the 1200 m/s waves even at near normal incidence. This data does not support the Biot medium model predictions⁴ of a second slower compressional wave in this unconsolidated water saturated porous medium.

Figures 8(b) and 10(b) show the pressure at 50 kHz for the smoothed and roughened interfaces, respectively. For the smoothed interface condition, Fig. 8(b) shows that the pressure has attenuated more than 60 dB relative to the maximum pressure in the plot for the source distance of 3.3 m from the hydrophone. The critical angle for this measurement configuration is at 28.1° (for the 1680 m/s wave) measured up from the interface, diagrammed in Fig. 8(b), and occurs for the source ~ 94 cm from the hydrophone. Near the top of Fig. 8(b), past the critical angle, there is a clear tran-

sition region where an evanescent wave field is measurable, but this attenuates quickly with depth. However, unlike the highly attenuated region of the smoothed interface measurement where there is little sound penetration at the shallow grazing angles, the roughened interface measurement shows a measurable increase in the pressure in the shallow grazing angle region of Fig. 10(b). The increase in the sound penetrating into the sand past the critical angle is interpreted to be a result of increased water-sand interface roughness scattering.

IV. CONCLUSIONS

The temporal analysis, the wave-number frequency analysis and the two-dimensional spatial analysis shows conclusively that a 1680 m/s "fast" wave is propagating in this 240 μm manufactured water-saturated sandy bottom. Any other slower waves, either compressional or shear, are at least 60 dB smaller than the 1680 m/s wave at normal incidence. For the smoothed interface, there is no evidence in these measurements of a slower, 1200 m/s wave in the bottom, even at shallow grazing angles. However, at shallow grazing angles, the penetration of sound into the sandy bottom can be enhanced by roughening the interface. Scattering from an increased interface roughness near the hydrophone manifests itself in the temporal measurements as a later time arrival at shallow grazing angles that produces a virtual slow wave having a wave speed of 1200 m/s. However, for near normal incidence, these measurements have enough temporal resolution to clearly show that this later time arrival is better described by the roughened interface scattering, which give the appearance of a wave front with a virtual wave speed of 1200 m/s at shallow ensonification angles, rather than by a 1200 m/s slow compressional wave.

ACKNOWLEDGMENT

This work was supported by the Office of Naval Research.

- ¹K. L. Williams and L. J. Satkowiak, "Linear and parametric array transmission across a water-sand interface-Theory, experiment, and observations of beam displacement," *J. Acoust. Soc. Am.* **86**, 311-325 (1989).
- ²R. A. Altenburg, N. P. Chotiros, and C. M. Faulkner, "Plane-wave analysis of acoustical signals in a sandy sediment," *J. Acoust. Soc. Am.* **89**, 165-170 (1991).
- ³D. 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).
- ⁴N. P. Chotiros, "Biot model of sound propagation in water-saturated sand," *J. Acoust. Soc. Am.* **97**, 199-214 (1995).
- ⁵E. I. Thorsos, D. R. Jackson, J. E. Moe, and K. L. Williams, "Modeling of subcritical penetration into sediments due to interface roughness," Conference Proceedings "High Frequency acoustics in shallow water," N. G. Pace *et al.*, NATO Saclant (1997).
- ⁶J. E. Moe, "Near and far-field acoustic scattering through and from two dimensional fluid-fluid rough interfaces," APL-UW TR 9606, October, 1996.
- ⁷K. L. Williams and D. R. Jackson, "A model for bistatic scattering into ocean sediments for frequencies from 10-100 kHz," APL-UW TR 9505, June, 1996.
- ⁸C. A. Nittrouer and R. W. Sternberg, "The formation of sedimentary strata in an allochthonous shelf environment: The Washington continental shelf," *Mar. Geol.* **42**, 201-232 (1981).

- ⁹E. L. Hamilton, "Geoacoustic modeling of the sea floor," J. Acoust. Soc. Am. **68**, 1313–1340 (1980).
- ¹⁰E. L. Hamilton, "Compressional-wave attenuation in marine sediments," Geophysics **37**, 620–646 (1972).
- ¹¹F. A. Boyle and N. P. Chotiros, "Experimental detection of a slow acoustic wave in sediment at shallow grazing angles," J. Acoust. Soc. Am. **91**, 2615–2619 (1992).
- ¹²H. J. Simpson, E. G. Williams, and B. H. Houston, "Analysis of Laboratory Measurements of Sound Propagating into an Unconsolidated Water-Saturated Porous Media," 16th ICA/135th ASA Proceedings, June (1998).
- ¹³C. B. Officer, *Introduction to the Theory of Sound Transmission* (McGraw-Hill, New York, 1958), Chap. 5, p. 188.

Acoustic scattering from two circular apertures in a thick hard plane

Jin S. Seo, Hyo J. Eom,^{a)} and Haeng S. Lee

Department of Electrical Engineering, Korea Advanced Institute of Science and Technology, 373-1, Kusong Dong, Yuseong Gu, Taejeon, Korea

(Received 5 October 1999; accepted for publication 3 February 2000)

A problem of acoustic wave scattering from two circular apertures in a thick hard plane is solved. The Hankel transform and mode matching is used to represent the scattered field in rapidly convergent series. The reflection coefficient, transmission coefficient, and far-zone field are rigorously derived and presented in numerically efficient forms. Numerical computations are performed to illustrate scattering behaviors in terms of aperture geometry. © 2000 Acoustical Society of America. [S0001-4966(00)01705-7]

PACS numbers: 43.20.Fn, 43.20.El [ANN]

INTRODUCTION

A problem of wave scattering from circular apertures in a plane is an important subject matter in acoustics and electromagnetics due to its practical applications. For instance, electromagnetic wave radiation from multiple circular apertures finds applications in antenna array problems.¹ Acoustic wave scattering from a single circular aperture in a thick plane has been studied in Refs. 2 and 3 and its scattering behavior is well understood. The investigation of acoustic wave scattering from multiple apertures in a plane, however, seems to be very little; hence, it is of theoretical interest to understand a coupling behavior between adjacent apertures in a plane. The purpose of the present paper is to solve a scattering problem of two circular apertures in a thick hard plane by extending the Hankel-transform technique as used in Ref. 3. A use of the Hankel transform allows us to obtain a rigorous, rapidly convergent series solution which is amenable to numerical computation. In the next sections, we present rapidly convergent series solution by enforcing the boundary conditions at the interfaces.

I. FIELD REPRESENTATIONS

Consider a plane acoustic wave impinging on two circular apertures in a thick hard plane (see Fig. 1). Our paper deals with a scalar acoustic problem of a plane compressional wave and no mode conversions are considered. For simplicity, the acoustic parameters in regions I–IV are assumed to be identical. The wave number of each region is k ($=\omega/c$, where ω is angular frequency and c is sound speed) and an $e^{-i\omega t}$ time-harmonic convention is suppressed. Note that a velocity potential Φ is related to the acoustic pressure $p=i\omega\rho_0\Phi$ (ρ_0 is constant equilibrium density), the velocity $\mathbf{v}=\nabla\Phi$. We will use circular cylindrical coordinates (r, ϕ, z) to describe scattering from circular apertures. In region I ($z>0$), an incident field Φ^i (velocity potential) impinges on

two circular apertures. Regions II and III ($-h<z<0$, $r<r_1$, $r'<r_2$) denote the circular apertures. Region IV ($z<-h$) denotes the lossless half-space.

In region I, the total field consists of the incident, reflected, and scattered waves. The incident (Φ^i) and reflected (Φ^r) components are

$$\Phi^{i,r}(x,y,z)=e^{ik\sin\theta_i(x\cos\phi_i+y\sin\phi_i)\mp ikz\cos\theta_i}. \quad (1)$$

Since $x=r\cos\phi-l/2=r'\cos\phi'+l/2$ and $y=r\sin\phi=r'\sin\phi'$, in view of Fig. 1

$$\begin{aligned} \Phi^{i,r}(r,\phi,z) &= e^{\mp ikz\cos\theta_i-(ikl\sin\theta_i/2)} \\ &\times \sum_{m=-\infty}^{\infty} i^m J_m(kr\sin\theta_i) e^{im(\phi-\phi_i)} \end{aligned} \quad (2)$$

$$\begin{aligned} \Phi^{i,r}(r',\phi',z) &= e^{\mp ikz\cos\theta_i+(ikl\sin\theta_i/2)} \\ &\times \sum_{m=-\infty}^{\infty} i^m J_m(kr'\sin\theta_i) e^{im(\phi'-\phi_i)}, \end{aligned} \quad (3)$$

where $J_m(\cdot)$ is the Bessel function of the first kind.

For convenience, the scattered wave is expressed as a sum of two components based on the superposition principle:

$$\Phi^s=\Phi^{s1}(r,\phi,z)+\Phi^{s2}(r',\phi',z), \quad (4)$$

$$\Phi^{s1}(r,\phi,z)=\sum_{m=-\infty}^{\infty} e^{im\phi}\int_0^\infty \tilde{\Phi}_m^{s1}(\zeta)J_m(\zeta r)e^{i\kappa z\zeta}d\zeta, \quad (5)$$

$$\Phi^{s2}(r',\phi',z)=\sum_{m=-\infty}^{\infty} e^{im\phi'}\int_0^\infty \tilde{\Phi}_m^{s2}(\zeta)J_m(\zeta r')e^{i\kappa z\zeta}d\zeta, \quad (6)$$

where $\kappa=\sqrt{k^2-\zeta^2}$. Note that the Hankel transform representation (5) is a solution to the wave equation $(\nabla^2+k^2)\Phi^{s1}(r,\phi,z)=0$ in cylindrical coordinates.³ A detailed derivation of (5), which is based on the technique of separation of variables, is available in Ref. 4. Since $\Phi^{s1}(r,\phi,z)$ should vanish as $z\rightarrow\infty$, we choose an eigenfunction $e^{i\kappa z}$ to satisfy the radiation condition.

^{a)}Electronic mail: hjeom@ee.kaist.ac.kr

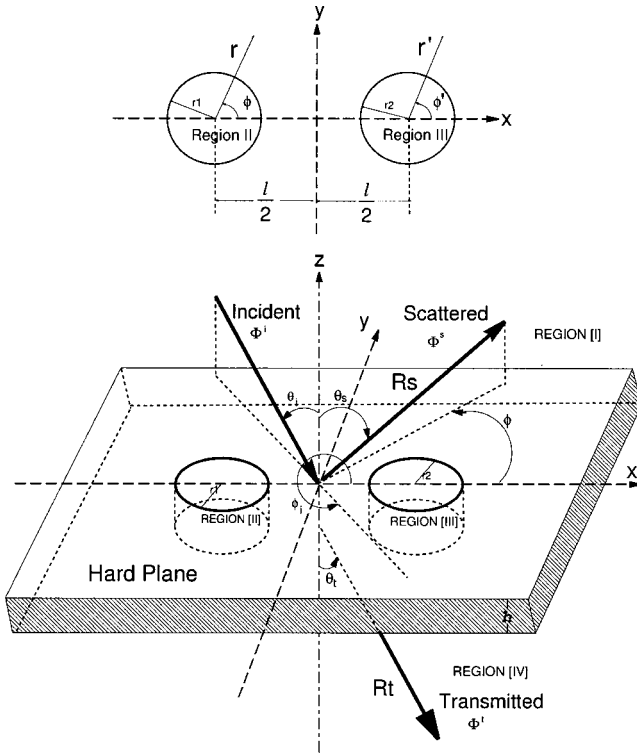


FIG. 1. Geometry of plane wave scattering by two circular apertures in a thick hard plane (radii r_1 and r_2 , thickness h , distance between apertures l , incident polar angle θ_i , and incident azimuth angle ϕ_i).

In regions II and III, the fields are, respectively, given as a superposition of discrete eigenfunctions in circular cylindrical coordinates:

$$\Phi^{\text{II}}(r, \phi, z) = \sum_{m=-\infty}^{\infty} e^{im\phi} \sum_{n=1}^{\infty} [a_n^m \sin k_z^m(z+h) + b_n^m \cos k_z^m(z+h)] J_m(k_n^m r), \quad (7)$$

$$\Phi^{\text{III}}(r', \phi', z) = \sum_{m=-\infty}^{\infty} e^{im\phi'} \sum_{n=1}^{\infty} [c_n^m \sin \bar{k}_z^m(z+h) + d_n^m \cos \bar{k}_z^m(z+h)] J_m(\bar{k}_n^m r'), \quad (8)$$

where $k_z^m = \sqrt{k^2 - (k_n^m)^2}$, $J'_m(k_n^m r_1) = 0$; $\bar{k}_z^m = \sqrt{k^2 - (\bar{k}_n^m)^2}$, $J'_m(\bar{k}_n^m r_2) = 0$, and the prime ' denotes differentiation with respect to the argument.

In region IV, the transmitted wave is a sum of two components,

$$\Phi^{\text{IV}} = \Phi^{t1}(r, \phi, z) + \Phi^{t2}(r', \phi', z) \quad (9)$$

$$\Phi^{t1}(r, \phi, z) = \sum_{m=-\infty}^{\infty} e^{im\phi} \int_0^{\infty} \tilde{\Phi}_m^{t1}(\zeta) J_m(\zeta r) e^{-i\kappa(\zeta+h)} \zeta d\zeta, \quad (10)$$

$$\begin{aligned} \Phi^{t2}(r', \phi', z) \\ = \sum_{m=-\infty}^{\infty} e^{im\phi'} \int_0^{\infty} \tilde{\Phi}_m^{t2}(\zeta) J_m(\zeta r') e^{-i\kappa(\zeta+h)} \zeta d\zeta. \end{aligned} \quad (11)$$

II. ENFORCEMENT OF BOUNDARY CONDITIONS

In order to determine unknown modal coefficients a_n^m , b_n^m , c_n^m , and d_n^m , the enforcement of boundary conditions at the interfaces is required. Explicitly, the boundary conditions include the continuity conditions for the interfaces I–II, I–III, IV–II, and IV–III, and “motion-free” boundary conditions for all interfaces between the acoustic media and the hard plane. On the surface of the hard plane, the incoming wave is completely reflected. The normal component of the particle velocity vanishes at the surface of the hard plane (surface boundary condition: $\nabla \Phi \cdot \hat{n} = 0$, \hat{n} is unit normal to the surface). The boundary conditions at $z=0$ yield

$$\left. \frac{\partial \Phi^{s1}}{\partial z} \right|_{z=0} = \begin{cases} \left. \frac{\partial \Phi^{\text{II}}}{\partial z} \right|_{z=0}, & r < r_1, \\ 0, & r > r_1, \end{cases} \quad (12)$$

$$\left. \frac{\partial \Phi^{s2}}{\partial z} \right|_{z=0} = \begin{cases} \left. \frac{\partial \Phi^{\text{III}}}{\partial z} \right|_{z=0}, & r' < r_2, \\ 0, & r' > r_2, \end{cases} \quad (13)$$

$$\Phi^{\text{I}}(r, \phi, 0) + \Phi^{\text{r}}(r, \phi, 0) + \Phi^{\text{s}}(r, \phi, 0) = \Phi^{\text{II}}(r, \phi, 0), \quad r < r_1, \quad (14)$$

$$\Phi^{\text{I}}(r', \phi', 0) + \Phi^{\text{r}}(r', \phi', 0) + \Phi^{\text{s}}(r', \phi', 0) = \Phi^{\text{III}}(r', \phi', 0), \quad r' < r_2, \quad (15)$$

Applying the Hankel transforms $\int_0^{\infty} (\bullet) J_m(\zeta r) r dr$ to (12) and (13) as well as the inverse Hankel transform theorem, respectively, yields

$$\tilde{\Phi}_m^{s1}(\zeta) = \frac{-i}{\kappa} \sum_{n=1}^{\infty} k_z^m (a_n^m \cos k_z^m h - b_n^m \sin k_z^m h) I_1^m(k_n^m, \zeta), \quad (16)$$

$$\tilde{\Phi}_m^{s2}(\zeta) = \frac{-i}{\kappa} \sum_{n=1}^{\infty} k_z^m (c_n^m \cos \bar{k}_z^m h - d_n^m \sin \bar{k}_z^m h) I_2^m(\bar{k}_n^m, \zeta), \quad (17)$$

where

$$\begin{aligned} I_{1,2}^m(\alpha, \beta) = \frac{r_{1,2}}{\alpha^2 - \beta^2} [\alpha J_{m+1}(\alpha r_{1,2}) J_m(\beta r_{1,2}) \\ - \beta J_m(\alpha r_{1,2}) J_{m+1}(\beta r_{1,2})] \quad (\alpha \neq \beta), \end{aligned} \quad (18)$$

$$I_{1,2}^m(\alpha, \alpha) = \frac{r_{1,2}^2}{2} [J_m^2(\alpha r_{1,2}) - J_{m-1}(\alpha r_{1,2}) J_{m+1}(\alpha r_{1,2})]. \quad (19)$$

In view of Fig. 2, the Graf's addition theorem⁵ gives

$$J_p(\zeta r') e^{ip\phi'} = \sum_{m=-\infty}^{\infty} J_{m-p}(\zeta l) J_m(\zeta r) e^{im\phi}, \quad (20)$$

$$J_p(\zeta r) e^{ip\phi} = \sum_{m=-\infty}^{\infty} J_{p-m}(\zeta l) J_m(\zeta r') e^{im\phi'}. \quad (21)$$

Substituting (16), (17), and (20) into (14), multiplying by $J_m(k_n^m r) r$, and integrating over $0 < r < r_1$ gives

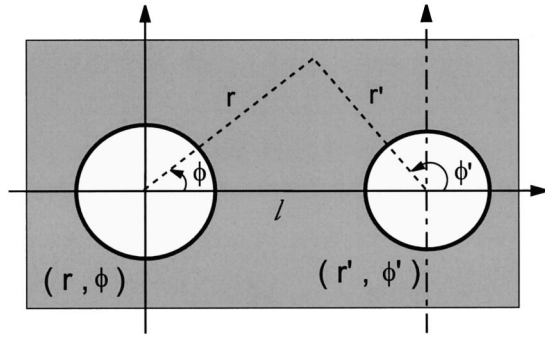


FIG. 2. Graf's addition theorem.

$$\begin{aligned}
 & 2e^{-(ikl/2) \sin \theta_i - im\phi_i} i^m I_2^m(k \sin \theta_i, k_q^m) \\
 & - \sum_{n=1}^{\infty} ik_2^m (a_n^m \cos k_z^m h - b_n^m \sin k_z^m h) I_{11} \\
 & - \sum_{p=-\infty}^{\infty} \sum_{n=1}^{\infty} i\bar{k}_z^p (c_n^p \cos \bar{k}_z^p h - d_n^p \sin \bar{k}_z^p h) I_{12} \\
 & = [a_q^m \sin k_z^m h + b_q^m \cos k_z^m h] I_1^m(k_q^m, k_q^m), \quad (22)
 \end{aligned}$$

where

$$I_{11} = \int_0^{\infty} \kappa^{-1} I_1^m(k_n^m, \zeta) I_1^m(k_q^m, \zeta) \zeta d\zeta, \quad (23)$$

$$I_{12} = \int_0^{\infty} \kappa^{-1} J_{m-p}(\zeta l) I_1^m(k_q^m, \zeta) I_2^p(\bar{k}_n^p, \zeta) \zeta d\zeta. \quad (24)$$

Note that the orthogonality condition of the Bessel function $\int_0^r J_m(k_q^m r) J_m(k_n^m r) dr = \delta_{qn}$ is utilized in obtaining (22) where δ_{qn} is the Kronecker delta.

Similarly from (16), (17), (21), and (15),

$$\begin{aligned}
 & 2e^{(ikl/2) \sin \theta_i - im\phi_i} i^m I_2^m(k \sin \theta_i, \bar{k}_q^m) \\
 & - \sum_{p=-\infty}^{\infty} \sum_{n=1}^{\infty} ik_z^p (a_n^p \cos k_z^p h - b_n^p \sin k_z^p h) I_{21} \\
 & - \sum_{n=1}^{\infty} i\bar{k}_z^m (c_n^m \cos \bar{k}_z^m h - d_n^m \sin \bar{k}_z^m h) I_{22} \\
 & = [c_q^m \sin \bar{k}_z^m h + d_q^m \cos \bar{k}_z^m h] I_2^m(\bar{k}_q^m, \bar{k}_q^m), \quad (25)
 \end{aligned}$$

where

$$I_{22} = \int_0^{\infty} \kappa^{-1} I_2^m(\bar{k}_n^m, \zeta) I_2^m(\bar{k}_q^m, \zeta) \zeta d\zeta, \quad (26)$$

$$I_{21} = \int_0^{\infty} \kappa^{-1} J_{p-m}(\zeta l) I_1^p(k_n^p, \zeta) I_2^m(\bar{k}_q^m, \zeta) \zeta d\zeta. \quad (27)$$

The boundary conditions at $z = -h$ require

$$\left. \frac{\partial \Phi^{I1}}{\partial z} \right|_{z=-h} = \begin{cases} \frac{\partial \Phi^{II}}{\partial z} \Big|_{z=-h}, & r < r_1, \\ 0, & r > r_1, \end{cases} \quad (28)$$

$$\left. \frac{\partial \Phi^{I2}}{\partial z} \right|_{z=-h} = \begin{cases} \frac{\partial \Phi^{III}}{\partial z} \Big|_{z=-h}, & r' < r_2, \\ 0, & r' > r_2, \end{cases} \quad (29)$$

$$\Phi^I(r, \phi, -h) = \Phi^{II}(r, \phi, -h), \quad r < r_1, \quad (30)$$

$$\Phi^I(r', \phi', -h) = \Phi^{III}(r', \phi', -h), \quad r' < r_2. \quad (31)$$

Note that $\Phi^I = \Phi^{I1} + \Phi^{I2}$ is given by (9). Applying the Hankel transforms and (28) and (29) gives

$$\tilde{\Phi}_m^{I1}(\zeta) = \frac{i}{\kappa} \sum_{n=1}^{\infty} k_z^m a_n^m I_1^m(k_n^m, \zeta), \quad (32)$$

$$\tilde{\Phi}_m^{I2}(\zeta) = \frac{i}{\kappa} \sum_{n=1}^{\infty} \bar{k}_z^m c_n^m I_2^m(\bar{k}_n^m, \zeta). \quad (33)$$

Substituting (20), (21), (32), and (33) into (30) and (31), and rearranging them, we obtain

$$\sum_{n=1}^{\infty} ik_z^m a_n^m I_{11} + \sum_{p=-\infty}^{\infty} \sum_{n=1}^{\infty} i\bar{k}_z^p c_n^p I_{12} = b_q^m I_1^m(k_q^m, k_q^m), \quad (34)$$

$$\sum_{p=-\infty}^{\infty} \sum_{n=1}^{\infty} ik_z^p a_n^p I_{22} + \sum_{n=1}^{\infty} i\bar{k}_z^m c_n^m I_{21} = d_q^m I_2^m(\bar{k}_q^m, \bar{k}_q^m). \quad (35)$$

The sets of simultaneous equations (22), (25), (34), and (35) can be solved numerically, yielding the unknown modal coefficients.

III. NUMERICAL COMPUTATIONS

The reflection coefficient ρ_ν (or transmission coefficient τ_ν) is defined as a ratio of the power reflected from (or transmitted through) aperture ν to the power impinging on the aperture ν where ν is 1 or 2. They are

$$\rho_\nu = \frac{-\int_S \text{Re}[p_\nu^* \mathbf{v}_\nu|_{z=0}] \cdot \hat{\mathbf{z}} ds}{\int_S \text{Re}[p_{\text{incident}}^* \mathbf{v}_{\text{incident}}|_{z=0}] \cdot \hat{\mathbf{z}} ds}, \quad (36)$$

$$\tau_\nu = \frac{\int_S \text{Re}[p_\nu^* \mathbf{v}_\nu|_{z=-h}] \cdot \hat{\mathbf{z}} ds}{\int_S \text{Re}[p_{\text{incident}}^* \mathbf{v}_{\text{incident}}|_{z=0}] \cdot \hat{\mathbf{z}} ds}, \quad (37)$$

where the symbols $\text{Re}(\dots)$ and $(\dots)^*$ denote the real part and complex conjugate of (\dots) . Note that $\hat{\mathbf{z}}$ is a unit vector normal to the surface, p_ν and \mathbf{v}_ν are, respectively, the acoustic pressure and velocity vector over the aperture ν , and S is a cross-sectional area of aperture ν . We define the reflection coefficient ρ and transmission coefficient τ of two circular apertures as a sum of two reflection and transmission coefficients, respectively,

$$\rho = \rho_1 + \rho_2, \quad (38)$$

with

$$\rho_1 = -\frac{2}{r_1^2 k \cos \theta_i} \sum_{m=-\infty}^{\infty} \sum_{n=1}^{\infty} \text{Im} \{ k_z^{m*} (a_n^m \cos k_z^m h - b_n^m \sin k_z^m h) * [(a_n^m \sin k_z^m h + b_n^m \cos k_z^m h) I_1^m(k_n^m, k_n^m) - 2i^m e^{-(ikl/2) \sin \theta_i - im\phi_i} I_1^m(k \sin \theta_i, k_n^m)] \}, \quad (39)$$

$$\rho_2 = -\frac{2}{r_2^2 k \cos \theta_i} \sum_{m=-\infty}^{\infty} \sum_{n=1}^{\infty} \text{Im} \{ \bar{k}_z^{m*} (c_n^m \cos \bar{k}_z^m h - d_n^m \sin \bar{k}_z^m h) * [(c_n^m \sin \bar{k}_z^m h + d_n^m \cos \bar{k}_z^m h) I_2^m(\bar{k}_n^m, \bar{k}_n^m) - 2i^m e^{-(ikl/2) \sin \theta_i - im\phi_i} I_2^m(k \sin \theta_i, \bar{k}_n^m)] \}, \quad (40)$$

and

$$\tau = \tau_1 + \tau_2, \quad (41)$$

with

$$\tau_1 = \frac{2}{r_1^2 k \cos \theta_i} \sum_{m=-\infty}^{\infty} \sum_{n=1}^{\infty} \text{Im} (k_z^{m*} b_n^m a_n^{m*}) I_1^m(k_n^m, k_n^m), \quad (42)$$

$$\tau_2 = \frac{2}{r_2^2 k \cos \theta_i} \sum_{m=-\infty}^{\infty} \sum_{n=1}^{\infty} \text{Im} (\bar{k}_z^{m*} d_n^m c_n^{m*}) I_2^m(\bar{k}_n^m, \bar{k}_n^m), \quad (43)$$

where $\text{Im}(\dots)$ denote the imaginary part of (\dots) .

Using the saddle-point approximation,⁶ the far-zone scattered and transmitted fields at distance R_s and R_t are

$$\Psi^s(R_s, \theta_s, \phi) = \frac{e^{i(kR_s - \pi/4)} \cos \theta_s}{R_s \sqrt{\sin \theta_s}} \sum_{m=-\infty}^{\infty} \sum_{p=-\infty}^{\infty} [e^{i(m\phi - p\pi/2)} \times \{ -ie^{i\zeta l/2} \tilde{\Psi}_p^{s1}(\zeta) + e^{-i\zeta l/2} \tilde{\Psi}_p^{s2}(\zeta) \} + (-1)^m e^{i(m\phi + p\pi/2)} \{ e^{-i\zeta l/2} \tilde{\Psi}_p^{s1}(\zeta) - ie^{i\zeta l/2} \tilde{\Psi}_p^{s2}(\zeta) \}]|_{\zeta = k \sin \theta_s} \quad (44)$$

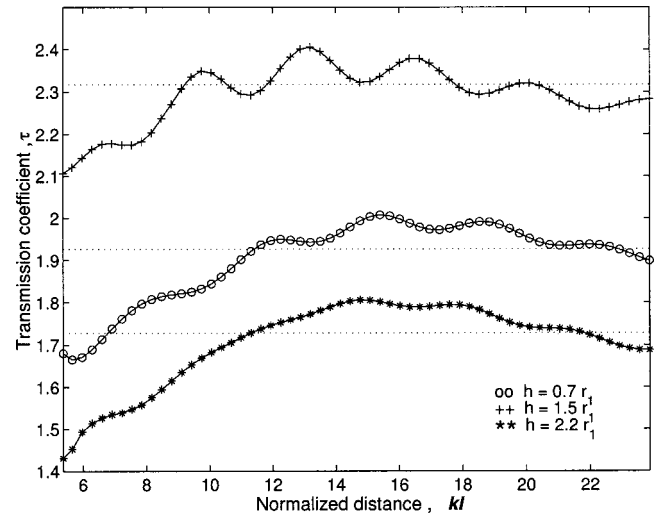


FIG. 4. Behavior of transmission coefficient versus normalized distance for $kr_1 = \pi$, $kr_2 = 0.7\pi$, $\theta_i = 45$ degrees, $\phi_i = 45$ degrees, --- a sum of two single-aperture transmission coefficients where the coupling is ignored.

and

$$\Psi^t(R_t, \theta_t, \phi) = \frac{e^{i(kR_t - \pi/4)} e^{-i(kh \cos \theta_t)} \cos \theta_t}{R_t \sqrt{\sin \theta_t}} \times \sum_{m=-\infty}^{\infty} \sum_{p=-\infty}^{\infty} [(-1)^m e^{i(m\phi - p\pi/2)} \times \{ ie^{i\zeta l/2} \tilde{\Psi}_p^{t1}(\zeta) - e^{-i\zeta l/2} \tilde{\Psi}_p^{t2}(\zeta) \} + e^{i(m\phi + p\pi/2)} \{ -e^{-i\zeta l/2} \tilde{\Psi}_p^{t1}(\zeta) + ie^{i\zeta l/2} \tilde{\Psi}_p^{t2}(\zeta) \}]|_{\zeta = k \sin \theta_t} \quad (45)$$

Figure 3 illustrates the behavior of the transmission co-

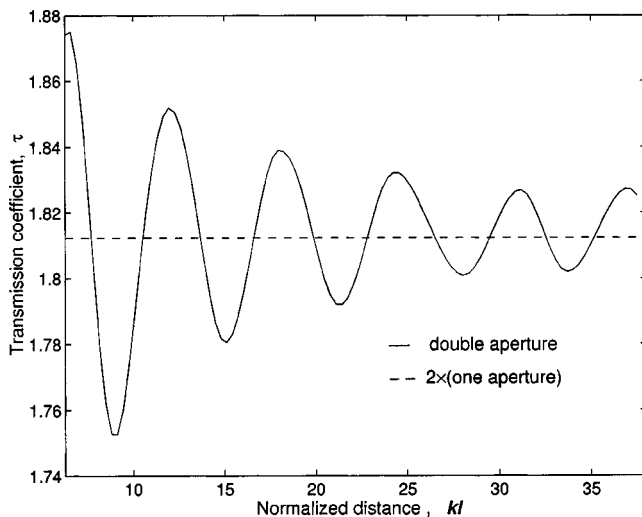


FIG. 3. Behavior of transmission coefficient versus normalized distance for $kr_1 = \pi$, $kr_2 = \pi$, $kh = 0.5\pi$, $\theta_i = 0$ degrees, and $\phi_i = 0$ degrees.

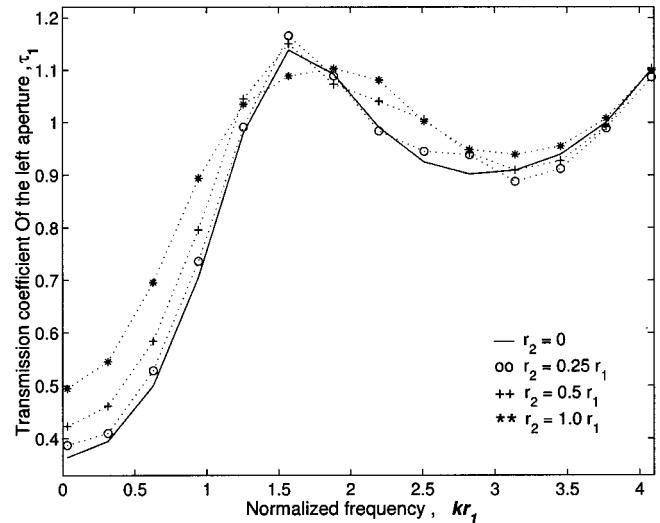


FIG. 5. Behavior of transmission coefficient τ_1 versus normalized frequency for $l = 2.1r_1$, $h = 0.7r_1$, $\theta_i = 0$ degrees, and $\phi_i = 0$ degrees.

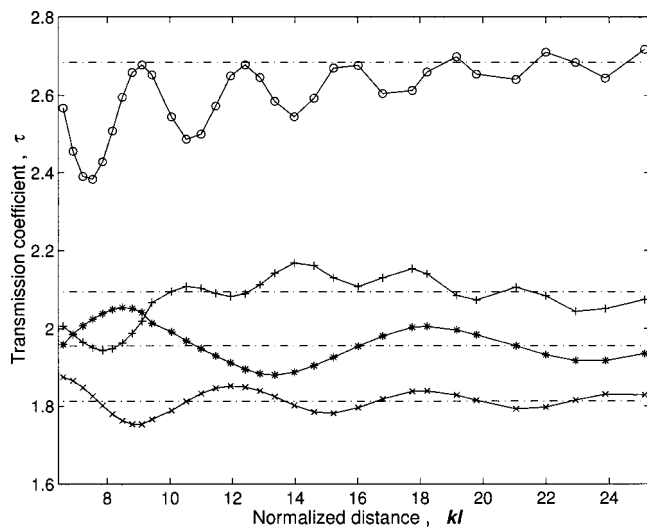


FIG. 6. Behavior of transmission coefficient versus normalized distance for $kr_1 = \pi$, $kr_2 = \pi$, $kh = 0.5\pi$, $\phi_i = 0$ degrees; $\times \times \times$ $\theta_i = 0$ degrees, $***$ $\theta_i = 22.5$ degrees, $+++$ $\theta_i = 45$ degrees, $\infty \infty$ $\theta_i = 67.5$, --- $2 \times (\text{one aperture } \tau)$ for each θ_i .

efficient τ through two apertures versus a distance (kl) between apertures at normal incidence. As kl increases, τ oscillates around 2×0.906 , where 0.906 is the transmission coefficient of a single aperture. The number of modes used in the computation is $m = 5$, which indicates a good convergence of our solution.

Figure 4 illustrates the behavior of τ versus kl for three different thickness $kh = 0.7\pi$, 1.5π , and 2.2π . As kl increases, τ tends to approach a sum of two single-aperture ($kr_1 = \pi$, $kr_2 = 0.7\pi$) transmission coefficients where the coupling between two apertures is ignored.

Figure 5 displays the behavior of transmission coefficient τ_1 through region II versus the normalized frequency kr_1 for four different r_2 . As r_2 increases from 0 to r_1 , the curve deviates further away from a case with $r_2 = 0$ (single aperture case), indicating that the coupling becomes stronger as r_2 increases.

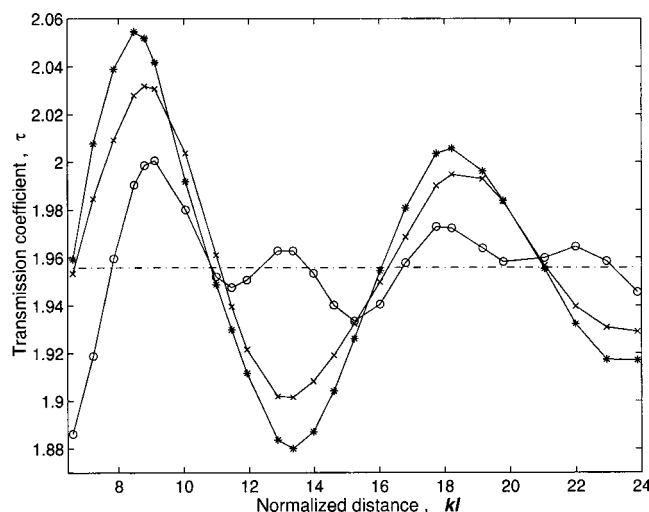


FIG. 7. Behavior of transmission coefficient versus normalized distance for $kr_1 = \pi$, $kr_2 = \pi$, $kh = 0.5\pi$, $\theta_i = 22.5$ degrees; $***$ $\phi_i = 0$ degrees, $\times \times \times$ $\phi_i = 45$ degrees, $\infty \infty$ $\phi_i = 90$ degrees, --- $2 \times (\text{one aperture } \tau)$.

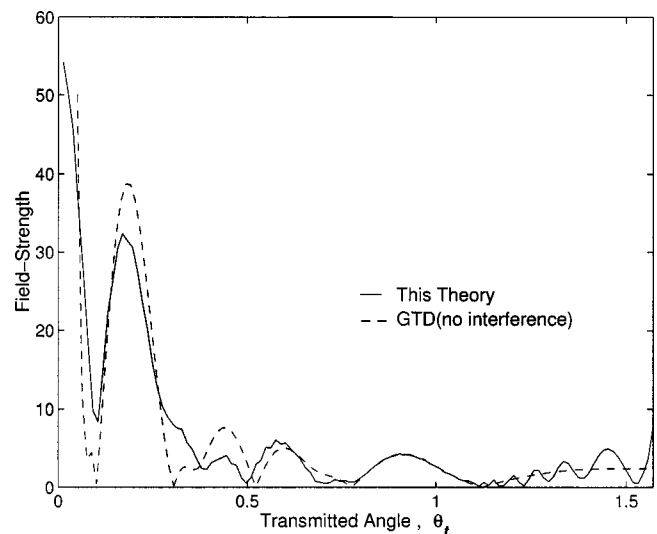


FIG. 8. Far-zone transmitted field versus the transmitted angle for $kr_1 = 3.5\pi$, $kr_2 = 3.5\pi$, $kh = 0$, $kl = 12\pi$, $\theta_i = 0$ degrees, and $\phi_i = 0$ degrees.

Figure 6 shows the transmission coefficient τ versus the normalized distance kl at four different incident polar angles θ_i . As kl increases, τ fluctuates around twice the transmission coefficient of a single aperture irrespective of θ_i . For small kl , a degree of interference (coupling) between two apertures is somewhat sensitive to the incident polar angle.

Figure 7 shows the behavior of τ versus kl for four incident azimuth angles ϕ_i . The largest (smallest) fluctuation in τ occurs when $\phi_i = 0$ degrees (90 degrees), indicating that an interference (coupling) between two aperture becomes most pronounced when $\phi_i = 0$ degrees.

Figure 8 shows the far-zone field transmitted through two apertures versus the transmitted angle. Our solution is compared with an approximate GTD (geometrical theory of diffraction) solution which ignores the coupling between two apertures. The GTD solution is a high-frequency approximation, which is based on the classical geometrical optics as well as the diffraction mechanism.⁷ The comparison between ours and the GTD solution shows a favorable agreement in general trend.

IV. CONCLUSION

A rigorous solution for scattering from two circular apertures in a hard plane is obtained. The solution is represented in fast-convergent series so that it is efficient for numerical computation. A degree of interference (coupling) between two apertures is discussed in terms of the distance between two apertures (kl), incident polar angle (θ_i) and azimuth angle (ϕ_i), and aperture radii (kr_1, kr_2). It is possible to extend our analytical approach developed for two-aperture scattering to a multiple-aperture scattering case. A study of multiple-aperture scattering may find practical applications in the area of noise-control using perforated plates.

¹T. S. Bird, "Improved solution for mode coupling in different sized circular apertures and its application," IEE Proc., Part H: Microwaves, Antennas Propag. **143**, 457-464 (1996).

²G. P. Willson and W. W. Soroka, "Approximation to the diffraction of

- sound by a circular aperture in a rigid wall of finite thickness,” J. Acoust. Soc. Am. **37**, 286–297 (1965).
- ³K. H. Jun and H. J. Eom, “Acoustic scattering from a circular aperture in a thick hard screen,” J. Acoust. Soc. Am. **98**, 2324–2327 (1995).
- ⁴J. D. Stratton, *Electromagnetic Theory* (McGraw-Hill, New York, 1941), pp. 361–372.
- ⁵M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions* (Dover, New York, 1965), p. 363.
- ⁶C. T. Tai, *Dyadic Green’s Functions in Electromagnetic Theory*, 2nd ed. (IEEE, New York, 1994), pp. 16–20.
- ⁷J. B. Keller, “Diffraction by an aperture,” J. Appl. Phys. **28**(4), 426–445 (1957).

Scale model experiments on the insertion loss of wide and double barriers

Glenn J. Wadsworth and James P. Chambers

National Center for Physical Acoustics, University of Mississippi, University, Mississippi 38677

(Received 23 February 1999; accepted for publication 11 January 2000)

The insertion loss of wide and double barriers is investigated through scale model experiments. Such configurations appear in outdoor sound propagation problems such as highway noise reduction and community noise control. The Biot–Tolstoy–Medwin (BTM) time domain wedge formulation for multiple diffraction [J. Acoust. Soc. Am. **72**, 1005–1013 (1982)] is used to predict the acoustic response of an impulsive source. Evaluation of the insertion loss at discrete frequencies is accomplished via the fast Fourier transform (FFT). Good agreement has been found between the BTM model and experimental data for all configurations tested. © 2000 Acoustical Society of America. [S0001-4966(00)03204-5]

PACS numbers: 43.20.Fn, 43.28.Fp, 43.50.Vt [ANN]

INTRODUCTION

The purpose of this study was to examine the diffraction effects caused by an acoustical barrier having two diffracting edges. The use of barriers to reduce noise near highways, railways, factories, etc. is widespread in large cities. The most common measure of the acoustic performance of a barrier is its insertion loss, which is the difference in sound pressure levels with and without the barrier. This quantity is used in the present work to describe the acoustic performance of a barrier.

The insertion loss of infinitely long barriers has been extensively studied. Methods of calculating barrier performance have been reviewed by Isei *et al.*¹ and Nicolas *et al.*² A thorough collection of references on the theory of sound diffraction can also be found in the more recent work by Salomons.³ The majority of the above-mentioned theories are based in the frequency domain and are utilized by many current computer prediction schemes.

Alternate time domain solutions presented in the late 1950's^{4,5} have recently been applied in modified forms^{6–14} to the wedge diffraction problem and have been validated with scale model experiments. The present work uses Medwin *et al.*'s formulation^{7,14,15} to calculate the insertion loss of wide and double barriers in the presence of a rigid ground surface.

Wide or three sided barriers were analyzed by Pierce¹⁶ but his model is not posed for multiple barriers that are not connected. Salomons³ offers a modification to the Pierce model to handle multiple arbitrarily placed barriers but his solution has some singularities that must be negotiated around. The current model investigated provides an alternate solution that may also be ideally suited to investigate transient phenomena or multiple finite length wedges. Its computational intensity, however, will probably not favor replacement of frequency based solutions such as Salomons for many practical noise prediction routines. Typical run times for the wide and double barrier problems ranged to a few minutes on a 100 MHz PC while the Pierce wide barrier solution averaged a few seconds.

The extension of the Biot–Tolstoy–Medwin (BTM) model to double-edge diffraction is briefly described in Sec.

I. The corresponding scale model experiment, using a spark as the sound source, is described in Sec. II. Section III compares the theoretical predictions with experimental results and conclusions are made in Sec. IV.

I. THEORY

A. Time domain diffraction

For a point source, the geometry used by Medwin^{7,15} is followed here and is shown in Fig. 1. Source coordinates are given by r_0 , θ_0 , and $z=0$; receiver coordinates are given by r , θ , and z . The wedge angle, θ_w , is measured from one face of the wedge through the propagating medium to the other face. For the thin plate, this angle is 2π , and such a barrier is referred to as a knife-edge barrier.

The diffracted pressure waveform is given in closed form by Medwin in Refs. 7 and 15 and is not repeated here for brevity. The impulse formulation can provide valuable insight into the relative contributions of each arriving pulse. In addition, ray-tracing techniques allow comparison of distinct reflected and diffracted components. It is often more desirable, however, to quantify the diffracted sound field in the frequency domain. The frequency response of the evaluated waveform is easily obtained by digital means. In order to represent the diffracted field digitally, instantaneous values of the pressure are calculated for discrete values of time and the frequency content of the diffracted field is obtained via the fast Fourier transform (FFT).

This discrete, digital representation of the pressure diffracted by a barrier follows from a classical Huygen's wavelet description of the spherical wave. Secondary sources are defined at interception points between the wavefront and the edge. The first secondary source is located on the intercept of the diffracting edge and the least time path. As the wavefront propagates, secondary sources on both sides of the least time path contribute to the sound field in pairs. The paths defined by the positions of the secondary sources are referred to as partial least time paths, and pulses propagating along these paths arrive at intervals of $n\Delta T$ after the first arrival. The complete time domain waveform is composed of N points

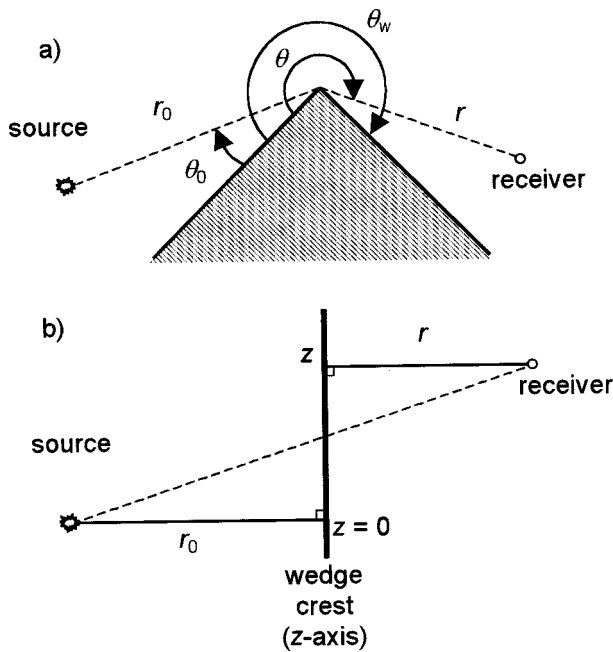


FIG. 1. BTM model wedge geometry: (a) side view; (b) top (unfolded) view. Dashed lines indicate the least time path.

whose values are written as $p(n\Delta T)$. The desired frequency range governs the choice of the discrete time step, ΔT , and the total duration, $N\Delta T$, of the impulse response. In this work the discrete time step was chosen to be $1 \mu\text{s}$ and the total duration was 4.096 ms. The pressure at $t = \tau_0$, where τ_0 is the least time over the barrier, is written as $p(0\Delta T)$. In the symmetric case where the source-receiver axis is normal to the diffracting edge, individual secondary sources make a contribution of $\frac{1}{2}p(n\Delta T)$ to the total waveform.

The model used in this work assumes that all surfaces, including the ground, are perfectly reflecting. Under this assumption, ray paths that include one or more ground bounces are treated as having originated from an image source and/or terminating at an image receiver whose locations are determined by Snell's Law. Once the images are in the proper location, the ground is removed and the ray path simply extends from image source to edge to image receiver (see, for example, Ref. 1).

When a single knife-edge barrier is placed on a rigid surface, four ray paths from the source to a receiver in the shadow zone are possible: (1) source-edge-receiver, (2)

source-ground-edge-receiver, (3) source-edge-ground-receiver, and (4) source-ground-edge-ground-receiver. The diffracted signal for each ray path is calculated independently using the method described above and then superimposed.

B. Double-edge diffraction

Multiple diffraction effects are easily handled by the BTM model in the context of the classical Huygen's Principle. These effects must be considered when predicting the insertion losses of thick and/or multiple barriers. Although the wide barrier has been addressed by Medwin *et al.*,¹⁵ Pierce¹⁶ and Salomons,³ it is reexamined in the presence of a rigid ground surface. In addition, diffraction over double knife-edge barriers of various separations is investigated. The three types of barriers, (1) single knife-edge, (2) wide, and (3) double knife-edge, are compared in terms of the insertion loss which will be addressed later.

The double diffraction method treats each diffracting edge independently. Each secondary source along the first edge in turn radiates a new, weaker pulse which diffracts over the second edge in the same manner as discussed above. Although similar, the calculations used for the wide and double knife-edge barriers have a few subtle differences; therefore the two barriers are addressed separately. The BTM model does not require that the diffracting edges of the barriers be parallel, nor does it place any stipulations on the relative heights of the two edges. In this discussion, however, the barriers are parallel with constant separation, W , and are the same height, h . Furthermore, the source-receiver axis is normal to the z -axis, or barriers, in all cases.

Cartesian coordinates are used to conveniently give the locations of the source and receiver throughout this discussion. For both the wide and double barrier, the origin of the BTM cylindrical coordinate system depends on which edge is being treated in the double diffraction method. In the Cartesian system, used only for measurement convenience, the origin, O , is placed at the point of contact between the first diffracting edge and the ground surface. The case of the double knife-edge barrier is illustrated in Fig. 2. The conversion is the same for the wide barrier.

The double diffraction method is illustrated in Fig. 3 for the wide barrier. The wide barrier is extended to form a single right-angle wedge, and a virtual receiver is placed at

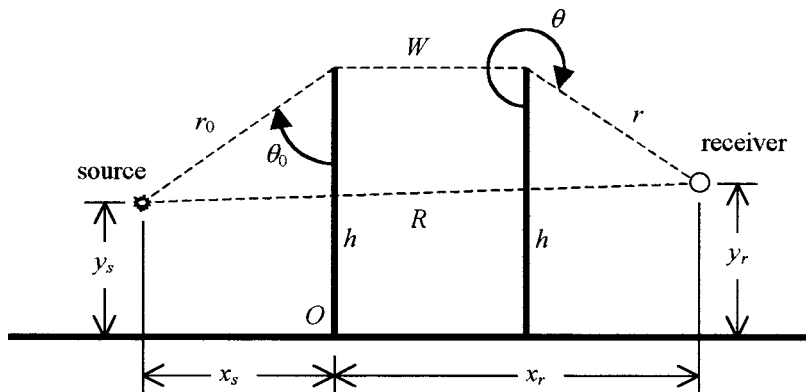


FIG. 2. Conversion from cylindrical to Cartesian coordinates for the wide or double knife-edge barrier. The origin of the Cartesian system is at point O .

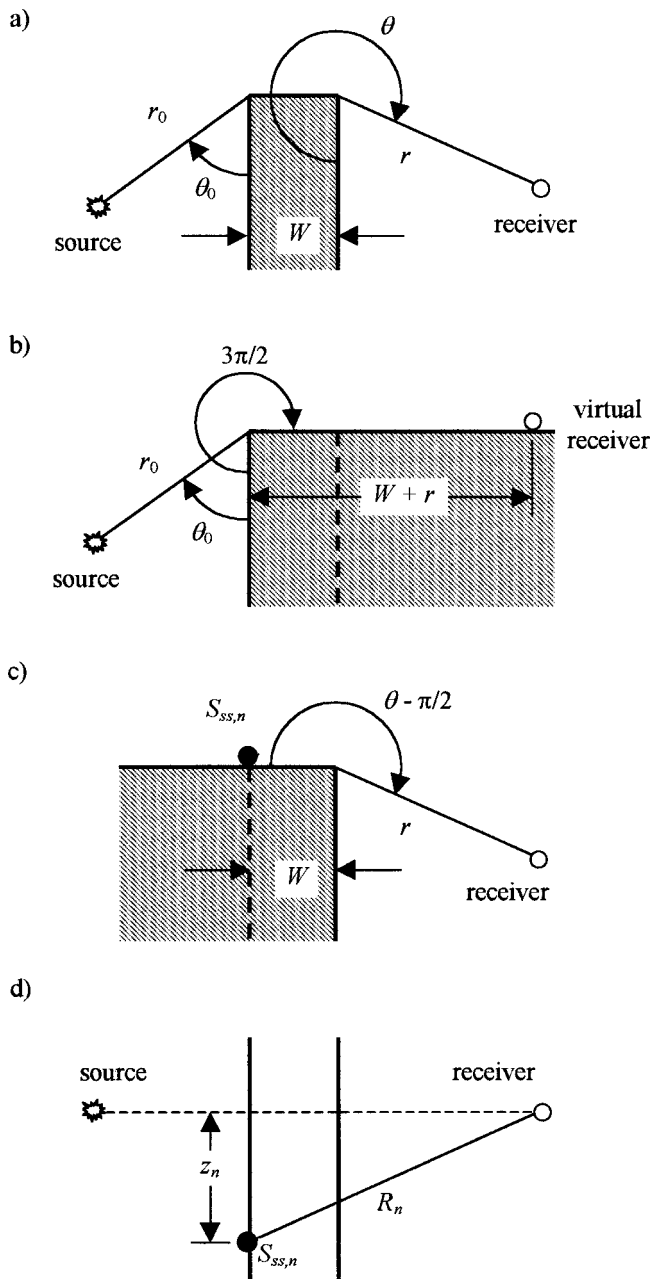


FIG. 3. Wide barrier double diffraction method. (a) Typical wide barrier geometry. (b) Placement of virtual receiver. (c) Diffraction over the second edge from the n th secondary source. (d) The n th partial least time path, top view.

coordinates $(r', \theta', 0)$ where $r' = W + r$ and $\theta' = \theta_w = 3\pi/2$. Using the single diffraction formulation, N values of $p(n\Delta T)$ are calculated at the virtual receiver. The fractional strength of the zeroth secondary source, $S_{ss,0}$, is then determined. This source strength is simply $S_{ss,0} = F_0 S$, where

$$F_0 = \frac{p(0\Delta T)}{p_{\delta,0}}. \quad (1)$$

The value of $p_{\delta,0}$ gives what the pressure “would have been” at the virtual receiver had the pulse originated from the zeroth secondary source location, and is given by

$$p_{\delta,0} = \frac{S\rho}{2\pi R_0 \Delta T}, \quad (2)$$

where $R_0 = W + r$, ρ is the density and S is the source strength of the original point source, typically having a value of 1. The factor of 2 in the denominator of Eq. (2) replaces the usual factor of 4 for a point source due to pressure doubling at the face of the wedge. Subsequent secondary source strengths are calculated in pairs as $S_{ss,n} = F_n S$, where

$$F_n = \frac{1}{2} \frac{p(n\Delta T)}{p_{\delta,n}} \quad (3)$$

and

$$p_{\delta,n} = \frac{S\rho}{2\pi R_n \Delta T} \quad (4)$$

for $R_n = [(W + r)^2 + z_n^2]^{1/2}$. Contributions from secondary sources accumulate until the cutoff criteria

$$F_n = \frac{S_{ss,n}}{S} < 0.5\% \quad (5)$$

is met. This cutoff criteria (also used by Medwin in Ref. 15) defines N_{ss} , the value of n when Eq. (5) is satisfied. Once these secondary source strengths have been found, diffraction over the second edge can be treated.

A second right-angle wedge is formed as shown in Fig. 3(c) such that each secondary source has coordinates (r'_0, θ'_0, z_n) where $r'_0 = W$ and $\theta'_0 = 0$. The N_{ss} secondary sources are used to calculate N_{ss} diffracted waveforms which are combined at the appropriate delay times of $n\Delta T$ to give the final pressure signal due to that ray path. These steps are repeated for each of the four possible ray paths. The superposition of these four signals gives the total impulse response for diffraction over the original wide barrier.

The same basic technique as described above is applied to the double knife-edge barrier to calculate the effects of double diffraction. That is, the impulse response of a point source diffracted over this double barrier is the superposition of separate ray paths, each of which is the superposition of N_{ss} individual waveforms originating from secondary sources along the first diffracting edge. However, since the double knife-edge barrier consists of two plates rather than a solid object, several distinct differences arise.

For both diffracting edges, the wedge angle, θ_w , is 2π rather than $3\pi/2$. An additional ground bounce in between the two plates is introduced to each of the four ray paths of the wide barrier, effectively doubling the number of possible ray paths to eight. Multiple diffraction effects beyond double diffraction, e.g., multiple scatter between the barriers, have been neglected.

The coordinates of the virtual receiver and secondary sources depend on the path taken. Paths that include the ground bounce between the two plates of the double knife-edge barrier use a virtual receiver at coordinates $(r', \theta', 0)$ where $r' = [W^2 + (2h)^2]^{1/2} + r$ and $\theta' = 2\pi - \arctan(W/2h)$ and secondary sources at (r'_0, θ'_0, z_n) where $r'_0 = [W^2 + (2h)^2]^{1/2}$ and $\theta'_0 = \arctan(W/2h)$. Paths that do not include the ground bounce have secondary source and virtual receiver coordinates that are the same as those of the wide barrier with the exception of θ'_0 , which is $\pi/2$. In using Eq. (2) and Eq. (4) to calculate the secondary source strengths,

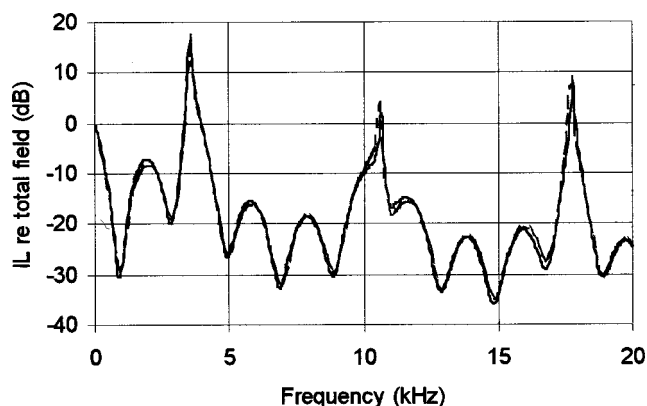


FIG. 4. Insertion loss, normalized to the total field, of a wide barrier as predicted by the BTM (solid line) and Pierce (dashed line) models. For this geometry, $W=20$ cm, $x_s=40$ cm, $y_s=15$ cm, $x_r=70$ cm, $y_r=20$ cm, and $h=30$ cm.

the effect of pressure doubling at the virtual receiver is no longer applicable. In both expressions, the factor of 2 in the denominators is replaced with the usual factor of 4 for spherical spreading.

The metric of interest in this work is the insertion loss or change in sound pressure level with and without a barrier present. Many authors investigate diffraction effects without incorporating ground reflections in either the pre- or post-barrier data sets. Such a measurement does provide a data set with minimal structure and allows for a more convenient comparison between diffraction models but is less ideal for analyzing actual barrier effectiveness. The present work investigates the sound pressure level present at the receiver for the actual source–receiver geometry used, including ground surfaces both when the barrier is present and when it is not. This normalization, with ground reflections included, is often referred to as the total field as opposed to the free field which refers to the direct signal only. Including ground bounces in the analysis provides a direct measure of a barrier's effectiveness including the possibility of increasing sound levels by removing the destructive interference of the ground bounce.¹⁷ Furthermore a “–” sign to force the loss to be a positive quantity was not included in order to compare with Medwin's data in the initial calculations (data not shown). Thus a negative value for the insertion loss indicates a decreased sound level and a positive value indicates an increased sound level.

Pierce's three sided barrier model¹⁶ was selected for comparison to the BTM model for the wide barrier. Using the Pierce model, values for the insertion loss were calculated every 100 Hz, while values for the insertion loss from the BTM model are determined from a digital transform of the sound field calculated in the time domain. Figure 4 shows the insertion loss of a wide barrier on a rigid ground surface. The agreement between these two models is excellent. It should be noted, however, that while the Pierce model is well suited to predict multiple diffraction effects, it requires adjacent diffracting edges to have a common side plane and therefore cannot be utilized as presented to examine the double knife-edge barrier. A comparison to extensions of the

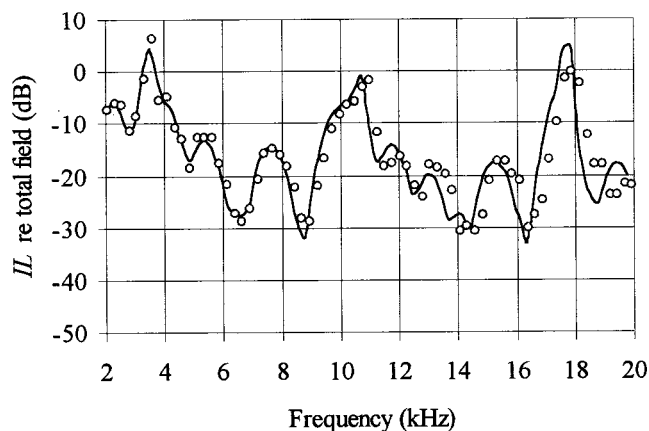


FIG. 5. Single knife-edge barrier insertion loss for $x_s=59.5$ cm, $y_s=15$ cm, $x_r=60.5$ cm, $y_r=20$ cm, and $h=31.7$ cm. The solid line is the BTM model prediction and the open circles are experimental data.

Pierce model for multiple barriers³ will be left to future work.

II. EXPERIMENT

Experiments were conducted to determine the insertion loss of three types of barriers: (1) the single knife-edge, (2) the wide, and (3) the double knife-edge. Since one of the goals of the work was to have an application to highway noise barriers, source and receiver coordinates were chosen to roughly correspond to typical test configurations used by the Federal Highway Administration.¹⁸ Barrier heights were chosen to correspond to the national mean barrier height as reported by Armstrong.¹⁹ These dimensions were scaled at 1:10 in order to be able to perform the experiments in a well-controlled indoor environment.

The sound source was produced by the discharge of an electric spark which generated a transient pulse of duration $\tau=100$ μ s. At 1.2 m, the signal-to-noise ratio was 40 dB \pm 2 dB in the range of 2 kHz to 20 kHz. Given the scaling factor, the experiments corresponded to a real world problem of 12 m propagation over a 3 m barrier in the frequency range of 200 Hz to 2 kHz. Since the spark source was non-

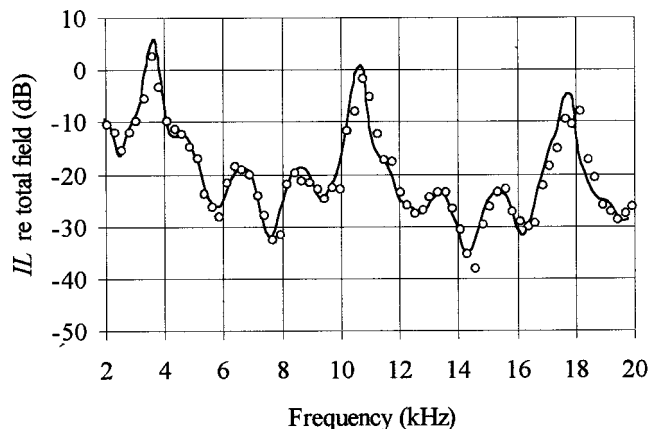


FIG. 6. Wide barrier insertion loss for $W=18.9$ cm, $x_s=50$ cm, $y_s=15$ cm, $x_r=70$ cm, $y_r=20$ cm, and $h=31.7$ cm. The solid line is the BTM model prediction and the open circles are experimental data.

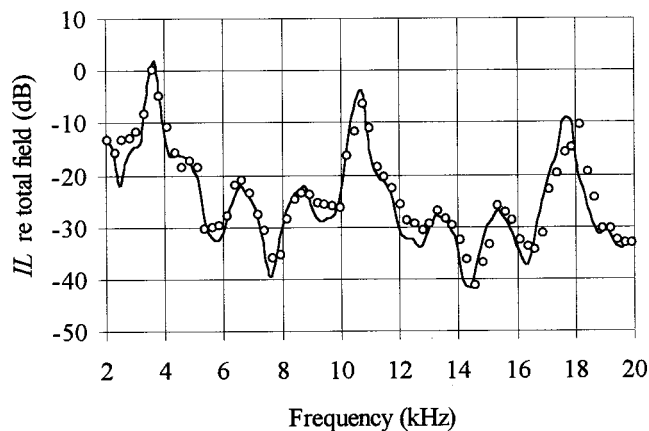


FIG. 7. Double knife-edge barrier insertion loss for $W=18.9$ cm, $x_s=50$ cm, $y_s=15$ cm, $x_r=70$ cm, $y_r=20$ cm, and $h=31.7$ cm. The solid line is the BTM model prediction and the open circles are experimental data.

repeatable, a reference microphone was used to calibrate each measurement. The source was found to be omnidirectional within ± 1.7 dB in angles of $\pm 50^\circ$ in both the horizontal and vertical planes. The direct sound pressure field was observed to decay as $1/r$ and nonlinear effects were negligible. All surfaces, with the exception of the knife-edge barriers, were constructed with smoothed, varnished wood.

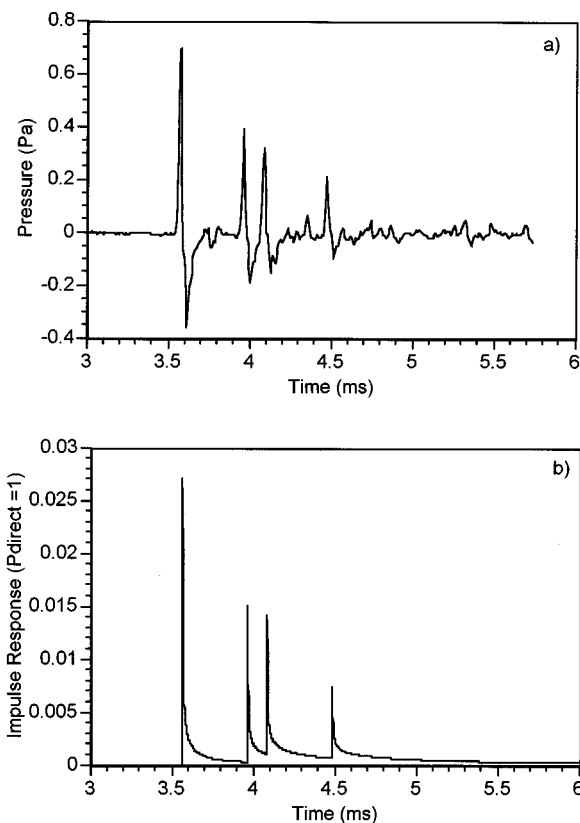


FIG. 8. Single knife-edge barrier (a) time domain data and (b) calculated impulse response for $x_s=59.5$ cm, $y_s=15$ cm, $x_r=60.5$ cm, $y_r=20$ cm, and $h=31.7$ cm. The impulse response has been normalized such that the direct signal impulse response (not present due to barrier shielding) has an amplitude of 1. The four arrivals are from paths SER , $SGER$, $SEGR$, $SGEGR$, where S is the source, G is the ground, E is the barrier edge and R is the receiver.

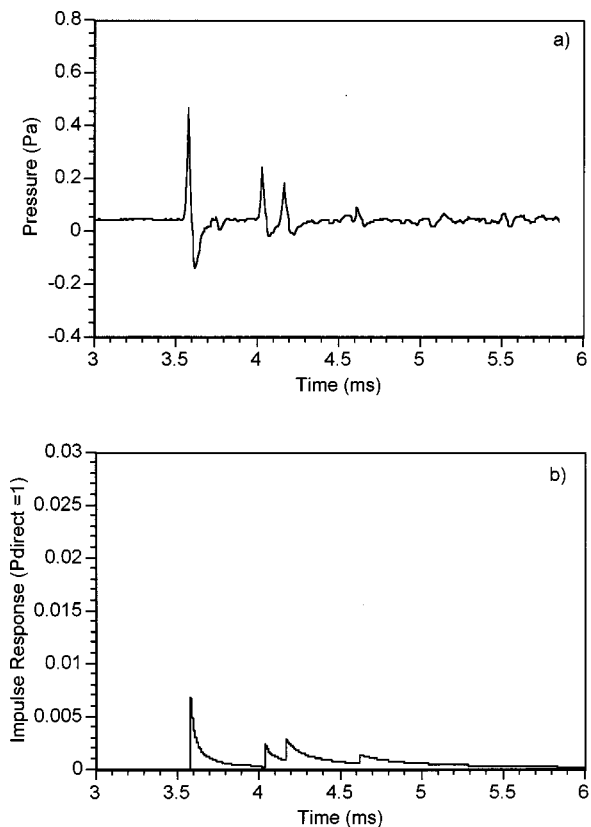


FIG. 9. Wide barrier (a) time domain data and (b) calculated impulse response for $W=18.9$ cm, $x_s=50$ cm, $y_s=15$ cm, $x_r=70$ cm, $y_r=20$ cm, and $h=31.7$ cm. The impulse response has been normalized such that the direct signal impulse response (not present due to barrier shielding) has an amplitude of 1. The four arrivals are from paths SE_1E_2R , SGE_1E_2R , SE_1E_2GR , SGE_1E_2GR , where S is the source, G is the ground, E_1 is the first barrier edge, E_2 is the second barrier edge and R is the receiver.

The knife-edge barriers were made from steel plates of thickness 3 mm. The receivers were $\frac{1}{4}$ in. condenser microphones with their diaphragms always oriented in a plane parallel to the least time propagation path in order to minimize the effects of directivity. The time domain signal was captured and transformed via FFT by an HP 35665A signal analyzer. The time domain signals and the frequency transforms were transferred to a PC for analysis and comparison to theory.

III. RESULTS AND COMPARISONS

The insertion loss data from the series of experiments were compared to the BTM predicted insertion loss for the three types of barriers investigated in this work. In each case, the receivers were in the shadow zones of the barriers and the diffracted sound field was normalized to the sound field for the same source–receiver geometry in the absence of the barriers. Several configurations of source and receiver positions relative to the barrier were tested.²⁰ Figures 5–7 show insertion loss curves for the single knife-edge, wide and double knife-edge barriers, respectively, for one such configuration. Each barrier was 31.7 cm high; the source height was 15 cm and the receiver height was 20 cm. For the single knife-edge barrier, the source was 59.5 cm from the barrier and the receiver was 60.5 cm from the barrier. For the wide and double knife-edge barriers, from Fig. 2, $x_s=50$ cm, y_s

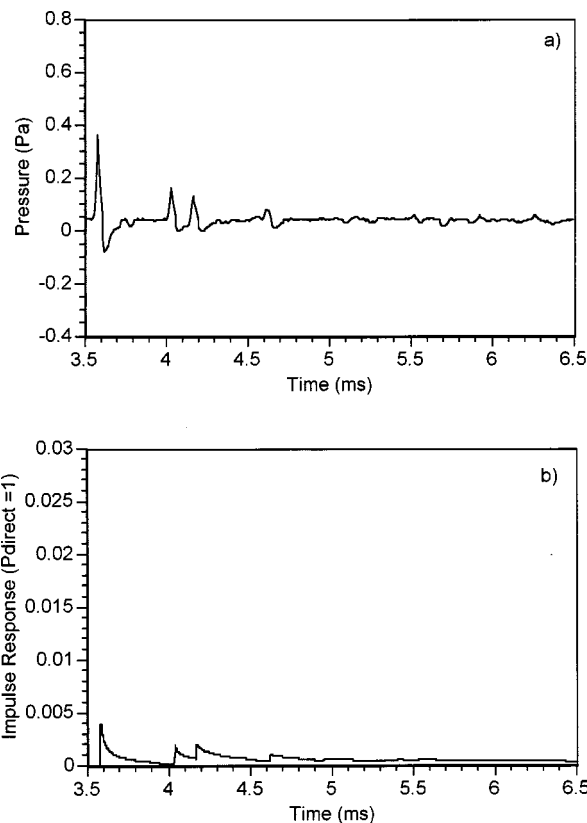


FIG. 10. Double knife-edge barrier (a) time domain data and (b) calculated impulse response for $W=18.9$ cm, $x_s=50$ cm, $y_s=15$ cm, $x_r=70$ cm, $y_r=20$ cm, and $h=31.7$ cm. The impulse response has been normalized such that the direct signal impulse response (not present due to barrier shielding) has an amplitude of 1. The eight arrivals are from paths SE_1E_2R , SGE_1E_2R , SE_1E_2GR , SGE_1E_2GR , SE_1GE_2R , SGE_1GE_2R , SE_1GE_2GR , SGE_1GE_2GR , where S is the source, G is the ground, E_1 is the first barrier edge, E_2 is the second barrier edge and R is the receiver.

$=70$ cm, and $W=18.9$ cm. The solid lines in the figures represent the BTM predicted insertion losses, and the dots represent the average of five firings of the spark for that configuration. The three prominent peaks in each of the insertion loss curves indicate the removal of the destructively interfering bounce path from the preinsertion sound field.

Figures 8–10 show the time domain data traces and calculated impulse responses for the knife-edge, wide and double knife-edge barriers respectively. Figure 11 shows the acoustic signal with no barrier present which is used as the reference signal. The acoustic data have been presented in units of pressure (Pa). The impulse response has been normalized such that the direct signal impulse response (not present due to barrier shielding) has an amplitude of 1. Thus, the free field signal from the spark could be convolved with the impulse response presented to reproduce the barrier data acquired (figure not shown).

The first four dominant arrivals are clearly identifiable in the data as the model impulse response aligns with the experimental data. Some late acoustic arrivals are evident but analysis indicated that they were weak reflections from the spark support system which can be seen in Fig. 11 where two weak signals follow the dominant direct and ground reflected signals. While these signals could have been windowed out of the reference signal they are embedded in the multiple

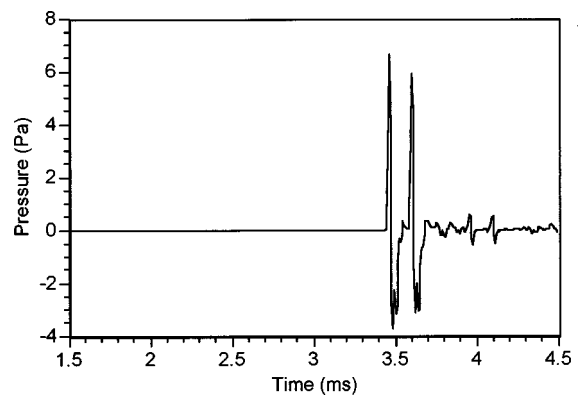


FIG. 11. Time domain data without a barrier present for $x_s=59.5$ cm, $y_s=15$ cm, $x_r=60.5$ cm, and $y_r=20$ cm.

diffraction data and were consequently retained as part of the reference signal. For the double knife-edge barrier data shown in Fig. 10, the last four arrivals, corresponding to ray paths that include a ground bounce between barriers, are not at all discernible in the data. Later calculations in which they were removed (data not shown) indicated that they had a negligible influence on the insertion loss. For shallower barriers or wider separations this observation may not hold.

IV. CONCLUSIONS

Excellent agreement is found in the audible spectrum between the BTM predictions and the experimental data for all three types of barriers included in this study. These findings should provide a good benchmark for the continued development of frequency domain or empirical models of multiple diffraction which may be incorporated into noise prediction models. The model is also capable of analyzing transient phenomena and finite length barriers.

An interesting observation is that the contributions from the ray paths containing the ground bounce between the two plates of the double knife-edge barrier appear to be negligible for the geometry examined. This can be seen from the similarities between the insertion losses of the wide and double knife-edge barriers. However, the double knife-edge barrier is seen to reduce the sound level by an additional 3–5 dB over a wide barrier of the same dimensions, which is attributed to the change in wall angle from $3\pi/2$ to 2π . This observation could lead to improvements in computation time by discarding negligible contributions to the diffracted field.

ACKNOWLEDGMENT

This work was supported by the Federal Highway Administration (FHWA).

¹T. Isei, T. F. W. Embleton, and J. E. Piercy, "Noise reduction by barriers on finite impedance ground," *J. Acoust. Soc. Am.* **67**, 46–58 (1980).

²J. Nicolas, T. F. W. Embleton, and J. E. Piercy, "Precise model measurements versus theoretical prediction of barrier insertion loss in presence of the ground," *J. Acoust. Soc. Am.* **73**, 44–54 (1983).

³E. M. Salomons, "Sound propagation in complex outdoor situations with a non-refracting atmosphere: model based on analytical solutions for diffraction and reflection," *Acustica* **83**, 436–454 (1997).

⁴F. Oberhettinger, "On the diffraction and reflection of waves and pulses by wedges and corners," *J. Res. Natl. Bur. Stand.* **61**, 343–365 (1958).

- ⁵M. A. Biot and I. Tolstoy, "Formulation of wave propagation in infinite media by normal coordinates with an Application to diffraction," J. Acoust. Soc. Am. **29**, 381–391 (1957).
- ⁶R. Raspet, J. Ezell, and S. Coggeshall, "Diffraction of an explosive transient," J. Acoust. Soc. Am. **79**, 1326–1334 (1986).
- ⁷H. Medwin, "Shadowing by finite noise barriers," J. Acoust. Soc. Am. **69**, 1060–1064 (1981).
- ⁸J. P. Chambers and Y. H. Berthelot, "Time-domain experiments on the diffraction of sound by a step discontinuity," J. Acoust. Soc. Am. **96**, 1887–1892 (1994).
- ⁹A. I. Papadopoulos and C. G. Don, "A study of barrier attenuation by using acoustic impulses," J. Acoust. Soc. Am. **90**, 1011–1018 (1991).
- ¹⁰D. Chu, "Impulse response of density contrast wedge using normal coordinates," J. Acoust. Soc. Am. **86**, 1883–1896 (1989).
- ¹¹D. Chu, "Exact solution for a density contrast shallow-water wedge using normal coordinates," J. Acoust. Soc. Am. **87**, 2442–2450 (1990).
- ¹²S. Li and C. S. Clay, "Sound transmission experiments from an impulsive source near rigid wedges," J. Acoust. Soc. Am. **84**, 2135–2143 (1988).
- ¹³S. Li, D. Chu, and C. S. Clay, "Time domain reflections and diffractions from facet-wedge constructions: Acoustic experiments including double diffractions," J. Acoust. Soc. Am. **96**, 3715–3720 (1994).
- ¹⁴H. Medwin and C. S. Clay, *Fundamentals of Acoustical Oceanography* (Academic, New York, 1997), p. 712.
- ¹⁵H. Medwin, E. Childs, and G. M. Jebsen, "Impulse studies of double diffraction: A discrete Huygens interpretation," J. Acoust. Soc. Am. **72**, 1005–1013 (1982).
- ¹⁶A. D. Pierce, "Diffraction of sound around corners and over wide barriers," J. Acoust. Soc. Am. **55**, 941–955 (1974).
- ¹⁷H. G. Jonasson, "Sound reduction by barriers on the ground," J. Sound Vib. **22**, 113–126 (1972).
- ¹⁸C. Menge, C. Rossano, G. Anderson, and C. Bajdek, *FHWA Traffic Noise Model Version 1.0 Technical Manual* (1995), p. 24.
- ¹⁹R. Armstrong, "Highway traffic noise barrier construction trends," Wall Journal **26**, 12–19 (1996).
- ²⁰G. J. Wadsworth, "Scale model experiments on the insertion loss of wide and double barriers," M.S. Thesis, University of Mississippi (1998).

“Fast” quasilongitudinal sagittally polarized surface waves in layer-substrate structures

A. N. Darinskii

Institute of Crystallography, Academy of Sciences of Russia, Leninskii pr. 59, Moscow 117333, Russia

I. S. Didenko and N. F. Naumenko

Moscow Steel and Alloys Institute, Leninskii pr. 4, Moscow 117936, Russia

(Received 14 September 1999; accepted for publication 28 January 2000)

The propagation of surface sagittally polarized waves on the plated surface of a semi-infinite medium with the phase velocity exceeding that of bulk quasitransverse sagittally polarized waves in the substrate is studied analytically and numerically. Such a surface wave is shown to exist in nonpiezoelectric and piezoelectric composites provided both the sagittal plane and the interface are planes of elastic symmetry of the layer and substrate. In the case of a nonpiezoelectric structure, the “fast” surface wave is one-component, involving only the quasi-longitudinal nonuniform mode in the substrate. In piezoelectrics this wave is two-component and incorporates the quasilongitudinal mode and the mode of electrical potential in the substrate. The “fast” wave exists at definite values of the velocity v_F and the layer thickness/wavelength ratio $(h/\lambda)_F$. It has been found that one or two surface waves can arise for the same direction of propagation at different v_F and $(h/\lambda)_F$, in the case of two surface solutions, both the values of $(h/\lambda)_F$ being small. The approximate expressions are derived which allow one to elucidate whether the structure under consideration can support the “fast” surface wave and to estimate the corresponding v_F and $(h/\lambda)_F$. A comparison is performed of numerical and analytical results. © 2000 Acoustical Society of America. [S0001-4966(00)00305-2]

PACS numbers: 43.20.Hq [ANN]

INTRODUCTION

The rapid development of electronics in the past few decades has provoked a growing interest in acoustics of layered media. The possibility of controlling the parameters of surface waves by means of deposition of a film on the substrate makes solid layered composites attractive for use in acoustoelectronic devices.

The theoretical analysis of wave solutions on the plated surface of a solid of generic anisotropy meets serious difficulties even if a one-layered structure is under investigation. The wave field in composites involves a great deal of partial plane modes coupled via the boundary conditions. Therefore, one usually studies isotropic media or assumes that the sagittal plane is a plane of symmetry of elastic properties. In either case the set of boundary equations decomposes into two subsets. One of them describes Love waves.^{1–4} The Love wave velocity is always less than the velocity of the bulk shear (SH) wave v_{SH} in the substrate. If the substrate and layer are isotropic, then the Love wave exists provided $v_{SH} > v_{SH}^{(f)}$, where $v_{SH}^{(f)}$ is the velocity of the bulk SH wave in the film deposited. However, when the substrate is coated with a piezoelectric film, v_{SH} need not be greater than $v_{SH}^{(f)}$ for the Love wave to appear (Ref. 5).

The second subset describes the propagation of sagittally polarized waves. In the substrate, such a surface wave involves nonuniform modes which can be viewed as continuations of quasilongitudinal (QL) and quasitransverse (QT) bulk modes onto the range $v < v_{QL}, v_{QT}$, where v_{QL} and v_{QT} are the minimum (limiting) velocities of bulk QL and QT waves along the surface of the substrate, respectively. Ac-

cordingly, in order to distinguish between the two types of modes we shall refer to them as QL and QT modes whether they are uniform or nonuniform.

The boundary conditions on the mechanically free surface as well as on the rigid contact between two solids commonly “mix” QL and QT modes. Therefore, it is a common practice to search for sagittally polarized surface waves only within the range $v < v_{QT}$, where both QL and QT modes are nonuniform, see Refs. 1–4, 6, 7. It is likely believed that in the interval $v_{QT} < v < v_{QL}$, perhaps except for extremely special cases, the wave field will necessarily incorporate the bulk reflected QT mode in the substrate so that the wave will be leaky.

To the best of our knowledge, no example of the “fast” pure surface wave in the range $v_{QT} < v < v_{QL}$ on the plated surface has been known until recently. Apparently Didenko *et al.*⁸ were the first to discover such a wave in the composite constituted of two piezoelectrics of symmetry 6 mm with the sixfold axes perpendicular to the interface, SiC (substrate) and ZnO (film). Later an analogous surface solution was obtained numerically for a ZnO-diamond composite.⁹ In particular, Didenko *et al.*⁸ computed the imaginary component v_l'' of the leaky wave velocity as a function of the film thickness/wavelength ratio (h/λ) and the magnitude of the elastic modulus c_{13} of SiC. It has been found that v_l'' becomes zero at a certain value of h/λ for the preset c_{13} ; the real component v_l' of the leaky wave velocity belongs to the range $v_{QT} < v < v_{QL}$. Obviously, the vanishing of v_l'' indicates that the leaky wave transforms into a pure surface wave. It has also been shown that the appropriate h/λ de-

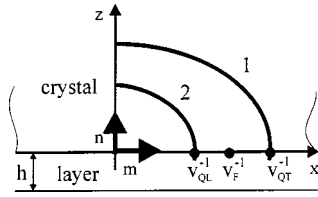


FIG. 1. Slowness curves (one-fourth) of quasitransverse (1) and quasi-longitudinal (2) bulk waves; v_{QT}^{-1} and v_{QL}^{-1} are the inverse velocities of the longitudinal and transverse bulk waves along the x -axis, respectively, v_F^{-1} is the inverse velocity of the “fast” surface wave.

creases with c_{13} of the substrate and tends to zero as $c_{13} \rightarrow 0$ if the piezoeffect in the substrate is ignored. In this limiting situation the surface solution goes into the bulk longitudinal wave that at $c_{13}=0$ satisfies the boundary conditions of a mechanically free surface; such a bulk wave is conventionally referred to as an exceptional bulk wave.¹⁰

The present paper studies the conditions for the existence of QL fast surface waves, such as the permissible crystallographic symmetry of the substrate and the layer, and the relationships between their material constants, and the effect of the electrical boundary conditions in piezoelectric composites. It appears that the existence conditions for these waves are not so stringent as one might think.

I. DISPERSION EQUATION FOR SAGITTALLY POLARIZED WAVES

Let the sagittal plane zx (Fig. 1) be a plane of symmetry of elastic properties both for the substrate and the layer. For the time being we disregard the piezoelectric effect. The contact between the layer and substrate is assumed to be perfectly rigid.

In considering the “fast” surface wave problem, we have two unknown quantities to be found from the boundary conditions. These are the velocity and h/λ ratio (or $H = 2\pi h/\lambda = kh$, where k is the tangential projection of the wave vector). In this section, we analyze the number of equations the velocity and H must obey for the surface wave to exist. Our discussion provides an explanation (but not a rigorous proof) of why the QL sagittally polarized localized wave can appear only when the interface is also the plane of elastic symmetry both for the substrate and layer; if the interface is oriented arbitrarily, then such a wave is unlikely to exist.

The wave equation has four partial solutions of the type

$$\mathbf{u}(\mathbf{r}, t) = \mathbf{A}_\alpha \exp(ik[x + p_\alpha z - vt])$$

at the preset v . These are two QT modes ($\alpha=1,3$) and two QL modes ($\alpha=2,4$). In the interval $v_{QT} < v < v_{QL}$ the QT modes are uniform, whereas the QL modes are nonuniform with $p_2 = p_4^*$. We choose the numbering, where

$$\text{Im}(p_2) > 0; \quad \alpha=1 \quad \text{is the reflected bulk wave.} \quad (1)$$

To find the parameters of partial modes, one commonly solves the Green-Christoffel equation, see, e.g., Ref. 2. However, there is an alternative way based on the Stroh eigenvalue problem (Refs. 11 and 12),

$$\hat{\mathbf{N}} \xi_\alpha = p_\alpha \xi_\alpha, \quad (2)$$

where $\hat{\mathbf{N}}$ is a 4×4 real matrix, ξ_α is a four-component vector column constructed from the polarization vector \mathbf{A}_α and traction \mathbf{L}_α of a mode α ,

$$\hat{\mathbf{N}} = - \begin{Bmatrix} \hat{\mathbf{N}}_{11} & \hat{\mathbf{N}}_{12} \\ \hat{\mathbf{N}}_{21} + \rho v^2 \hat{\mathbf{I}} & (\hat{\mathbf{N}}_{11})^t \end{Bmatrix}, \quad \xi_\alpha = \begin{pmatrix} \mathbf{A}_\alpha \\ \mathbf{L}_\alpha \end{pmatrix}, \quad (3)$$

and p_α is the decay parameter. In (3), $\hat{\mathbf{N}}_{11} = (nn)^{-1}(nm)$, $\hat{\mathbf{N}}_{12} = (nn)^{-1}$, and $\hat{\mathbf{N}}_{21} = (mn)(nn)^{-1}(nm) - (mm)$, where (mm) , (nm) , and (mn) are the 2×2 matrices with elements $(mm)_{ij} = c_{i11j}$, $(nm)_{ij} = (mn)_{ji} = c_{i31j}$, $i, j = 1, 3$ the matrix $(nn)^{-1}$ is the inverse of the matrix $(nn)_{ij} = c_{i33j}$, $\hat{\mathbf{I}}$ is the 2×2 unit matrix, ρ in $\hat{\mathbf{N}}$ is the density of the medium, and the symbol $(\dots)^t$ stands for transposition. In our case, $\hat{\mathbf{N}}$ (3) is the “reduced” Stroh matrix obtained from the “full” 6×6 Stroh matrix^{11,12} by deleting two rows and two columns associated with SH partial modes.

Unless otherwise stated, we use Stroh’s normalization condition for ξ_α ,

$$\xi_\alpha \cdot \hat{\mathbf{T}} \xi_\beta = \delta_{\alpha\beta}, \quad \hat{\mathbf{T}} = \begin{pmatrix} \hat{\mathbf{O}} & \hat{\mathbf{I}} \\ \hat{\mathbf{I}} & \hat{\mathbf{O}} \end{pmatrix}, \quad (4)$$

where $\delta_{\alpha\beta}$ is the Kronecker symbol, $\hat{\mathbf{O}}$ is the zero matrix, respectively. The orthogonality $\xi_\alpha \cdot \hat{\mathbf{T}} \xi_\beta = 0$, $\alpha \neq \beta$, follows from the relation $(\hat{\mathbf{T}} \hat{\mathbf{N}})^t = \hat{\mathbf{T}} \hat{\mathbf{N}}$. If Stroh’s normalization is accepted, then ξ_α ’s corresponding to the reflected and incident bulk waves are purely imaginary and real, respectively, providing a coordinate axis is directed along the internal normal to the surface of the medium (Fig. 1).¹⁰ As to the vectors ξ_α associated with complex conjugate p_α , we introduce them complex conjugate. Summed up, in view of (1) for the plane modes in the substrate we have

$$\xi_1 = -\xi_1^*, \quad \xi_3 = \xi_3^*, \quad \xi_2 = \xi_4^*. \quad (5)$$

Let us discuss the number of equations for the surface wave velocity and the appropriate H in the range $v_{QT} < v < v_{QL}$. The boundary conditions in the composite (the continuity of the displacement and traction at $y=0$ and the vanishing of the traction at $y=-h$) can be written in the form $\sum_\alpha b_\alpha \mathbf{L}_{h\alpha} = 0$, where the summation is carried out over the substrate modes involved in the problem, b_α is the partial amplitude of the mode α , and the “effective” traction $\mathbf{L}_{h\alpha}$ is composed of the last two components of the vector

$$\xi_{h\alpha} = e^{-iH \hat{\mathbf{N}}_f} \xi_\alpha = \left(\sum_{\alpha=1}^4 e^{-ip_{f\alpha} H} \xi_{f\alpha} \otimes \hat{\mathbf{T}} \xi_{f\alpha} \right) \xi_\alpha, \quad (6)$$

where $\hat{\mathbf{N}}_f$ is the 4×4 Stroh matrix for the film, $(\xi_{f\alpha}, p_{f\alpha})$ are eigensolutions of Stroh’s eigenvalue problem for $\hat{\mathbf{N}}_f$ and associated with plane modes in the film, and the symbol \otimes denotes dyadic multiplication. The spectral decomposition of the propagator $\exp(-iH \hat{\mathbf{N}}_f)$ in the form (6) stems from relation (4) applied to the vectors $\xi_{f\alpha}$.

By (1), the surface wave may involve only the mode $\alpha=2$ in the substrate so that the boundary conditions merely reduce to $\mathbf{L}_{h2} = 0$. Thus, there are in a general manner four equations for v_F and H , because \mathbf{L}_{h2} is a complex two-

component vector. However, due to certain relations the tractions of plane modes obey it can be shown that v_F and H corresponding to the surface wave are less severely conditioned.

Indeed, consider the equation for the leaky wave velocity v_l ,

$$||\mathbf{L}_{h1}\mathbf{L}_{h2}|| \equiv (\mathbf{L}_{h1})_x(\mathbf{L}_{h2})_z - (\mathbf{L}_{h1})_z(\mathbf{L}_{h2})_x = 0. \quad (7)$$

The leaky wave incorporates the nonuniform mode $\alpha=2$ and the reflected mode $\alpha=1$ in the substrate. Its velocity $v_l = v'_l - iv''_l$ is complex valued ($v_{QT} < v'_l < v_{QL}$, $v''_l > 0$). But if a root of (7) becomes real, then $\mathbf{L}_{h2}=0$ rather than $||\mathbf{L}_{h1}||\mathbf{L}_{h2}$. This fact is obvious from the law of energy conservation; moreover, using relations (9) and (13), see below, one can formally prove that \mathbf{L}_{h1} cannot be collinear to \mathbf{L}_{h2} at real v from the range $v_{QT} < v < v_{QL}$.

So, a real root of Eq. (7) corresponds to the surface solution. This equation seemingly imposes two conditions upon v_F and H , $\text{Re}(||\mathbf{L}_{h1}\mathbf{L}_{h2}||)=0$, $\text{Im}(||\mathbf{L}_{h1}\mathbf{L}_{h2}||)=0$. However, in fact this is not necessarily the case. Let us introduce the vector columns

$$\eta_{h\alpha} \equiv (\mathcal{A}_\alpha, \mathcal{L}_\alpha)^t = e^{iH\hat{\mathbf{N}}_f} \xi_\alpha. \quad (8)$$

Due to (4) and the relation $(\hat{\mathbf{T}}\hat{\mathbf{N}}_f)^t = \hat{\mathbf{T}}\hat{\mathbf{N}}_f$,

$$\eta_{h\alpha} \cdot \hat{\mathbf{T}} \xi_{h\alpha} = \delta_{\alpha\beta}, \quad (9)$$

and then

$$\sum_{\alpha=1}^4 \eta_{h\alpha} \otimes \hat{\mathbf{T}} \xi_{h\alpha} = \begin{pmatrix} \hat{\mathbf{I}} & \hat{\mathbf{O}} \\ \hat{\mathbf{O}} & \hat{\mathbf{I}} \end{pmatrix}, \quad (10)$$

whence it follows that

$$\sum_{\alpha=1}^4 \mathcal{L}_{h\alpha} \otimes \mathbf{L}_{h\alpha} = \hat{\mathbf{O}}. \quad (11)$$

Let $\text{Re}(||\mathbf{L}_{h1}\mathbf{L}_{h2}||)=0$. Hence,

$$i \text{Im}(||\mathbf{L}_{h1}\mathbf{L}_{h2}||) = ||\mathbf{L}_{h1}\mathbf{L}_{h2}|| = - (||\mathbf{L}_{h1}\mathbf{L}_{h2}||)^* = ||\mathcal{L}_{h1}\mathcal{L}_{h4}||, \quad (12)$$

since owing to (5) for real v ,

$$\xi_{h1} = -\eta_{h1}^*, \quad \xi_{h3} = \eta_{h3}^*, \quad \xi_{h2} = \eta_{h4}^*, \quad \xi_{h4} = \eta_{h2}^*. \quad (13)$$

Multiplying identity (11) by the vector $\mathcal{L}_{h1}^\times = ((\mathcal{L}_{h1})_z, -(\mathcal{L}_{h1})_x)^t$ from the left and by the vector $\mathbf{L}_{h1}^\times = ((\mathbf{L}_{h1})_z, -(\mathbf{L}_{h1})_x)^t$ from the right we obtain, accounting for (12) and (13),

$$||\mathbf{L}_{h1}\mathbf{L}_{h3}||^2 \equiv -2 \text{Im}(||\mathbf{L}_{h1}\mathbf{L}_{h2}||) \text{Im}(||\mathbf{L}_{h1}\mathbf{L}_{h4}||). \quad (14)$$

As a result, the vanishing of $\text{Im}(||\mathbf{L}_{h1}\mathbf{L}_{h2}||)$ implies $||\mathbf{L}_{h1}\mathbf{L}_{h3}||=0$. In the general case $||\mathbf{L}_{h1}\mathbf{L}_{h3}||$ is a complex function with independent real and imaginary parts [we cannot prove that $\text{Re}(||\mathbf{L}_{h1}\mathbf{L}_{h3}||)$ and $\text{Im}(||\mathbf{L}_{h1}\mathbf{L}_{h3}||)$ are generally somehow dependent functions]. Therefore, $\text{Re}(||\mathbf{L}_{h1}\mathbf{L}_{h3}||)=0$ and $\text{Im}(||\mathbf{L}_{h1}\mathbf{L}_{h3}||)=0$ should be viewed as independent equations so that (7) imposes three rather than two conditions upon v_F and H .

Thus, v_F and H corresponding to a QL surface wave must obey at least three equations. In other words, the num-

ber of variables is less than that of relationships to be fulfilled. Hence, the conjecture that in a general manner such waves do not exist seems to be reasonable.

Note that in the range $v < v_{QT}$ there are two nonuniform modes ($\alpha=1,2$) for surface wave construction. Using (11) one can prove that the boundary equation $||\mathbf{L}_{h1}\mathbf{L}_{h2}||=0$ imposes only one condition on v in this interval. Therefore, it is quite natural that surface solutions commonly exist in the interval $v < v_{QT}$, their velocities being functions of H .

Now let us show that if the interface xy is a plane of symmetry both for the layer and substrate, then $||\mathbf{L}_{h1}\mathbf{L}_{h3}||$ is a purely imaginary function in the interval $v_{QT} < v < v_{QL}$. In this symmetric configuration the equation for the parameter p_α is biquadratic, i.e., both p_α and $-p_\alpha$ are solutions. One can then appreciate, even without calculating explicitly \mathbf{A}_α and \mathbf{L}_α , that the Stroh-normalized eigenvectors ξ_α of the matrix $\hat{\mathbf{N}}$ [see (4) with $c_{15}=c_{35}=0$] obey the following relations.

p_α corresponds to a reflected mode,

$$\xi_\alpha = i(A_{\alpha x}, A_{\alpha z}, L_{\alpha x}, L_{\alpha z})^t, \quad p_\beta = -p_\alpha, \quad \xi_\beta = (-A_{\alpha x}, A_{\alpha z}, L_{\alpha x}, -L_{\alpha z})^t. \quad (15)$$

p_α is purely imaginary valued,

$$\xi_\alpha = e^{-i\pi/4}(ia_{\alpha x}, a_{\alpha z}, l_{\alpha x}, il_{\alpha z})^t, \quad p_\beta = p_\alpha^* = -p_\alpha, \quad \xi_\beta = \xi_\alpha^*. \quad (16)$$

Here $A_{\alpha x}$, etc., in (15) and $a_{\alpha x}$, etc., in (16) are real valued quantities. It may be that the biquadratic equation has two complex conjugate roots, p_α^2 and $p_\beta^2 = (p_\alpha^*)^2$. In this instance,

$$p_1 = a + ib, \quad \xi_1 = (A_{1x}, A_{1z}, L_{1x}, L_{1z})^t, \quad p_2 = -a + ib, \quad \xi_2 = i(A_{1x}^*, -A_{1z}^*, -L_{1x}^*, L_{1z}^*)^t, \quad (17)$$

$$p_{\alpha+2} = p_\alpha^*, \quad \xi_{\alpha+2} = \xi_\alpha^*, \quad \alpha = 1, 2.$$

In (17) A_{1x} , etc., are complex valued.

Making use of the spectral representation for $\exp(-iH\hat{\mathbf{N}}_f)$, applying Eqs. (15)–(17) to the eigensolutions of $\hat{\mathbf{N}}_f$, we obtain that for all possible alternatives the real and imaginary parts of the propagator are of the form

$$\text{Re} = \begin{Bmatrix} \times & 0 & 0 & \times \\ 0 & \times & \times & 0 \\ 0 & \times & \times & 0 \\ \times & 0 & 0 & \times \end{Bmatrix}, \quad \text{Im} = \begin{Bmatrix} 0 & \times & \times & 0 \\ \times & 0 & 0 & \times \\ \times & 0 & 0 & \times \\ 0 & \times & \times & 0 \end{Bmatrix}, \quad (18)$$

where the crosses \times denote nonzero elements. Combining (15) and (18) one can readily check that

$$(\mathbf{L}_{h1})_x \equiv i[(\mathbf{L}_{h3})_x]^*, \quad (\mathbf{L}_{h1})_z \equiv -i[(\mathbf{L}_{h3})_z]^*, \quad (19)$$

whence it follows that $||\mathbf{L}_{h1}\mathbf{L}_{h3}||$ is purely imaginary valued, which completes the proof.

Thus, with regard to (14), $||\mathbf{L}_{h1}\mathbf{L}_{h2}||=0$ can be thought to impose only two conditions upon v_F and H in the symmetric configuration.

It is straightforward to check that for this configuration the boundary equation taken in the form $\mathbf{L}_{h2}=0$ also reduces to two conditions upon v_F and H . Indeed, multiplying ξ_2 by the propagator yields, accounting for (16) and (18), $(L_{h2})_x = e^{-i\pi/4}L_x$ and $(L_{h2})_z = e^{i\pi/4}L_z$, where L_x and L_z are real valued quantities. Hence, $\mathbf{L}_{h2}=0$ is equivalent to $L_x=L_z=0$.

To summarize, when the interface is a plane of elastic symmetry of the layer and substrate, v_F and H are subject to only two equations. Therefore, there are strong grounds to believe that for such a configuration the surface wave in the range $v_{QT} < v < v_{QL}$ can exist at particular v_F and H . Of course, the final conclusion can be made only after the solvability of the boundary equations is analyzed.

II. EXISTENCE CONSIDERATIONS FOR ONE-COMPONENT WAVES

We consider the solvability of the boundary equations assuming that the parameter H is small, $H \ll 1$, and that the modulus c_{13} of the substrate is small as compared to c_{33} ; the xy -plane is assumed to be a plane of symmetry. Our investigation will rely on perturbation theory. As a basis for perturbation expansions, we choose the eigenvectors of Stroh's matrix of the substrate at $v=v_{QL}$ and $c_{13}=0$.

With $v=v_{QL}$ (c_{13} need not to be zero), the Stroh matrix becomes non-semi-simple degenerate, Ref. 12. Stroh's eigenvalue problem then reads as

$$\hat{\mathbf{N}}\xi_\alpha = p_\alpha \xi_\alpha, \quad \alpha=1,3, \quad \hat{\mathbf{N}}\xi_{d2}=0, \quad \hat{\mathbf{N}}\xi_{d4}=-\xi_{d2}, \quad (20)$$

where the nondegenerate eigensolutions (ξ_α, p_α) , $\alpha=1,3$ correspond to the QT bulk waves, the degenerate eigensolution $(\xi_{d2}, 0)$ is associated with the bulk longitudinal wave propagating along the x -axis, and ξ_{d4} is the generalized eigenvector. Due to $(\hat{\mathbf{T}}\hat{\mathbf{N}})^t = \hat{\mathbf{T}}\hat{\mathbf{N}}$ it is possible to arrange such that^{13,14}

$$\xi_\alpha \cdot \hat{\mathbf{T}}\xi_\alpha = \xi_{d2} \cdot \hat{\mathbf{T}}\xi_{d4} = 1, \quad \xi_\alpha \cdot \hat{\mathbf{T}}\xi_\beta = \xi_\alpha \cdot \hat{\mathbf{T}}\xi_\gamma = \xi_\gamma \cdot \mathbf{T}\xi_\gamma = 0, \\ \alpha \neq \beta, \quad \alpha=1,3, \quad \gamma=d2,d4. \quad (21)$$

Note that the vectors $\xi_{d2,d4}$ obeying (21) appear to be real. Expanding a vector ξ with respect to ξ_α , $\alpha=1,3,d2,d4$, one obtains

$$\xi = \sum_{\alpha=1,3} (\xi_\alpha \cdot \hat{\mathbf{T}}\xi) \xi_\alpha + (\xi_{d2} \cdot \hat{\mathbf{T}}\xi) \xi_{d4} + (\xi_{d4} \cdot \hat{\mathbf{T}}\xi) \xi_{d2}. \quad (22)$$

Let us ascribe the superscript “(0)” to the eigenvectors of $\hat{\mathbf{N}}$ at $v=v_{QL}$ and $c_{13}=0$. It is not difficult to check that $\mathbf{L}_{d2}^{(0)}=0$, i.e., the longitudinal bulk wave leaves the surface of the substrate free of traction, and that $i\mathbf{L}_1^{(0)} - \mathbf{L}_3^{(0)}=0$. By analogy with Ref. 13, the vectors ξ_α and $\xi_\alpha^{(0)}$ can be interrelated with a real 4×4 matrix $\hat{\mathbf{M}} = \sum_{n=1}^{\infty} (c_{13}/c_{33})^n \hat{\mathbf{M}}^{(n)}$, specifically, $\xi_\alpha = (\hat{\mathbf{I}} + \hat{\mathbf{M}})\xi_\alpha^{(0)}$. The elements of $\hat{\mathbf{M}}$ are computed using perturbation theory.¹³

For $v < v_{QL}$, the vectors ξ_α , $\alpha=2,4$ are sought in the form of expansions with respect to the small quantity $\Delta p = i\sqrt{\rho}|\mathbf{A}_{d2}|(v_{QL}^2 - v^2)^{1/2}$, Refs. 13 and 14,

$$\xi_2 = \xi_{d2} - \Delta p \xi_{d4} + (\Delta p)^2 \sum_{\alpha=1,3,d2} a_\alpha \xi_\alpha + \dots, \quad (23)$$

$$\xi_4 = \xi_{d2} + \Delta p \xi_{d4} + (\Delta p)^2 \sum_{\alpha=1,3,d2} a_\alpha \xi_\alpha + \dots,$$

where a_α are constants; $\xi_{2,4}$ are subject to the relations

$$\xi_2 \cdot \hat{\mathbf{T}}\xi_2 = -2\Delta p, \quad \xi_4 \cdot \hat{\mathbf{T}}\xi_4 = 2\Delta p, \quad \xi_2 \cdot \hat{\mathbf{T}}\xi_4 = 0. \quad (24)$$

The corresponding p_α have the form $p_2 = \Delta p + \sum_{n=2}^{\infty} (\Delta p)^n m_n$ and $p_4 = -\Delta p + \sum_{n=2}^{\infty} (-\Delta p)^n m_n$, where m_n are real constants. By contrast, the vectors $\xi_{1,3}$ as well as $p_{1,3}$ vary as $\delta\xi_{1,3}, \delta p_{1,3} \propto (\Delta p)^2 \propto v_{QL}^2 - v^2$ near v_{QL} .

Consider (7) in the vicinity of v_{QL} . Taking into account the properties of ξ_α and $\xi_\alpha^{(0)}$, expanding the propagator in power series with respect to H , using Eq. (22) and grouping addends, we can bring $||\mathbf{L}_{h1}\mathbf{L}_{h2}||$ into the form

$$||\mathbf{L}_{h1}\mathbf{L}_{h2}|| = -||\mathbf{L}_1\mathbf{L}_{d4}|| \left\{ \Delta p + iH(\hat{N}_f)_{22} + 0.5 \frac{c_{13}^2}{c_{33}^2} (\hat{M}^{(1)})_{2R}^2 + \dots \right\}, \quad (25)$$

where

$$(\hat{N}_f)_{22} = \xi_{d2}^{(0)} \cdot \hat{\mathbf{T}}\hat{\mathbf{N}}_f \xi_{d2}^{(0)} = \frac{c_{11} - c_{55}}{c_{11}c_{55}} N, \quad (26)$$

$$N = c_{11}^{(f)} - \rho^{(f)} v_{QL}^2 - \frac{c_{13}^{(f)2}}{c_{33}^{(f)}},$$

$$(\hat{M}^{(1)})_{2R} = \xi_{d2}^{(0)} \cdot \hat{\mathbf{T}}\hat{\mathbf{M}}^{(1)} \mathbf{R} = \frac{c_{33}}{\sqrt{2(c_{11}c_{33})^{1/2}}} \sqrt{\frac{c_{11} - c_{55}}{c_{11}c_{55}}}, \quad (27)$$

$\mathbf{R} = i\xi_1^{(0)} - \xi_3^{(0)}$ is a real vector, $v_{QL} = (c_{11}/\rho)^{1/2}$. Note that the terms proportional to H^2 , Hc_{13}/c_{33} , and $(\Delta p)^2$ do not appear in (25).

By (25), $\text{Re}\{||\mathbf{L}_{h1}\mathbf{L}_{h2}||\} \approx -||\mathbf{L}_1\mathbf{L}_{d4}||\{\Delta p + iH(\hat{N}_f)_{22} + \dots\}$ vanishes at $\Delta p \approx -iH(\hat{N}_f)_{22} + \dots$ provided $(\hat{N}_f)_{22} < 0$; here we have taken into account the definition of Δp . On the other hand, it is difficult to analyze directly the solubility of the equation $\text{Im}\{||\mathbf{L}_{h1}\mathbf{L}_{h2}||\} = 0$, since $\text{Im}\{||\mathbf{L}_{h1}\mathbf{L}_{h2}||\} = i||\mathbf{L}_1\mathbf{L}_{d4}||\{0.5c_{13}^2/c_{33}^2(\hat{M}^{(1)})_{2R}^2 + \dots\}$ does not involve linear terms.

In this connection we proceed as follows. With regard to (24), instead of (14), one obtains

$$||\mathbf{L}_{h1}\mathbf{L}_{h3}||^2 \equiv \frac{i}{2\Delta p} \text{Im}(|\mathbf{L}_{h1}\mathbf{L}_{h2}|) \text{Re}(|\mathbf{L}_{h1}\mathbf{L}_{h4}|). \quad (28)$$

Due to (23) $\text{Re}(|\mathbf{L}_{h1}\mathbf{L}_{h4}|) \neq 0$ if $\text{Re}(|\mathbf{L}_{h1}\mathbf{L}_{h2}|) = 0$ at least for small H . Hence, the vanishing of the purely imaginary function $||\mathbf{L}_{h1}\mathbf{L}_{h3}||$ is equivalent to $\text{Im}(|\mathbf{L}_{h1}\mathbf{L}_{h2}|) = 0$. Expanding $||\mathbf{L}_{h1}\mathbf{L}_{h3}||$ with respect to c_{13}/c_{33} , H , and $v_{QL}^2 - v^2$ gives after some manipulations,

$$\begin{aligned}
||\mathbf{L}_{h1}\mathbf{L}_{h3}|| = & -||\mathbf{L}_1\mathbf{L}_{d4}|| \left\{ \frac{c_{13}}{c_{33}}(\hat{\mathbf{M}}^{(1)})_{2R} + (\hat{\mathbf{D}})_{2R}\rho(v^2 - v_{\text{QL}}^2) \right. \\
& - 0.5H^2[(\hat{\mathbf{N}}_f)_{22}(\hat{\mathbf{N}}_f)_{4R} - 0.5(\hat{\mathbf{N}}_f)_{2G}(\hat{\mathbf{N}}_f)_{RR}] \\
& \left. + \dots \right\}, \quad (29)
\end{aligned}$$

where $(\hat{\mathbf{D}})_{2R} = \xi_{d2}^{(0)} \cdot \hat{\mathbf{T}}\hat{\mathbf{D}}\mathbf{R}$, $\hat{\mathbf{D}}$ is the real 4×4 matrix relating $\partial \xi_{\alpha}^{(0)} / \partial v^2$ with $\xi_{\alpha}^{(0)}$ at $v = v_{\text{QL}}$ and $c_{13} = 0$, $(\hat{\mathbf{N}}_f)_{4R} = \xi_{d4}^{(0)} \cdot \hat{\mathbf{T}}\hat{\mathbf{N}}_f\mathbf{R}$, $(\hat{\mathbf{N}}_f)_{RR} = \mathbf{R} \cdot \hat{\mathbf{T}}\hat{\mathbf{N}}_f\mathbf{R}$, $(\hat{\mathbf{N}}_f)_{2G} = \xi_{d2}^{(0)} \cdot \hat{\mathbf{T}}\hat{\mathbf{N}}_f\mathbf{G}$, and $\mathbf{G} = i\xi_1^{(0)} + \xi_3^{(0)}$ is a real vector.

From (29) one sees that the equation $||\mathbf{L}_{h1}\mathbf{L}_{h3}|| = 0$ in conjunction with $\text{Re}(||\mathbf{L}_{h1}\mathbf{L}_{h2}||) = 0$ permits us to find c_{13}/c_{33} as a function of H . Substituting $\Delta p \approx -iH(\hat{\mathbf{N}}_f)_{22}$ in (29), using explicit expressions for $\hat{\mathbf{N}}_f$ and the parameters of plane modes, we obtain in the lowest approximation,

$$\frac{c_{13}}{c_{33}} = \Xi H^2 + O(H^4), \quad (30)$$

where

$$\Xi = \frac{c_{11} - c_{55}}{c_{11}^2 c_{55}} N^2 - 0.5 \left\{ \frac{N}{c_{33}} \left(1 - \frac{2\rho^{(f)}}{\rho} \right) + \frac{\rho^{(f)}}{\rho} \frac{c_{13}^{(f)}}{c_{33}^{(f)}} \frac{c_{11}}{c_{33}} \right\}. \quad (31)$$

Note that c_{13}/c_{33} does not depend on H raised in an odd power, since due to (13) $i||\mathbf{L}_{h1}\mathbf{L}_{h3}||$ is an even function of H .

Thus, we are led to the conclusion that the velocity of the QL surface wave and the value of c_{13} , at which this wave exists, can be expressed in terms of H from the boundary equations. However, of practical interest is the question whether, for the given c_{13} , one can find H and v such that the one-component surface wave would exist. According to (30), for small c_{13}/c_{33} this can be done if c_{13} and Ξ (31) are of the same sign; we have then

$$\begin{aligned}
H_F &= \sqrt{\frac{1}{\Xi} \frac{c_{13}}{c_{33}}} \left[1 + O\left(\frac{c_{13}}{c_{33}}\right) \right], \\
v_F &= v_{\text{QL}} \left[1 - \frac{c_{11} - c_{55}}{2c_{11}^2 c_{55}} N^2 \frac{1}{\Xi} \frac{c_{13}}{c_{33}} + O\left(\frac{c_{13}^2}{c_{33}^2}\right) \right]. \quad (32)
\end{aligned}$$

Summed up, with small c_{13}/c_{33} , the one-component surface wave appears in the range $v_{\text{QT}} < v < v_{\text{QL}}$ provided: (1) $N < 0$, see (26); (2) c_{13} and Ξ (31) are of the same sign.

Having in view the above discussion, one can suggest the following “mechanism” of the surface wave formation. Should c_{13} equal zero, the longitudinal bulk wave traveling along the x -axis at $v = v_{\text{QL}}$ would satisfy the boundary conditions on the free surface. Once $c_{13} \neq 0$, the QL mode satisfies the boundary conditions of mechanically free surface only in combination with QT modes; in other words, c_{13} can be said to mix QL and QT modes. The film deposited on the substrate also mixes the modes. Under certain conditions, the two channels of mixing add up destructively, decoupling QL and QT modes at definite v and H and producing the one-component localized wave.

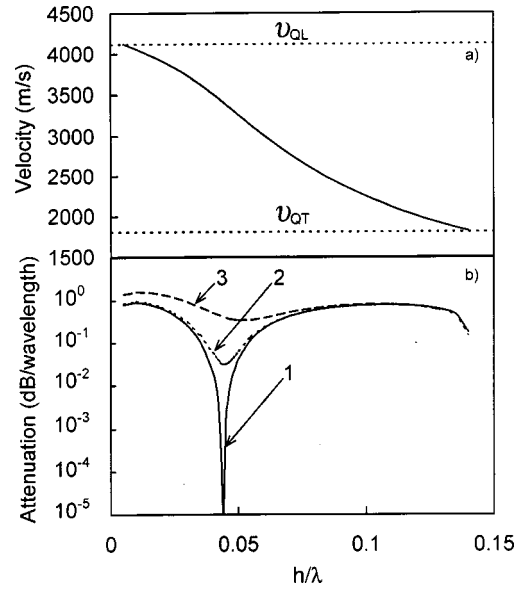


FIG. 2. Velocity v'_l (a) and attenuation δ (b) of the leaky wave on the In/PbS structure, as functions of h/λ . Curve 1 shows the h/λ dependence of δ for the “symmetrical” case. Curves 2 and 3 refer to “nonsymmetrical” orientations with the four-fold axis in the substrate rotated about the y -axis through angles of 3 and 10°, respectively.

Notice that on the mechanically free surface the one-component surface wave arises only under extremely specific conditions; they cannot be fulfilled in the range $v_{\text{QT}} < v < v_{\text{QL}}$, Ref. 15.

III. NUMERICAL RESULTS. COMPARISON WITH ANALYTICAL ESTIMATIONS

According to the data published, in “actual” crystals $|c_{13}/c_{33}| = 0.05 - 0.2$ or greater. Hence, the smallest values of H_F are of order 0.1, i.e., the “small” parameters are not very small in fact. In this connection, it is of special interest to check to what extent the above analytical study is adequate.

As the first example, we consider an In(layer)–PbS structure. PbS is of symmetry $m3m$. It is assumed that the layer is also a single crystal; the In crystal is of symmetry $4/mmm$. The x - and y -axes (in Fig. 1) are directed along the [100] and [001] crystallographic directions both in PbS and In. The material constants are borrowed from Refs. 2 (In) and 16 (PbS).

Equation (7) generally has solutions at complex velocities $v'_l - iv''_l$ associated with leaky waves. Figure 2 shows v'_l [Fig. 2(a)] and the attenuation coefficient δ [Fig. 2(b)] vs h/λ in the interval $v_{\text{QT}} < v < v_{\text{LT}}$ [curve 1 in Fig. 2(b), v''_l is related to δ as $v''_l/v'_l = \delta/40\pi \lg(e) \approx 0.01832\delta$]. Attenuation is seen to be fairly strong, except for a narrow range of h/λ , where δ abruptly decreases and goes to zero. Numerical computations yield $v_F = 3415$ m/s, $(h/\lambda)_F = 0.044$. The estimations made using (32) agree well with numerical findings; $v_F \approx 3480$ m/s, $(h/\lambda)_F \approx 0.048$. The existence conditions for the surface solution are satisfied, i.e., N (26) is negative and c_{13}/Ξ is positive.

Let the substrate be rotated about the y -axis so that the fourfold symmetry axis is not perpendicular to the interface; the plane xz still coincides with a plane of symmetry but the

interface is not a plane of symmetry for the substrate. From Fig. 2(b) (curves 2 and 3) we infer that the sharp minimum on $\delta(h/\lambda)$ quickly disappears as the angle φ between the symmetry axis and the z -axis increases, suggesting the disappearance of the pure surface wave. This favors our guess that the surface wave does not exist unless the interface is a plane of symmetry both for the layer and substrate.

The rotation of the layer also breaks the symmetry required. Notice, however, that for In the moduli c_{11} and c_{33} are very close ($c_{33}/c_{11}=0.996$). Therefore, the tetragonal In layer “behaves” like a cubic medium under rotation about the y -axis, and at $\varphi=\pi/4$ the attenuation of the leaky wave must be nearly zero (according to our computations, δ falls below 10^{-5} dB/ λ).

The surface solution exists in the In–PbS composite if the x -axis is the [110] direction for PbS (in In the fourfold axis is parallel to the z -axis and the x -axis is the [100] direction). The numerically obtained values of v_F and $(h/\lambda)_F$ equal 3039 m/s and 0.052, respectively; from (32) we find $v_F \approx 3110$ m/s and $(h/\lambda)_F \approx 0.056$.

Thus, there are examples of a good agreement between numerical and analytical results. For some other structures, discrepancy between numerical and analytical estimations is more substantial, especially when c_{13}/c_{33} is not small ($c_{13}/c_{33}=0.23$ for PbS). However, of interest is the fact that the analytical description developed in the previous section can happen to be inapplicable even if c_{13}/c_{33} is fairly small.

We have found the following example. Consider four composites: (i) ZnO(b) layer–KH₂PO₄ (KDP) substrate; (ii) ZnO(f) layer–KDP substrate; (iii) ZnO(b) layer–(NH₄)H₂PO₄ (ADP) substrate; (iv) ZnO(f) layer–ADP substrate. The sixfold axis in hexagonal (6 mm) ZnO(b) and ZnO(f) is directed along the z -axis. ZnO(b) and ZnO(f) have the elastic moduli representative of bulk single crystals and of ZnO film,¹⁷ respectively. The piezoeffect in ZnO is not taken into account. KDP and ADP are of symmetry $\bar{4}2m$ and oriented such as the fourfold and twofold axes are directed along the coordinate y - and x -axes, respectively (the xz - and yz -planes are then the plane of elastic symmetry for the substrate and layer, although they are not planes of crystallographic symmetry). The material constants of KDP and ADP are taken from Ref. 18. Since the z -axis is not parallel to the fourfold axis, in our case $c_{11}=c_{33}=c'_{11}$, $c_{13}=c'_{12}$, and $c_{55}=c'_{66}$, where c'_{ij} are the elastic moduli with respect to the crystallographic coordinate frame.

According to Sec. III, the boundary equations are solvable with respect to real v and the modulus of the substrate c_{13} at the preset h/λ providing the quantity N (26) is negative; this condition is satisfied for all the four structures. In Fig. 3 the dependences of c_{13}/c_{33} vs h/λ (curves 1–3) found numerically from (7) are plotted. All other moduli, including c_{33} , were set equal to the appropriate moduli of KDP or ADP. Apparently, the graphs in Fig. 3 show what value c_{13} must have for the one-component surface wave to exist at the preset h/λ .

For the ZnO(b)–KDP composite, c_{13}/c_{33} decreases with increasing h/λ (curve 1) nearly according to parabola (30) (curve 1'). By (30) $v_F \approx 5370$ m/s, $(h/\lambda)_F \approx 0.054$ (c_{13}/c_{33}

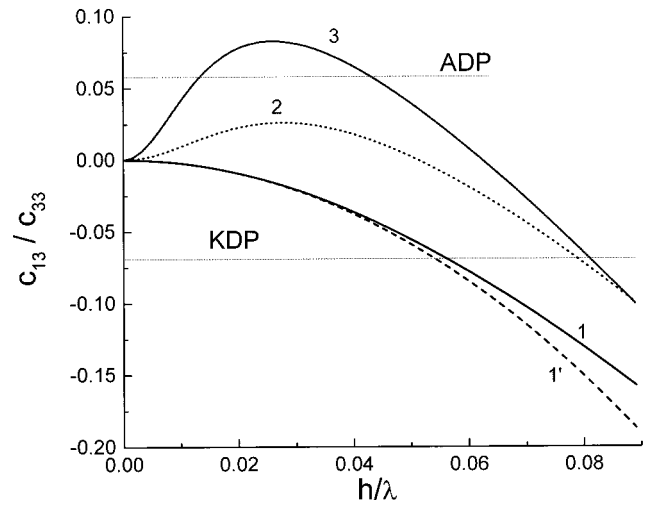


FIG. 3. The magnitude of c_{13}/c_{33} as a function of h/λ obtained from Eq. (7). 1—ZnO(b)–KDP; 2—ZnO(f)–KDP; 3—ZnO(f)–ADP. Curve 1' is parabola (30) for ZnO(b)–KDP. Thin horizontal lines indicate the value of c_{13}/c_{33} in ADP and KDP.

$= -0.069$ for KDP); the roots obtained numerically are $v_F = 5425$ m/s and $(h/\lambda)_F = 0.056$.

On the other hand, curves 2 and 3 resemble parabola only in a very close neighborhood of zero. Such an “anomalous” behavior results in three alternatives. The first one arises in the ZnO(f)–KDP structure. From (31) we find that c_{13} and Ξ are of different sign. Hence, according to Sec. III, the surface wave should not exist (in Fig. 3 curve 2 first goes up, while $c_{13}=c'_{12}$ is negative for KDP). But in fact the wave exists, $(h/\lambda)_F = 0.079$.

The second case is illustrated by curve 3 [ZnO(f)–ADP]. In this structure two surface waves appear for the same direction of propagation ($c_{13}/c_{33}=0.058$ for ADP). In the h/λ range between the two surface waves there exists the low-attenuated leaky wave with $\delta < 0.007$ dB/ λ . Analytical consideration predicts only the solution with smaller h/λ . The estimation gives $v_F \approx 5900$ m/s and $(h/\lambda)_F \approx 0.0097$. The “true” values are $v_F = 5886$ m/s and $(h/\lambda)_F = 0.013$.

At last, it may be that the surface wave does not exist, while the analytical calculation leads to a directly opposed conclusion, i.e., from (30) we obtain that c_{13} and Ξ have the like sign. This occurs in ZnO(b)–ADP. The curve c_{13}/c_{33} against h/λ is similar to curve 3. But $(c_{13}/c_{33})_{\max} < 0.058$ and, hence, Eq. (7) has no solution at real velocities in this composite.

As a result, one sees that the analytical estimations on the basis of the expressions derived in Sec. III are not always reliable and, hence, numerical computations are needed. It is not, however, improbable that the enormous effect of higher-order terms in ADP (KDP) composites is due to the large value of the ratio c_{11}/c_{55} (≈ 11 in ADP and KDP). This ratio appears in numerical factors before H^2 via the contractions involving the vector $\xi_{d2}^{(0)}$, because upon normalization according to (21) $A_{d2x}^{(0)} = \sqrt{(c_{11}-c_{55})/c_{11}c_{55}}$. An extra increase of the numerical factor can be provided by the quantity $c_{11}^{(f)}/c_{11} \approx 2.5-3$. Supposing that the factor before H^{2n} is the large quantity $(c_{11}/c_{55})^n (c_{11}^{(f)}/c_{11})^{2n}$, see (26) and (29), multiplied by a combination of the material constants, which is

small for ZnO(b)–KDP but becomes of order of unity for three other composites, we have, thus, a plausible explanation of the behavior of the curves in Fig. 3.

The piezoelectric coupling in ZnO–KDP (ADP) structures being considered destroys the surface wave, since the piezoactive modes in KDP and ADP, oriented as has been pointed out, are SH-polarized. In the presence of the ZnO film SH-bulk waves become involved in the solution and therefore, instead of the surface wave, one obtains a nonsagittally polarized leaky wave.

On the other hand, if all the piezoactive modes are sagittally polarized both in the substrate and layer, then the “fast” sagittally polarized surface wave can be constructed. This problem is outlined below.

IV. “FAST” SURFACE WAVES IN PIEZOELECTRIC COMPOSITES

Let us discuss briefly the existence of surface QL waves in piezoelectric structures. In a piezoelectric medium, like in a nonpiezoelectric one, the parameters of plane modes can be found from an eigenvalue problem $\hat{\mathbf{N}}\xi_\alpha = p_\alpha \xi_\alpha$ for a real matrix $\hat{\mathbf{N}}$.^{19,20} Apart from \mathbf{A}_α and \mathbf{L}_α , the eigenvector ξ_α involves the amplitude of electrical potential and the normal projection of the electrical induction divided by $-ik$, Φ_α and D_α , respectively. For a medium of arbitrary symmetry, $\hat{\mathbf{N}}$ is an 8×8 matrix.^{19,20} When the sagittal plane coincides with a plane of symmetry, the sagittally polarized modes are described by the “reduced” 6×6 matrix and $\xi_\alpha = (A_{\alpha x}, A_{\alpha z}, \Phi_\alpha, L_{\alpha x}, L_{\alpha z}, D_\alpha)^t$. The matrix $\hat{\mathbf{T}}\hat{\mathbf{N}}$ is again symmetric; here $\hat{\mathbf{T}}$ is of the same block structure as $\hat{\mathbf{T}}$ in (4) but now $\hat{\mathbf{O}}$ and $\hat{\mathbf{I}}$ are the 3×3 zero and unit matrices, respectively. Accordingly, $\xi_\alpha \cdot \hat{\mathbf{T}}\xi_\beta = 0$ if $p_\alpha \neq p_\beta$.

In order to represent the boundary conditions in a convenient compact form, one can also introduce the propagator

$$\hat{\mathbf{P}} = e^{-iH\hat{\mathbf{N}}_f} = \sum_{\alpha=1}^6 e^{-ip_{f\alpha}H} \xi_{f\alpha} \otimes \hat{\mathbf{T}}\xi_{f\alpha}, \quad (33)$$

and the vectors $\xi_{h\alpha} = e^{-iH\hat{\mathbf{N}}_s} \xi_\alpha$. Like in the previous section, we ascribe the indices $\alpha=1,3$ and $\alpha=2,4$ to the QT and QL modes, respectively. The indices $\alpha=5,6$ are assigned to the modes of electrical potential in the substrate such that $\alpha=5$ will be “physical” nonuniform mode, i.e., $\text{Im}(p_5) > 0$.

As distinct from nonpiezoelectric composites, in piezoelectric layered structures there are two nonuniform modes available for surface wave construction in the interval $v_{QT} < v < v_{QL}$. These are the “acoustic” QL mode $\alpha=2$ and the mode of electrical potential $\alpha=5$. However, now the wave field must satisfy both the mechanical and electrical boundary conditions. We assume that the external face of the film is mechanically free and consider the following three types of the electrical boundary conditions.

- (i) A dielectric coating with the metallized external surface. The boundary conditions for the surface wave can be written as $\mathbf{V}_{h2} + b_5 \mathbf{V}_{h5} = 0$, where $\mathbf{V}_{h\alpha} = (\mathbf{L}_{h\alpha}, \Phi_{h\alpha})^t$. The surface wave velocity obeys the equation

$$F(v, H) \equiv \|\mathbf{V}_{h1} \mathbf{V}_{h2} \mathbf{V}_{h5}\| = 0. \quad (34)$$

The symbol $\|\dots\|$ denotes the scalar mixed product of the enclosed vector.

- (ii) The substrate is coated with a metallic film. One can introduce the vectors $\mathbf{W}_{h\alpha} = (\mathbf{L}_{h\alpha}, \Phi_\alpha)^t$, where Φ_α is the potential involved in the vector ξ_α , to represent the boundary conditions in the form $\mathbf{W}_{h2} + b_5 \mathbf{W}_{h5} = 0$. The equation for the surface wave velocity reads similarly to (34).
- (iii) The substrate is overlaid by a dielectric film with the nonmetallized external face. In this instance, the acoustic waves traveling in the composite produce an electrical field outside the structure, i.e., in vacuum. Let us ascribe the index “ e ” to the decaying mode of electrical potential in vacuum ($p_e = -i$) and define $\Phi_e = e^{i\pi/4}/\sqrt{2\varepsilon_0}$ and $D_e = -e^{-i\pi/4}/\sqrt{2\varepsilon_0}$ so that $2\Phi_e D_e = -1$. In terms of the vectors $\mathbf{S}_{h\alpha} = (\mathbf{L}_{h\alpha}, \Phi_{h\alpha}, D_{h\alpha})^t$ and $\mathbf{S}_{he} = (\mathbf{0}, \Phi_e, D_e)^t$ (Ref. 21) the boundary conditions read $\mathbf{S}_{h2} + b_5 \mathbf{S}_{h5} = b_e \mathbf{S}_e$ and the equation for the surface wave velocity becomes

$$F(v, H) \equiv \|\mathbf{S}_{h1} \mathbf{S}_{h2} \mathbf{S}_{h5} \mathbf{S}_e\| = 0. \quad (35)$$

The velocity of the pure surface wave is a real root of the equation $F(v, H) = 0$. By analogy with Sec. II, making use of the properties of the vectors $\mathbf{V}_{h\alpha}$, $\mathbf{W}_{h\alpha}$, and $\mathbf{S}_{h\alpha}$, one can show that if $\text{Re}(F(v, H)) = 0$, then $\text{Im}(F(v, H)) = 0$ implies the vanishing of a complex function $J(v, H)$, where

$$J(v, H) \equiv \|\mathbf{V}_{h1} \mathbf{V}_{h3} \mathbf{V}_{h5}\| \text{ or } \|\mathbf{W}_{h1} \mathbf{W}_{h3} \mathbf{W}_{h5}\| \text{ or } \|\mathbf{S}_{h1} \mathbf{S}_{h3} \mathbf{S}_{h5} \mathbf{S}_e\|. \quad (36)$$

In the general case $\text{Re}[J(v, H)]$ and $\text{Im}[J(v, H)]$ should be viewed as independent functions. This means that $F(v, H) = 0$ imposes at least three conditions upon v and H . Therefore, like in nonpiezoelectric structures, it is reasonable to believe that the surface wave in a general manner does not exist in the range $v_{QT} < v < v_{QL}$.

Assume that the interface is also a plane of elastic symmetry for the layer and substrate. On calculating explicitly the matrix $\hat{\mathbf{N}}$, or taking advantage of the Green-Christoffel equation and constitutive connections (see, e.g., Refs. 2 and 22), one can check that the components of ξ_α then obey rules (15)–(17), Φ_α and D_α “changing” similarly to $A_{\alpha z}$ and $L_{\alpha z}$, respectively, or vice versa. Using these relations one can show that in all three cases $|\text{Re}[J(v, H)]| \equiv |\text{Im}[J(v, H)]|$ so that $F(v, H) = 0$ imposes only two conditions upon v and H . Hence, when the interface is a plane of symmetry, the pure surface solution may exist in the range $v_{QT} < v < v_{QL}$ at particular values of v and H (for example, such a wave arises on SiC coated with piezoelectric ZnO film⁸).

The analysis of the solvability of the boundary equations can be performed in the spirit of Sec. III when the substrate is a weak piezoelectric (the layer may be a strong piezoelectric). As a basis of perturbation theory, we choose the eigenvectors of Stroh’s matrix at $c_{13} = e_{ij} = 0$ and $v = v_{QL}$, where

e_{ij} are the piezomoduli of the substrate. Expanding the appropriate functions with respect to c_{13}/c_{33} , H , Δp , and e_{ij} , we arrive at expressions similar to (25) and (29). These expressions will involve terms proportional to even powers of e_{ij} , since neither Δp nor c_{13} can depend on the sign of piezomoduli.

V. CONCLUDING REMARKS

The sagittally polarized surface waves on the plated surface of a solid exist in the velocity range $v_{QT} < v < v_{QL}$ providing the material constants of the substrate and layer satisfy certain conditions. However, these conditions are not stringent and such waves “easily” arise both in nonpiezoelectric and piezoelectric structures.

Unlike the phase velocity of conventional “slow” localized solutions in the range $v < v_{QT}$, the velocity v_F of the “fast” surface wave is not a function of h/λ . Both v_F and h/λ are to be determined from the boundary equations.

We have shown that “fast” surface waves propagate along the interface coinciding, like the sagittal plane, with a plane of elastic symmetry of the substrate and the layer. Such a geometry of propagation is “favorable” for this type of wave, because the unknown v_F and h/λ must then obey only two equations (both in nonpiezoelectric and piezoelectric composites). If the interface is not a plane of symmetry either for the substrate or layer, a third equation appears. We have not found a rigorous proof that the pure localized wave does not exist on the nonsymmetric interface. Nevertheless, in our opinion, the surface wave is very unlikely to exist when the interface is generically oriented.

The analytical consideration of “fast” surface waves developed in this paper is based on the expansion of the boundary equations with respect to the small parameters c_{13}/c_{33} , H , and Δp (and $e_{ij}/\sqrt{c_{ij}\epsilon}$ in piezoelectrics). Allowance has been made for first nonvanishing terms. We have derived approximate expressions for the solutions of these equations and the criterion for the existence of surface waves. It has been found that analytical calculations can agree well with the results of numerical analysis. At the same time it may be that, as applied to some structures, the theory does not work. Strictly speaking it remains unclear why this happens; apparently, the comments given at the end of Sec. III are not satisfactory. Therefore, in searching for “fast” surface waves numerical computations are always necessary. The analytical estimations and predictions can be useful for checking and better understanding numerical results. From analytical study of the problem it also follows that the smaller the ratio c_{13}/c_{33} for the substrate, the more likely the surface wave exists.

Our computations indicate that the velocity and suitable h/λ may be fairly sensitive to changes in the values of the elastic moduli of the substrate and the layer. This fact should be taken into account in preparing experimental samples. Note that the surface wave can be observed not only “directly” but also using the resonant nonspecular reflection of an acoustic bounded beam from a composite immersed into liquid. In the presence of the liquid, there appears the pseudosurface wave with double leakage — into liquid and sub-

strate. Hence, on striking the surface of the structure, the wave incident, e.g., from the liquid, converts into refracted and reflected waves, see Refs. 23 and 24. The abrupt change of the imaginary component of the leaky wave velocity associated with leakage into the substrate (Fig. 2) should result in a strong angular and frequency dependence of the amplitude and shape of the reflected and transmitted acoustic bounded beams.

It should be said that surface waves with phase velocity higher than the velocity of transverse bulk waves can appear even if the sagittal plane is not a plane of symmetry for the structure. The polarization vector of the quasilongitudinal decaying mode in the substrate has then an out-of-plane component. For example, these nonsagittally polarized waves have been found in a ZnO–Al₂O₃ composite.⁹ The question of how the existence of such waves is correlated with the crystallographic symmetry of the structure is considered in Ref. [25].

ACKNOWLEDGMENTS

We are grateful to V. I. Alshits, V. N. Lyubimov, and A. L. Shuvalov for valuable comments. A. N. Darinskii thanks the Russian Foundation for Basic Investigation (Grant No. 98-02-16077) for partial financial support. N. F. Naumenko thanks SAWTEK Inc. for financial support.

¹L. M. Brekhovskikh and O. A. Godin, *Acoustics of Layered Media* (Springer-Verlag, New York, 1992).

²E. Dieulesaint and D. Royer, *Elastic Waves in Solids: Applications to Signal Processing* (Wiley, New York, 1980).

³A. H. Nayfeh, *Wave Propagation in Layered Media* (North-Holland, Amsterdam, 1995).

⁴*Acoustic Surface Waves*, edited by A. Oliner (Springer, Berlin, 1978).

⁵G. G. Kessinikh, V. N. Lyubimov, and V. V. Filippov, “Transverse surface acoustic waves on an isotropic substrate with a piezoelectric overlaid,” *Akust. J.* **31**(4), 492–495 (1985).

⁶G. G. Kessinikh, “Rayleigh surface waves in a soft transverse isotropic layer on a transverse isotropic substrate,” *Akust. J.* **34**(4), 657–662 (1988).

⁷V. I. Alshits, W. Gierulski, V. N. Lyubimov, and A. Radowicz, “Resonance excitation of quasi-Rayleigh waves in plates on soft or hard substrates,” *Crystallogr. Rep.* **42**(1), 20–27 (1997).

⁸I. S. Didenko, F. S. Hickernell, and N. F. Naumenko, “The experimental and theoretical characterization of the SAW propagation properties for zinc oxide on silicon carbide,” *IEEE Trans. Ultrason. Ferroelect., Freq. Contr.* **47**(1), 179–187 (2000).

⁹N. F. Naumenko and I. S. Didenko, “High-velocity surface acoustic waves in diamond and sapphire with zinc film” *Appl. Phys. Lett.* **75**(19), 3029–3031 (1999).

¹⁰V. I. Alshits and J. Lothe, “Comments on the relation between surface wave theory and the theory of reflection,” *Wave Motion* **3**, 297–310 (1981).

¹¹A. N. Stroh, “Steady state problems in anisotropic elasticity,” *J. Math. Phys.* **41**, 77–103 (1962).

¹²P. Chadwick and G. D. Smith, “Foundation of the theory of surface waves in anisotropic elastic media,” *Adv. Appl. Mech.* **17**, edited by C.-S. Yih (Academic, New York, 1977), pp. 303–376.

¹³A. N. Darinskii, “Quasi-bulk Rayleigh waves in semi-infinite media of arbitrary anisotropy,” *Wave Motion* **27**(1), 79–93 (1998).

¹⁴A. N. Darinskii, “Leaky waves and the elastic wave resonance reflection on a crystal-thin solid layer interface. II. Leaky waves given rise to by an exceptional bulk wave,” *J. Acoust. Soc. Am.* **103**, 1845–1854 (1998).

¹⁵D. M. Barnett, P. Chadwick, and J. Lothe, “The behaviour of elastic surface waves polarized in a plane of material symmetry. Addendum to part 1,” *Proc. R. Soc. London, Ser. A* **433**, 699–715 (1991).

¹⁶G. W. Farnell, “Properties of elastic surface waves,” in *Physical Acous-*

- tics, edited by W. P. Mason and R. N. Thurston (Academic, New York, 1970), pp. 109–166.
- ¹⁷J. G. Gualtieri, J. A. Kosinski, and A. Ballato, “Piezoelectric materials for acoustic wave applications,” *IEEE Trans. Ultrason. Ferroelectr. Freq. Control* **41**, 53–58 (1994).
 - ¹⁸A. A. Blistanov *et al.*, *Acoustic Crystals*, edited by M. P. Shaskol’skaya (Nauka, Moscow, 1982) (in Russian).
 - ¹⁹J. Lothe and D. M. Barnett, “Integral formalism for surface waves in piezoelectric crystals. Existence consideration,” *J. Appl. Phys.* **47**, 1799 (1976).
 - ²⁰J. Lothe and D. M. Barnett, “Further development of the theory for surface waves in piezoelectric crystals,” *Phys. Norv.* **8**(4), 239–254 (1976).
 - ²¹V. I. Alshits, A. N. Darinskii, and A. L. Shuvalov, “Theory of reflection of acoustoelectric waves in semi-infinite piezoelectric media,” *Sov. Phys. Crystallogr.* **34**(6), 308–312 (1989).
 - ²²G. A. Maugin, *Continuum Mechanics of Electromagnetic Solids* (North-Holland, Amsterdam, 1988).
 - ²³O. Arkan, E. Telatar, and A. Atalar, “Reflection coefficient null of acoustic waves at a liquid–anisotropic solid interface,” *J. Acoust. Soc. Am.* **85**, 1 (1989).
 - ²⁴V. I. Alshits, A. N. Darinskii, and A. L. Shuvalov, “Resonant reflection and refraction of sound at a liquid–crystal interface,” *Sov. Phys. Solid State* **34**(8), 1337–1346 (1992).
 - ²⁵A. N. Dazinskii, “Symmetry aspects of the existence of high-velocity SAW in layered composites,” *Phys. Lett. A* **266**, 183–186 (2000).

Circular asymmetric Helmholtz resonators

A. Selamet^{a)} and Z. L. Ji

Department of Mechanical Engineering and The Center for Automotive Research,
The Ohio State University, Columbus, Ohio 43210-1107

(Received 29 December 1998; revised 21 July 1999; accepted 10 January 2000)

A three-dimensional (3D) analytical approach is developed to account for the nonplanar wave propagation in the cavity and neck of “piston-driven” circular asymmetric Helmholtz resonators. The present 3D analytical results are compared with (1) the numerical predictions from the boundary element method (BEM) to evaluate the analytical approach; and (2) the one-dimensional (1D) solution to examine the effect of nonplanar waves at area discontinuity between the neck and the cavity. In order to improve the 1D solution, the end correction is also determined by using the 3D analytical approach. The effect of neck offset on the resonance frequency of circular asymmetric Helmholtz resonators is investigated. Predictions of resonance frequency and transmission loss from the present 3D and corrected 1D analytical approaches are, respectively, identical and close to the BEM results, while the corrected 1D approach provides a better accuracy compared to the 1D solutions with Ingard’s correction. Finally, the boundary element method is employed to determine the wave attenuation performance of the “pipe-mounted” Helmholtz resonators to examine the effect of multidimensional waves in the vicinity of the main duct and neck junction. © 2000 Acoustical Society of America. [S0001-4966(00)02604-7]

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

INTRODUCTION

The narrow-band behavior of Helmholtz resonators requires an accurate prediction of the acoustic attenuation and the resonance frequency. The one-dimensional theory^{1,2} based on the axial wave propagation in the circular Helmholtz resonators yields the relationship $\tan kl_c \tan kl_v = A_c/A_v$ for the resonance frequency, with l_c and l_v , A_c and A_v being the lengths and cross-sectional areas of the neck and cavity, respectively. To improve the accuracy of the one-dimensional prediction, the neck length is usually modified by adding a correction factor for each end, thereby replacing l_c by $l'_c = l_c + \delta_p + \delta_v$, δ_v and δ_p being the added lengths corresponding to the cavity volume and main duct interfaces, respectively. For a short neck ($kl_c \ll 1$), the foregoing expression for the resonance frequency may be reduced (considering the end corrections) to $\cot kl_v = k(l_c + \delta_v + \delta_p) \times (A_v/A_c)$, which is the expression used by Chanaud.^{3,4} Several other interesting and useful works on the Helmholtz resonators have been cited, for example, in Refs. 1–3, and will not be discussed here.

The simple one-dimensional method may provide an approximate estimate of the resonance frequency and the wave attenuation behavior. However, the method clearly excludes the effect of higher-order modes, and such neglect results in incorrect prediction of attenuation behavior, particularly the location of resonance frequency. In view of the inherent narrow-band acoustic characteristics of the Helmholtz resonator, such inaccuracy leads to improper designs. In order to examine the effect of nonplanar waves in the cavity and neck generated by the area discontinuity on the acoustic suppression performance, Selamet *et al.*¹ developed a two-dimensional (2D), axisymmetric analytical approach for the

concentric Helmholtz resonators by using a “piston-driven” model. The 2D analytical approach is employed to determine the resonance frequency and acoustic attenuation performance. The analytical results are also compared with those from experiments, BEM, and the one-dimensional methods with and without end correction. The space constraint in practical applications (such as underhood), however, may require the neck to be oriented eccentrically relative to the cavity leading to noncoaxial (asymmetric) Helmholtz resonators. For such configurations, particularly with short cavity and large neck offset, the circumferential higher-order modes are also excited, resulting in a need for three-dimensional analysis for the accurate prediction of resonance frequency and acoustic attenuation.

The objective of the present study is then (1) to develop a *three-dimensional analytical approach* to determine the acoustic suppression performance of circular *asymmetric* Helmholtz resonators, and (2) to obtain an end correction to improve the 1D predictions. The 3D analytical approach is then used to investigate the effect of neck offset from the center of the cavity on the resonance frequency and acoustic attenuation. The resonance frequency and transmission loss predictions from the present 3D analytical approach, the boundary element method, and the 1D analytical methods are compared to discuss the effect of multidimensional waves in the cavity, neck, and main duct.

Following the Introduction, Sec. I describes briefly the pipe-mounted Helmholtz resonator model, Sec. II develops the three-dimensional analytical approach for a piston-driven Helmholtz resonator, and Sec. III discusses the one-dimensional limit and determines the end correction. The results from the three-dimensional analytical approach are compared with the boundary element predictions and the

^{a)}Electronic mail: Selamet.1@osu.edu

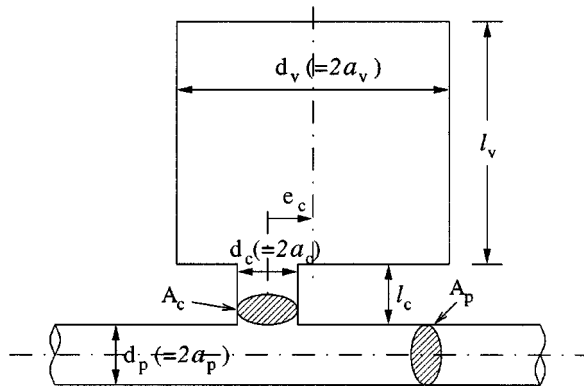


FIG. 1. Pipe-mounted asymmetric Helmholtz resonator.

one-dimensional solutions with and without end correction in Sec. IV. The study is concluded with final remarks in Sec. V.

I. PIPE-MOUNTED HELMHOLTZ RESONATOR

The geometry of a pipe-mounted asymmetric Helmholtz resonator considered in this study is shown in Fig. 1. Assuming one-dimensional wave propagation in the main duct, negligible viscous losses, and constant pressures at the main duct–neck junction and using the conservation of volumetric flow at the same junction, the acoustic theory yields the transmission loss of a Helmholtz resonator as

$$TL = 10 \log_{10} \left| 1 + \frac{A_c}{2A_p} \left(\frac{1}{Z_c} \right) \right|^2, \quad (1)$$

where $Z_c = (P/\rho c U)$ is the acoustic impedance of resonator at the main duct–neck junction. Once Z_c is obtained, the transmission loss and resonance frequency can be determined. To evaluate the acoustic impedance Z_c that accounts for the multidimensional waves generated at the area discontinuity of the neck and cavity, a three-dimensional analytical approach is developed, as described next.

II. THREE-DIMENSIONAL ANALYTICAL APPROACH

Following Ref. 1, the piston-driven Helmholtz resonator model shown in Fig. 2 is used to evaluate the acoustic impedance Z_c . For a circular asymmetric Helmholtz resonator, the sound field in the each cylindrical section may be expressed analytically as described in Appendix A. Equations (A2) and (A8) can be used for waves A and C traveling in the positive x direction, and Eqs. (A3) and (A9) for waves B and D traveling in the negative x direction. The piston with oscillating velocity amplitude U_p requires

$$(U_A + U_B)|_{x=-l_c} = U_p. \quad (2)$$

Substitution of Eqs. (A8) and (A9) into Eq. (2) gives

$$A_{00}e^{jk_{l_c}} - B_{00}e^{-jk_{l_c}} = \rho c U_p, \quad (3)$$

$$A_{mn}e^{-jk_{c,mn}l_c} - B_{mn}e^{jk_{c,mn}l_c} = 0; \quad (4)$$

where $k_{c,mn} = k[1 - (\alpha_{mn}/ka_c)^2]^{1/2}$. At the endplate of the cavity, the rigid wall boundary condition gives

$$(U_C + U_D)|_{x=l_v} = 0. \quad (5)$$

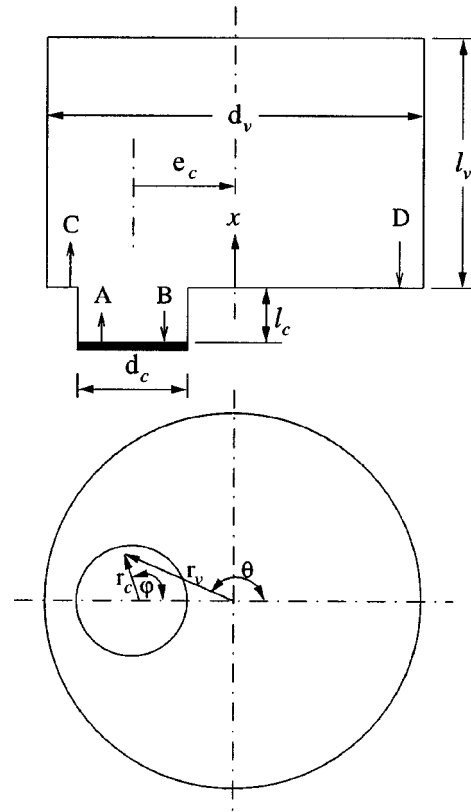


FIG. 2. Piston-driven asymmetric Helmholtz resonator.

Substitution of Eqs. (A8) and (A9) into Eq. (5) yields

$$D_{00} = C_{00}e^{-j2kl_v}, \quad (6)$$

$$D_{mn} = C_{mn}e^{j2k_{v,mn}l_v}, \quad (7)$$

where $k_{v,mn} = k[1 - (\alpha_{mn}/ka_v)^2]^{1/2}$. At the expansion from the neck to the volume, the continuity conditions reveal, for the pressure,

$$(P_A + P_B)|_{x=0} = (P_C + P_D)|_{x=0} \quad (\text{on } A_c), \quad (8)$$

and the continuity and boundary conditions, for the velocity,

$$(U_A + U_B)|_{x=0} = (U_C + U_D)|_{x=0} \quad (\text{on } A_c), \quad (9)$$

$$(U_C + U_D)|_{x=0} = 0 \quad (\text{on } A_v - A_c). \quad (10)$$

To establish a set of independent equations in determining the amplitude coefficients, the foregoing equations are coupled with the orthogonality of Fourier–Bessel functions for the pressure and velocity conditions.

For the pressure boundary condition, multiply both sides of Eq. (8) by $J_t(\alpha_{ts}r_c/a_c)\cos(t\varphi)dA$ and integrate over A_c to get, for $t=0$ and $s=0$,

$$(A_{00} + B_{00}) \frac{a_c^2}{2} = C_{00}(1 + e^{-j2kl_v}) \frac{a_c^2}{2} + \sum_{m=0}^{\infty} \sum_{n=0}^{\infty} C_{mn}(1 + e^{j2k_{v,mn}l_v}) \frac{a_v a_c}{\alpha_{mn}} \times J_m(\alpha_{mn}e_c/a_v) J_1(\alpha_{mn}a_c/a_v); \quad (11)$$

for $t=0$ and $s=1, 2, \dots, \infty$,

$$(A_{0s} + B_{0s}) \frac{a_c^2}{2} J_0(\alpha_{0s}) = \sum_{m=0}^{\infty} \sum_{n=0}^{\infty} ' C_{mn} (1 + e^{j2k_v, mn l_v}) \times J_m(\alpha_{mn} e_c / a_v) \times \frac{\alpha_{mn} a_c / a_v J_0'(\alpha_{mn} a_c / a_v)}{(\alpha_{0s} / a_c)^2 - (\alpha_{mn} / a_v)^2}; \quad (12)$$

and for $t = 1, 2, \dots, \infty$ and $s = 0, 1, \dots, \infty$,

$$(A_{ts} + B_{ts}) \frac{a_c^2}{2} \left(1 - \frac{t^2}{\alpha_{ts}^2}\right) J_t(\alpha_{ts}) = \sum_{m=0}^{\infty} \sum_{n=0}^{\infty} ' C_{mn} (1 + e^{j2k_v, mn l_v}) [J_{m+t}(\alpha_{mn} e_c / a_v) + (-1)^t J_{m-t}(\alpha_{mn} e_c / a_v)] \frac{\alpha_{mn} a_c / a_v J_t'(\alpha_{mn} a_c / a_v)}{(\alpha_{ts} / a_c)^2 - (\alpha_{mn} / a_v)^2}. \quad (13)$$

For the two velocity conditions, multiply both Eqs. (9) and (10) by $J_t(\alpha_{ts} r_v / a_v) \cos(t\theta) dA$ and integrate Eq. (9) over A_c and Eq. (10) over $A_v - A_c$, and then add these two integral equations to yield, for $t=0$ and $s=0$,

$$(A_{00} - B_{00}) a_c^2 = C_{00} (1 - e^{-j2kl_v}) a_v^2, \quad (14)$$

for $t=0$ and $s = 1, 2, \dots, \infty$,

$$k(A_{00} - B_{00}) \frac{a_v a_c}{\alpha_{0s}} J_0(\alpha_{0s} e_c / a_v) J_1(\alpha_{0s} a_c / a_v) - \sum_{m=0}^{\infty} \sum_{n=0}^{\infty} ' k_{c, mn} (A_{mn} - B_{mn}) J_m(\alpha_{0s} e_c / a_v) \times \frac{\alpha_{0s} a_c / a_v J_m(\alpha_{mn}) J_m'(\alpha_{0s} a_c / a_v)}{(\alpha_{mn} / a_c)^2 - (\alpha_{0s} / a_v)^2} = -k_{v, 0s} C_{0s} (1 - e^{j2k_v, 0s l_v}) \frac{a_v^2}{2} J_0^2(\alpha_{0s}), \quad (15)$$

and for $t = 1, 2, \dots, \infty$ and $s = 0, 1, \dots, \infty$,

$$2k(A_{00} - B_{00}) \frac{a_v a_c}{a_{ts}} J_t(a_{ts} e_c / a_v) J_1(\alpha_{ts} a_c / a_v) - \sum_{m=0}^{\infty} \sum_{n=0}^{\infty} ' k_{c, mn} (A_{mn} - B_{mn}) [J_{m+t}(\alpha_{ts} e_c / a_v) + (-1)^t J_{m-t}(\alpha_{ts} e_c / a_v)] \times \frac{\alpha_{ts} a_c / a_v J_m(\alpha_{mn}) J_m'(\alpha_{ts} a_c / a_v)}{(\alpha_{mn} / a_c)^2 - (\alpha_{ts} / a_v)^2} = -k_{v, ts} C_{ts} (1 - e^{j2k_v, ts l_v}) \frac{a_v^2}{2} \left(1 - \frac{t^2}{\alpha_{ts}^2}\right) J_t^2(\alpha_{ts}). \quad (16)$$

The detailed derivation of Eqs. (13) and (16) is deferred to Appendix B.

Equations (3), (4), and (11)–(16) give a large (theoretically infinite) number of relations $3(t+1)(s+1)$ for a large

number of unknowns $3(m+1)(n+1)$. The unknowns are the pressure magnitudes for waves A_{mn} , B_{mn} , and C_{mn} . Since the higher-order modes have a diminishing effect on the solution, t and m can be truncated to p terms and s and n to q terms resulting in $3(p+1)(q+1)$ equations with $3(p+1)(q+1)$ unknowns. The truncation of terms and the convergence have been discussed in Ref. 5, and will not be repeated here. For the geometries and frequencies investigated here, $p=5$ and $q=5$ were found to be sufficient. To determine the acoustic impedance of the Helmholtz resonator, $\rho c U_p$ is chosen to be unity for convenience. Once Eqs. (3), (4), and (11)–(16) are solved, the acoustic impedance Z_c can be evaluated by

$$Z_c = \bar{P} / (\rho c U_p) = \int_{A_c} (P_A + P_B) dA = (A_{00} e^{jkl_c} + B_{00} e^{-jkl_c}), \quad (17)$$

and then the transmission loss can be determined by using Eq. (1).

III. ONE-DIMENSIONAL LIMIT AND END CORRECTION

Setting $p=0$ and $q=0$ in Eqs. (3), (11), (14), and (17) readily gives the classical transmission loss of a one-dimensional Helmholtz resonator as¹

$$TL = 10 \log_{10} \left[1 + \left(\frac{A_c}{2A_p} \frac{\tan kl_c + (A_v / A_c) \tan kl_v}{1 - (A_v / A_c) \tan kl_c \tan kl_v} \right)^2 \right]. \quad (18)$$

As described in the Introduction for one-dimensional acoustic analysis, l_c in Eq. (18) needs to be modified by an end correction, particularly for short necks, to obtain a reasonable prediction for resonance frequency. One such correction that accounts for the nonplanar wave propagation in the circular cylindrical duct was suggested by Ingard⁶ as

$$\delta_v = \sum_{m=0}^{\infty} \sum_{n=0}^{\infty} ' \frac{4a_v J_1^2(\alpha_{mn} a_c / a_v) J_m^2(\alpha_{mn} e_c / a_v)}{\alpha_{mn}^3 (1 - m^2 / \alpha_{mn}^2) J_{mn}^2(\alpha_{mn})}, \quad (19)$$

who replaced the neck with a piston oscillating into an expanded duct of infinite length (hereafter referred to as “single-infinite-duct” model). In Eq. (19), the prime on the summation sign implies that the term representing the (0, 0) fundamental mode is excluded. Further discussion on this approach is deferred to Appendix C. Karal⁷ investigated the acoustic impedance for discontinuities at pipes of circular cross section by using an approach similar to Ingard’s work. The model, however, excludes the nonplanar waves in the smaller duct (for example, the neck of a Helmholtz resonator). Kergomard and Garcia⁸ examined the difference between the foregoing configuration and a pair of circular ducts connected through a sudden cross-sectional area change (hereafter referred to as “double-infinite-duct” model). They employed the mode-matching technique to determine the Karal correction factor of a coaxial discontinuity. Using an analytical approach and the finite element method, Peat⁹ also studied the effect of evanescent modes for coaxial discontinuities and developed an empirical relationship for the Karal correction factor in terms of a double Chebyshev series.

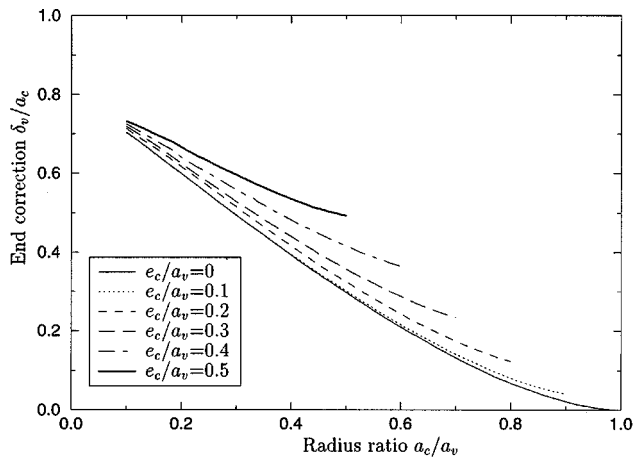


FIG. 3. End correction of infinite duct.

Later, by curve fitting the finite element results, Sahasrabudhe *et al.*¹⁰ provided polynomial expressions for the Karal correction factor for both coaxial and offset configurations. Their studies demonstrated the importance of the higher-order modes in two ducts.

The present study employs the three-dimensional analytical approach to determine the end correction of asymmetric discontinuity and discuss the effect of neck offset on the end correction. The end correction may be expressed as^{8,9}

$$\delta_v = \frac{P_{00}^c - P_{00}^v}{j\rho\omega U_{00}^c} = \frac{(A_{00} + B_{00}) - (C_{00} + D_{00})}{jk(A_{00} - B_{00})}, \quad (20)$$

where the subscript 00 denotes the (0,0) fundamental mode. To determine the end correction, Eqs. (11)–(16) are used to calculate the pressure amplitudes for a specified planar wave A . For low frequencies (for example, $ka_v \ll \alpha_{01}$ for the concentric configuration, or $ka_v \ll \alpha_{10}$ for the offset configuration), $k_{v,mn} \approx j\alpha_{mn}/a_v$ and $k_{c,mn} \approx j\alpha_{mn}/a_c$, resulting in an end correction independent of frequency. The effect of number of terms on end correction needs to be examined, and the study shows that the number of modes needed for convergence increases as radius ratio decreases. For the range of $a_c/a_v \geq 0.1$, using $p=q=20$ may provide a reasonably converged end correction. The difference between predictions with $p=q=20$ and $p=q=40$ is less than 2% for $a_c/a_v \geq 0.1$. Figure 3 depicts the dimensionless end correction (δ_v/a_c) versus radius ratio (a_c/a_v) of infinite duct from the present approach with $p=q=20$ for different relative neck offsets (e_c/a_v). It is interesting to note that for a given radius ratio a_c/a_v , the end correction appears to increase with increasing eccentricity e_c/a_v , rather than varying linearly with e_c/a_v . For the concentric configuration with small radius ratio (a_c/a_v), an approximate expression for the end correction may be suggested as

$$\delta_v = 0.82a_c(1 - 1.33a_c/a_v), \quad (21)$$

which compares well with Eq. (20) for $a_c/a_v < 0.4$, as illustrated in Fig. 4.

The end-correction predictions from Eqs. (19) and (20) are shown in Fig. 5 for the special value $a_c/a_v = 0.15$. As expected, the dotted line for the single-infinite-duct model matches Fig. 6 of Ref. 6. The significant difference in end

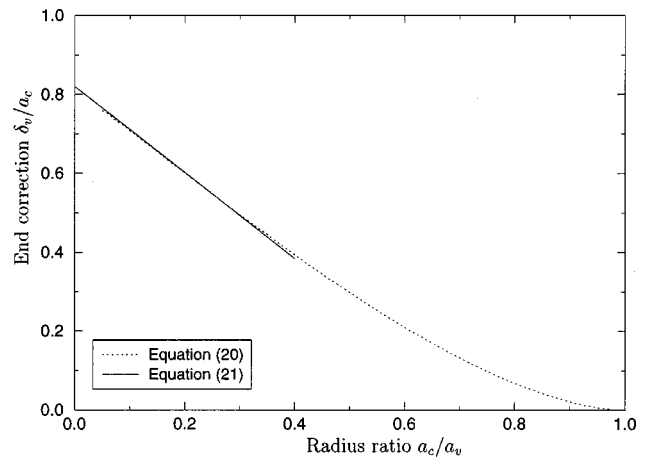


FIG. 4. End correction of concentric infinite duct: dotted line, Eq. (20) with $q=80$; solid line, Eq. (21).

correction between the double- and the single-infinite-duct models, particularly for the asymmetric configurations, demonstrates the importance of the higher-order modes in the region of neck neighboring the volume.

Unlike infinite-duct models, the formulation of the present study incorporates the reflections from the end wall. The end corrections for the *finite* versus *infinite* cylindrical cavities in a double-duct configuration are then depicted in Fig. 6 for the symmetric case, and Fig. 7 for the asymmetric case. The examination of these two figures reveals that the difference in the end correction due to finite versus infinite duct becomes negligible for the length-to-diameter ratio $l_v/d_v > 0.3$ for the concentric configuration and $l_v/d_v > 0.6$ for the asymmetric configuration. Thus, for the finite cavity with large l_v/d_v , the end correction for the infinite duct may be used as a reasonable approximation.

IV. RESULTS AND DISCUSSION

To investigate the effect on resonance frequencies, consider, for example, one of the eight configurations fabricated for the experiments in Ref. 1, $l_v = 24.420$ cm, $d_v = 15.319$ cm, $l_c = 8.50$ cm, $d_c = 4.044$ cm, $d_p = 4.859$ cm;

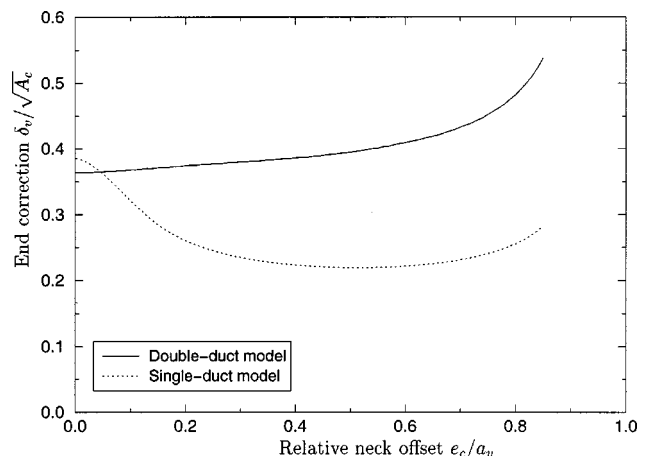


FIG. 5. End correction for the configuration with $a_c/a_v = 0.15$: dotted line, Eq. (19) for the single-infinite-duct model; solid line, Eq. (20) for the double-infinite-duct model.

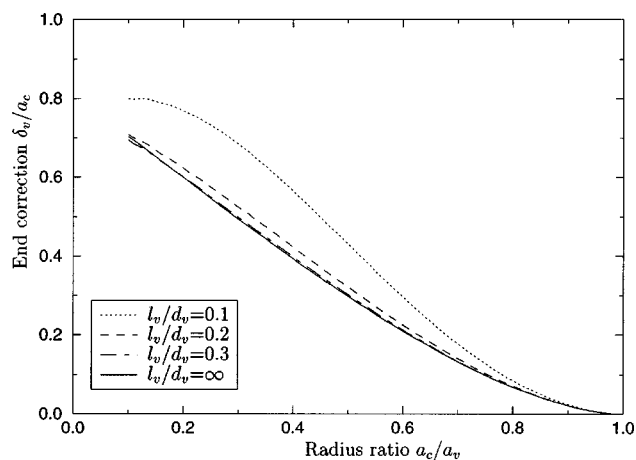


FIG. 6. End correction for finite symmetric cavity in a double-duct configuration.

thus, $l_v/d_v = 1.59$ and $a_c/a_v = 0.264$. The location/eccentricity of the neck is, however, allowed to vary in this work. Five approaches are employed to determine the resonance frequency of double-duct geometry with different neck offsets: (1) the simple 1D result from Eq. (18), (2) the 1D combined with end correction for the junction of neck and cavity, (3) the present three-dimensional analytical approach, (4) the boundary element method for the piston-driven model, and (5) the boundary element method for the pipe-mounted model.

The resonance frequency predictions from the 1D approaches without and with end corrections are depicted in Fig. 8, and compared with those from the 3D approaches. The simple 1D approach without end correction is transparent, as expected from Eq. (18), to neck offsets, and thus gives a single resonance frequency (96.7 Hz), while the 1D method combined with end corrections [l_c is replaced by $l_c + \delta_v$ in Eq. (18)] yields different resonance frequencies for different neck offsets, since the effect of offset is considered in δ_v . The present three-dimensional analytical approach explicitly includes the nonplanar waves both in the neck and the cavity, and thus predicts resonance frequencies varying with the neck offset. The resonance frequency predictions

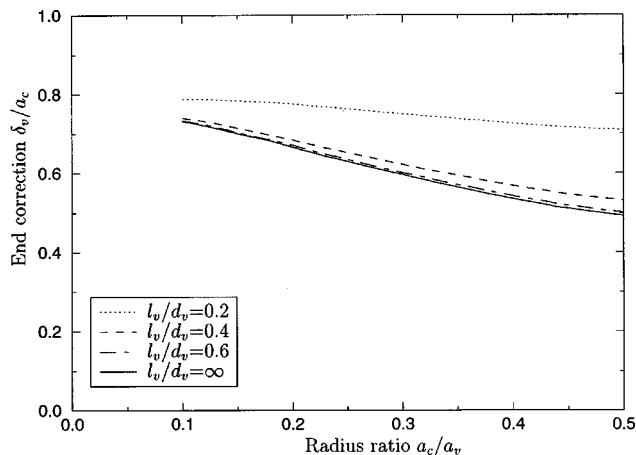


FIG. 7. End correction for finite asymmetric cavity ($e_c/a_v = 0.5$) in a double-duct configuration.

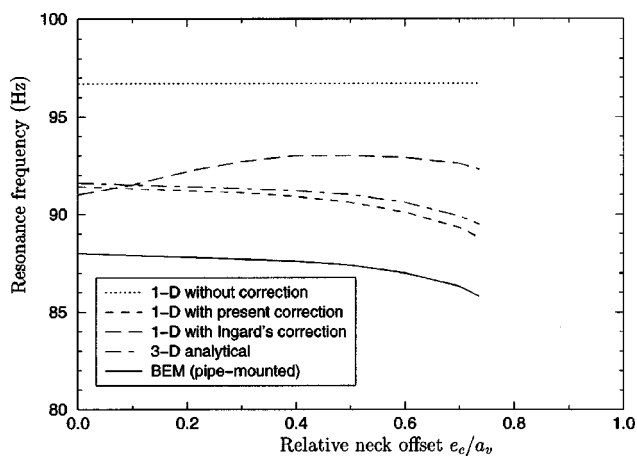


FIG. 8. Comparison for resonance frequency of Helmholtz resonators [$l_v = 24.420$ cm, $d_v = 15.319$ cm ($l_v/d_v = 1.59$), $l_c = 8.50$ cm, $d_c = 4.044$ cm ($a_c/a_v = 0.264$), $d_p = 4.859$ cm, $c_0 = 343.7$ m/s; all analytical results are for piston-driven configuration].

from BEM for the piston-driven model are identical to those from the present 3D analytical approach; thus, a single curve designated as 3D analytical (piston-driven) represents both. The 1D approach with the present end correction and the 3D predictions agree well (within 1% relative error), showing that the resonance frequency is reduced as the neck offset from cavity axis e_c is increased. A similar trend is also observed in measured results of Chanaud⁴ for the circular Helmholtz resonators with a thin orifice. However, the trend with the 1D method with Ingard's end correction does not appear to be accurate. The discrepancy in resonance frequency predictions from the 3D analytical approach and the 1D method with Ingard's end correction is attributed to the fact that the effect of higher-order modes in the region of neck neighboring the volume is neglected in the latter. The boundary element method is also used to calculate the characteristic impedance Z_c of the piston-driven Helmholtz resonator of Fig. 2, and then to determine the transmission loss by substituting Z_c into Eq. (1).

In order to examine the effect of multidimensional wave propagation in the main duct, neck, and cavity, the boundary element method is also employed to determine the transmission loss for the entire duct-neck-cavity configuration (the pipe-mounted model), as well as the resonance frequencies which are depicted in Fig. 8. In the present BEM calculations, the boundary surfaces are discretized into eight-node quadrilateral and six-node triangular elements. The details of the numerical procedure using quadratic isoparametric elements for three-dimensional analysis can be found elsewhere.¹¹ In general, the boundary element method provides a reliable prediction of acoustic attenuation performance due to its ability to incorporate the complicated geometry and nonplanar wave propagation at the duct-neck junction, which was left uncorrected in the analytical approaches.

The transmission loss results from the 1D approaches with (present and Ingard) and without the end corrections, the present 3D analytical approach for the piston-driven model, and the boundary element method for the pipe-mounted model are depicted in Fig. 9 for the concentric con-

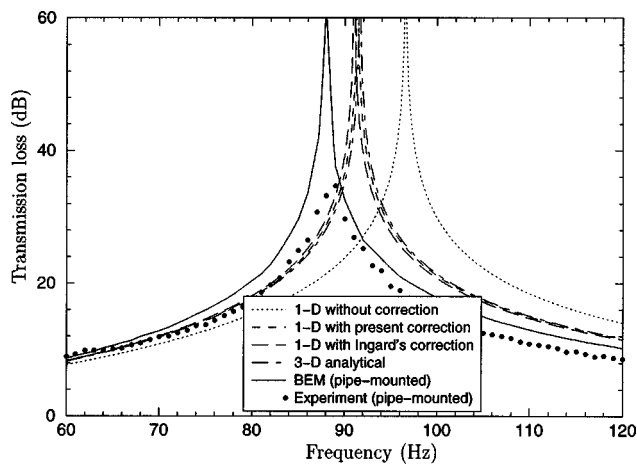


FIG. 9. Transmission loss of concentric Helmholtz resonator (dimensions same as in Fig. 8; all analytical results are for piston-driven configuration).

figuration and Fig. 10 for the asymmetric configuration with the extreme neck offset of $e_c/a_v = 0.736$. For comparison purposes, Fig. 9 also includes the experimental results.¹ The examination of accurate 3D predictions demonstrates the shift in resonance to lower frequencies from piston-driven to pipe-mounted configurations. Thus, to improve the analytical predictions for pipe-mounted resonators, the length correction at the main duct–neck junction needs to be obtained and incorporated into the 3D and corrected 1D analytical approaches. These two figures also illustrate the close representation of the 3D results by 1D approach with the present correction.

V. CONCLUDING REMARKS

This study has shown the effect of geometry and nonplanar waves on the resonance frequency and wave attenuation characteristics of circular *asymmetric* Helmholtz resonators. A three-dimensional analytical approach was developed for a piston-driven finite double-duct configuration to examine the nonplanar wave propagation at the area discontinuity of the neck–cavity interface. The resonance frequency predictions from this approach agree well with the boundary

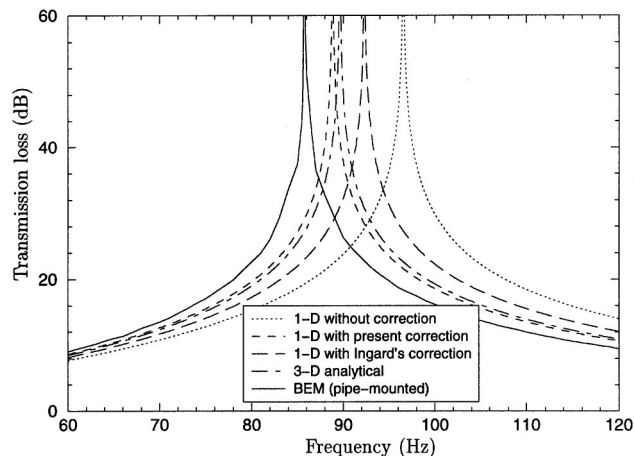


FIG. 10. Transmission loss of asymmetric Helmholtz resonator with $e_c/a_v = 0.736$ (other dimensions same as in Fig. 8; all analytical results are for piston-driven configuration).

element results, as expected. The predictions from the 1D approach with present correction and multidimensional approaches show that the resonance frequency decreases as the neck offset from the center of the cavity is increased, whereas the 1D method with Ingard's correction does not follow the same trend. The difference in resonance frequency predictions between the 3D analytical approach and the 1D method with Ingard's correction is attributed to the fact that the effect of nonplanar waves in the neck near the opening to the volume is excluded in the derivation of end correction for the corrected 1D analysis of Ingard. In order to improve the corrected 1D method, the 3D analytical approach is also employed to calculate the end correction for the double-duct model. The 1D predictions coupled with the present end correction agree well with those from the 3D analytical approach.

The boundary element method was also employed to examine the effect of multidimensional waves in the vicinity of the main duct–neck junction on the acoustic characteristics of the pipe-mounted Helmholtz resonator. The resonance frequency was observed to shift towards a lower value in comparison with the piston-driven model. Such shift demonstrates the need to account for nonplanar behavior at the main duct–neck junction for an accurate prediction of resonance frequency and acoustic attenuation of pipe-mounted resonators.

APPENDIX A: SOUND PROPAGATION IN CIRCULAR DUCT

For the three-dimensional sound propagation in circular duct, the governing equation is the well-known Helmholtz equation¹²

$$\nabla^2 P + k^2 P = 0, \quad (\text{A1})$$

where P is the acoustic pressure, $k = \omega/c$ is the wave number, ω is the angular frequency, and c is the speed of sound. By employing the separation of variables and considering the presence of symmetry, the following general solutions of the sound pressure can be written, for a wave C traveling in the positive x direction:

$$P_C = C_{00}e^{-jkx} + \sum_{m=0}^{\infty} \sum_{n=0}^{\infty} C_{mn}e^{jk_{mn}x} J_m(\alpha_{mn}r/a) \cos(m\theta), \quad (\text{A2})$$

and, for a wave D traveling in the negative x direction,

$$P_D = D_{00}e^{jkx} + \sum_{m=0}^{\infty} \sum_{n=0}^{\infty} D_{mn}e^{-jk_{mn}x} J_m(\alpha_{mn}r/a) \cos(m\theta), \quad (\text{A3})$$

where C_{mn} and D_{mn} are, respectively, the constants corresponding to the waves traveling in the positive and negative x directions, a is the radius of duct, the prime on the summation sign denotes that the planar wave term ($m=0, n=0$) is excluded, J_m is the Bessel function of the first kind of order m ; α_{mn} is the root satisfying the radial boundary condition of

$$J'_m(\alpha_{mn}) = 0, \quad (\text{A4})$$

and m and n denote the asymmetrical and radial mode numbers; and

$$k_{mn} = k[1 - (\alpha_{mn}/ka)^2]^{1/2} \quad (\text{A5})$$

is the axial wave number of the mode (m, n) . Examining Eq. (A5) for any high-order mode (m, n) , k_{mn} will be imaginary when

$$f < \frac{c}{2\pi} \left(\frac{\alpha_{mn}}{a} \right). \quad (\text{A6})$$

The sign difference between the planar and higher-order modes in the exponential terms of Eqs. (A2) and (A3) ensures that for a wave traveling, for example, in the positive direction, the magnitude of all modes will decrease exponentially to zero with increasing distance when Eq. (A6) is satisfied. The axial particle velocities for waves C and D are obtained, in view of the harmonic time dependence $\exp(j\omega t)$, from the momentum equation,

$$j\rho\omega\mathbf{U} = -\nabla p, \quad (\text{A7})$$

as

$$U_C = \frac{1}{\rho\omega} \left\{ kC_{00}e^{-jkx} - \sum_{m=1}^{\infty} \sum_{n=0}^{\infty} k_{mn}C_{mn}e^{jk_{mn}x} \times J_m(\alpha_{mn}r/a)\cos(m\theta) \right\}, \quad (\text{A8})$$

and

$$U_D = -\frac{1}{\rho\omega} \left\{ kD_{00}e^{jkx} - \sum_{m=1}^{\infty} \sum_{n=0}^{\infty} k_{mn}D_{mn}e^{-jk_{mn}x} \times J_m(\alpha_{mn}r/a)\cos(m\theta) \right\}, \quad (\text{A9})$$

where ρ is the density of the medium.

APPENDIX B: DERIVATION OF EQS. (13) AND (16)

For the pressure boundary condition, multiply both sides of Eq. (8) by $J_t(\alpha_{ts}r_c/a_c)\cos(t\varphi)dA$, and integrate over A_c to get

$$\begin{aligned} & \int_0^{2\pi} \int_0^{a_c} \left\{ (A_{00} + B_{00}) + \sum_{m=0}^{\infty} \sum_{n=0}^{\infty} (A_{mn} + B_{mn}) \right. \\ & \quad \times J_m(\alpha_{mn}r_c/a_c)\cos(m\varphi) \left. \right\} J_t(\alpha_{ts}r_c/a_c) \\ & \quad \times \cos(t\varphi)r_c dr_c d\varphi \\ & = \int_0^{2\pi} \int_0^{a_c} \left\{ C_{00}(1 + e^{-j2klv}) \right. \end{aligned}$$

$$\begin{aligned} & + \sum_{m=0}^{\infty} \sum_{n=0}^{\infty} C_{mn}(1 + e^{j2k_v, mn^l v}) J_m(\alpha_{mn}r_v/a_v) \\ & \quad \times \cos(m\theta) \left. \right\} J_t(\alpha_{ts}r_c/a_c)\cos(t\varphi)r_c dr_c d\varphi. \quad (\text{B1}) \end{aligned}$$

Using Graf's addition theorem for Bessel functions¹³

$$J_m(\mu r_v)\cos(m\theta) = \sum_{p=-\infty}^{\infty} J_{m+p}(\mu e_c)J_p(\mu r_c)\cos(p\varphi) \quad (\text{B2})$$

for the right side of Eq. (B1) to transform the coordinates of the cavity to those of the neck, and then integrating over φ from 0 to 2π for both sides yields

$$\begin{aligned} & \int_0^{a_c} \sum_{n=0}^{\infty} (A_{tn} + B_{tn}) J_t(\alpha_{tn}r_c/a_c) J_t(\lambda_{ts}r_c/a_c) r_c dr_c \\ & = \int_0^{a_c} \sum_{m=0}^{\infty} \sum_{n=0}^{\infty} C_{mn}(1 + e^{j2k_v, mn^l v}) \\ & \quad \times [J_{m+t}(\alpha_{mn}e_c/a_v) + (-1)^t J_{m-t}(\alpha_{mn}e_c/a_v)] \\ & \quad \times J_t(\alpha_{mn}r_c/a_v) J_t(\alpha_{ts}r_c/a_c) r_c dr_c. \quad (\text{B3}) \end{aligned}$$

Integrating over r_c from 0 to a_c , and applying Eq. (A4) for the radial boundary condition gives

$$\begin{aligned} & (A_{ts} + B_{ts}) \frac{a_c^2}{2} \left(1 - \frac{t^2}{\alpha_{ts}^2} \right) J_t(\alpha_{ts}) \\ & = \sum_{m=0}^{\infty} \sum_{n=0}^{\infty} C_{mn}(1 + e^{j2k_v, mn^l v}) [J_{m+t}(\alpha_{mn}e_c/a_v) \\ & \quad + (-1)^t J_{m-t}(\alpha_{mn}e_c/a_v)] \frac{\alpha_{mn}a_c/a_v J'_t(\alpha_{mn}a_c/a_v)}{(\alpha_{ts}/\alpha_c)^2 - (\alpha_{mn}/a_v)^2}, \quad (\text{B4}) \end{aligned}$$

which is identical to Eq. (13).

For the velocity boundary conditions, multiply both sides of Eqs. (9) and (10) by $J_t(\alpha_{ts}r_c/a_v)\cos(t\theta)dA$ and integrate Eq. (9) over A_c and Eq. (10) over $A_v - A_c$, and then add these two integral equations to get

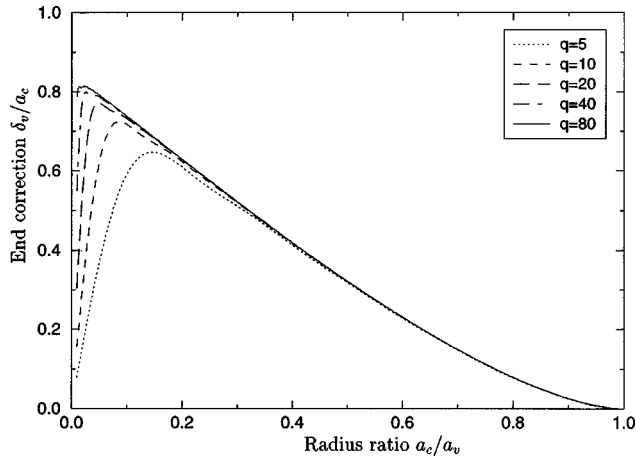


FIG. C1. End correction of concentric configuration.

$$\begin{aligned}
 & \int_0^{2\pi} \int_0^{a_c} \left\{ k(A_{00} - B_{00}) \right. \\
 & \quad \left. - \sum_{m=0}^{\infty} \sum_{n=0}^{\infty} k_{c,mn}(A_{mn} - B_{mn}) J_m(\alpha_{mn} r_c / a_c) \cos(m\varphi) \right\} \\
 & \quad \times J_t(\alpha_{ts} r_v / a_v) \cos(t\theta) r_c dr_c d\varphi \\
 & = \int_0^{2\pi} \int_0^{a_v} \left\{ k C_{00} (1 - e^{-j2kl_v}) - \sum_{m=0}^{\infty} \sum_{n=0}^{\infty} k_{v,mn} C_{mn} \right. \\
 & \quad \left. \times (1 - e^{j2k_{v,mn} l_v}) J_m(\alpha_{mn} r_v / a_v) \cos(m\theta) \right\} \\
 & \quad \times J_t(\alpha_{ts} r_v / a_v) \cos(t\theta) r_v dr_v d\theta. \quad (B5)
 \end{aligned}$$

Using the Graf's addition theorem for Bessel functions (B2) for the left side of Eq. (B5) to transform the coordinates of the cavity to those of the neck, and then integrating over φ for the left side and over θ for the right side, yields

$$\begin{aligned}
 & \int_0^{a_c} \left\{ 2k(A_{00} - B_{00}) J_t(\alpha_{ts} e_c / a_v) J_0(\alpha_{ts} r_c / a_v) \right. \\
 & \quad \left. - \sum_{m=0}^{\infty} \sum_{n=0}^{\infty} k_{c,mn}(A_{mn} - B_{mn}) \right. \\
 & \quad \left. \times [J_{m+t}(\alpha_{ts} e_c / a_v) + (-1)^t J_{m-t}(\alpha_{ts} e_c / a_v)] \right. \\
 & \quad \left. \times J_m(\alpha_{mn} r_c / a_c) J_m(\alpha_{ts} r_c / a_v) \right\} r_c dr_c \\
 & = - \int_0^{a_v} \sum_{n=0}^{\infty} k_{v,tn} C_{tn} (1 - e^{j2k_{v,tn} l_v}) \\
 & \quad \times J_t(\alpha_{tn} r_v / a_v) J_t(\alpha_{ts} r_v / a_v) r_v dr_v. \quad (B6)
 \end{aligned}$$

Integrating over r_c and r_v for the left and right sides of Eq. (B6) yields

$$\begin{aligned}
 & 2k(A_{00} - B_{00}) \frac{a_v a_c}{\alpha_{ts}} J_t(\alpha_{ts} e_c / a_v) J_1(\alpha_{ts} a_c / a_v) \\
 & - \sum_{m=0}^{\infty} \sum_{n=0}^{\infty} k_{c,mn}(A_{mn} - B_{mn}) \\
 & \quad \times [J_{m+t}(\alpha_{ts} e_c / a_v) + (-1)^t J_{m-t}(\alpha_{ts} e_c / a_v)] \\
 & \quad \times \frac{\alpha_{ts} a_c / a_v J_m(\alpha_{mn}) J'_m(\alpha_{ts} a_c / a_v)}{(\alpha_{mn} / a_c)^2 - (\alpha_{ts} / a_v)^2} \\
 & = -k_{v,ts} C_{ts} (1 - e^{j2k_{v,ts} l_v}) \frac{a_v^2}{2} \left(1 - \frac{t^2}{\alpha_{ts}^2} \right) J_t^2(\alpha_{ts}), \quad (B7)
 \end{aligned}$$

which is identical to Eq. (16).

APPENDIX C: INGARD'S END CORRECTION

Equation (19) for the end correction includes infinite modes. Since the higher modes have a diminishing effect on the solution, m and n can be truncated to p and q terms. The number of terms, p and q , needed for a converged solution depends on the magnitude of the area transition. Figure C1 shows the effect of number of terms in Eq. (19) on end correction as a function of radius ratio a_c/a_v . As a_c/a_v decreases, the number of modes needed for convergence increases, which suggests that care should be exercised in choosing the number of terms for adequate convergence particularly at low a_c/a_v . In the limit of $a_c/a_v = 0$, the number of modes approaches infinity. For the symmetric configuration with small values of a_c/a_v , Kergomard and Garcia⁹ simplified Eq. (19) to yield

$$\delta_v = \sqrt{\pi A_c} \left(\frac{8}{3\pi^2} - 0.3523 \frac{a_c}{a_v} \right),$$

or, by rearranging

$$\frac{\delta_v}{a_c} = \frac{8}{3\pi} \left(1 - 1.30 \frac{a_c}{a_v} \right) = 0.85 \left(1 - 1.30 \frac{a_c}{a_v} \right), \quad (C1)$$

which may be compared to the Ingard's approximation⁶ of

$$\frac{\delta_v}{a_c} = 0.85 \left(1 - 1.25 \frac{a_c}{a_v} \right) \quad (C2)$$

for $a_c/a_v < 0.4$. Clearly, both expressions approach $\delta_v/a_c = 8/3\pi$, Rayleigh's upper limit,¹⁴ as $a_c/a_v \rightarrow 0$. The difference in shapes is reflected in Fig. C2 which compares Eqs. (19) with $q = 80$, (C1), and (C2). For small radius ratios (particularly below 0.3), Eq. (C1) provides some improvement over Eq. (C2). Figure C1 illustrates that using $q = 40$ can provide a reasonably converged end correction for the range of $a_c/a_v > 0.05$. For an asymmetric configuration with $e_c/a_v = 0.5$, the results of end-correction calculations are depicted in Fig. C3 for different mode numbers, which illustrates that using $p = 40$ and $q = 40$ can provide a converged value of end correction for the range of $a_c/a_v > 0.05$.

Ingard¹⁵ also examined the effect of *finite* length on the inertial mass of a concentric aperture on the circular cylin-

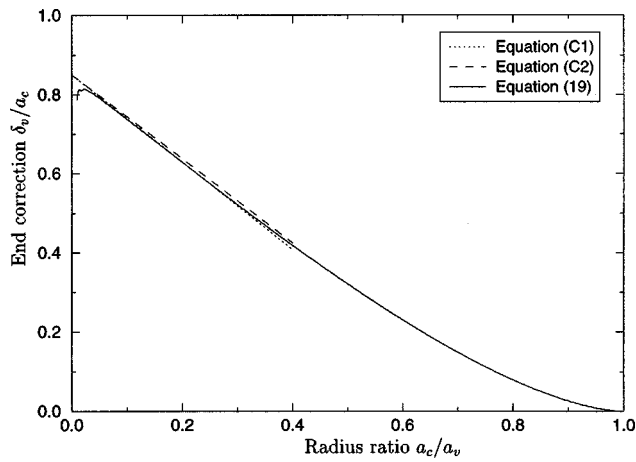


FIG. C2. End correction of concentric configuration: dotted line, Eq. (C1); dashed line, Eq. (C2); solid line, Eq. (19) with $q=80$.

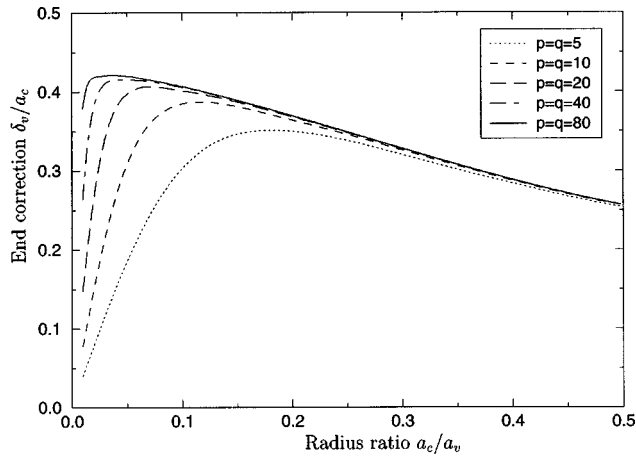


FIG. C3. End correction of offset configuration.

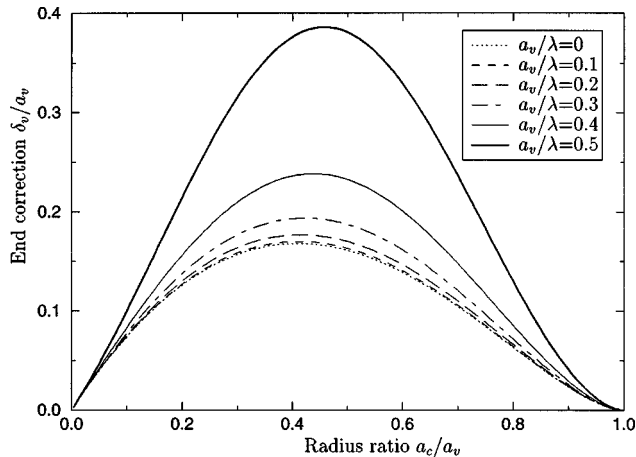


FIG. C4. End correction (inertial mass) of plane piston in a finite circular duct ($l_v/\lambda=0.2$).

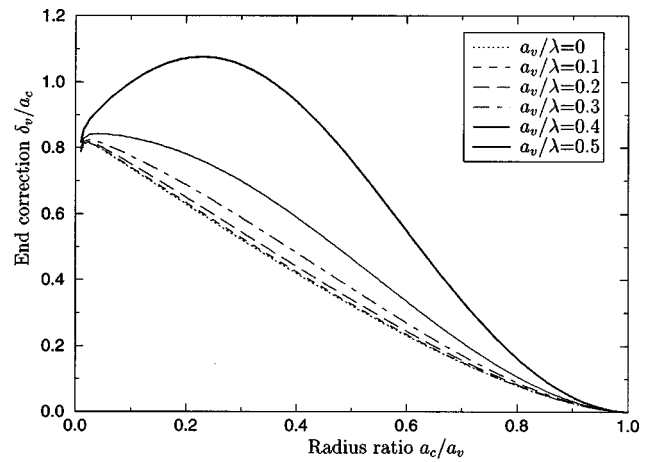


FIG. C5. End correction (inertial mass) of plane piston in a finite circular duct ($l_v/\lambda=0.2$).

drical cavity. The dimensionless inertial mass ($M/\rho A_c a_v$) is equivalent to the dimensionless end correction (δ_v/a_v), which may be expressed as

$$\frac{M}{\rho A_c a_v} \equiv \frac{\delta_v}{a_v} = \sum_{\text{atten. waves}} \frac{4J_1^2(\alpha_{0n} a_c/a_v)}{\alpha_{0n}^2 J_0^2(\alpha_{0n})} \times \frac{\coth(k_{v,0n} l_v)}{[\alpha_{0n}^2 - (2\pi a_v/\lambda)^2]^{1/2}}. \quad (\text{C3})$$

For illustration, the dimensionless end correction is calculated from Eq. (C3) for $l_v/\lambda=0.2$ and depicted in Fig. C4, which corresponds to Fig. 20 in Ref. 15. Note the difference between the results for the case of $a_v/\lambda=0.5$. The vertical coordinate of δ_v/a_v of Fig. C4 has been transformed to δ_v/a_c in Fig. C5.

¹A. Selamet, P. M. Radavich, N. S. Dickey, and J. M. Novak, "Circular concentric Helmholtz resonators," *J. Acoust. Soc. Am.* **101**, 41–51 (1997).

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

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

⁴R. C. Chanaud, "Effects of geometry on the resonance frequency of Helmholtz resonators, Part II," *J. Sound Vib.* **204**, 829–834 (1997).

⁵A. Selamet and Z. L. Ji, "Acoustic attenuation performance of circular flow-reversing chambers," *J. Acoust. Soc. Am.* **104**, 2867–2877 (1998).

⁶U. Ingard, "On the theory and design of acoustic resonators," *J. Acoust. Soc. Am.* **25**, 1037–1061 (1953).

⁷F. C. Karal, "The analogous acoustical impedance for discontinuities of circular cross section," *J. Acoust. Soc. Am.* **25**, 327–334 (1953).

⁸J. Kergomard and A. Garcia, "Simple discontinuities in acoustic waveguides at low frequencies: Critical analysis and formulae," *J. Sound Vib.* **114**, 465–479 (1987).

⁹K. S. Peat, "The acoustical impedance at discontinuities of ducts in the presence of a mean flow," *J. Sound Vib.* **127**, 123–132 (1988).

- ¹⁰A. D. Sahasrabudhe, M. L. Munjal, and S. Anantha Ramu, "Analysis of inertance due to the higher order mode effects in a sudden area discontinuity," *J. Sound Vib.* **185**, 515–529 (1995).
- ¹¹Z. L. Ji, Q. Ma, and Z. H. Zhang, "Application of the boundary element method to predicting acoustic performance of expansion chamber mufflers with mean flow," *J. Sound Vib.* **173**, 57–71 (1994).
- ¹²M. L. Munjal, *Acoustics of Ducts and Mufflers* (Wiley, New York, 1987).
- ¹³C. J. Tranter, *Bessel Functions with Some Physical Applications* (Hart, New York, 1969).
- ¹⁴J. W. S. Rayleigh, *The Theory of Sound* (Dover, New York, 1945), Vol. II.
- ¹⁵U. Ingard, "On the radiation of sound into a circular tube, with an application to resonators," *J. Acoust. Soc. Am.* **20**, 665–682 (1948).

Transfer and Green functions based on modal analysis for Lamb waves generation

Ismael Núñez

Laboratorio de Acústica Ultrasonora, Instituto de Física, Facultad de Ciencias, Montevideo, Uruguay

Ros K. Ing

Laboratoire Ondes de Acoustique, Université Paris 7, ESPCI Paris, France

Carlos Negreira

Laboratorio de Acústica Ultrasonora, Instituto de Física, Facultad de Ciencias, Montevideo, Uruguay

Mathias Fink

Laboratoire Ondes et Acoustique, Université Paris 7, ESPCI Paris, France

(Received 28 July 1999; accepted for publication 28 January 2000)

This work presents an easy way to deduce the tensorial transfer and Green functions for Lamb waves generated in isotropic elastic plates. These functions could be applied to obtain the response of each propagating mode in the ensemble of excited modes arising from any sort of pulsed excitation (wedge transducers, lasers, etc.). The transfer function is based on modal analysis development. Not only is it easy to manipulate but also allows the avoidance of laborious calculations for each kind of Lamb waves source. Theoretical predictions are compared with those of Viktorov [I. A. Viktorov, *Rayleigh and Lamb Waves* (Plenum, New York, 1967)] and with experimental measurements of Lamb waves generated by the wedge-transducer method. © 2000 Acoustical Society of America. [S0001-4966(00)00705-0]

PACS numbers: 43.20.Mv [ANN]

INTRODUCTION

Lamb waves are dispersive waves that propagate along thin plates. They are especially useful for detecting surface and subsurface flaws.¹⁻³

Only a discrete set of wave vectors k_n can propagate for a given product *frequency of excitation times plate thickness*.⁴ Each of these wave vectors defines a normal mode of propagation. The response to an arbitrary excitation can be developed as a sum of normal mode waves, with suitable amplitudes. In the area of nondestructive testing, it is important to control the amplitudes of the excited Lamb modes for a given source of excitation.

Several works have been published relating different kinds of excitations to the corresponding Lamb modes observed in the plate. Viktorov⁴ used the potential field method, with appropriate boundary conditions, to solve monochromatic excitation surface displacements produced by a wedge transducer. Ditri and Rose⁵ developed the problem using the normal-mode method proposed by Auld⁶ and Kino.⁷ They took into account diffraction effects and the transient loading induced by the wedge transducer. Jia⁸ used the normal-mode method for liquid wedge transducers, taking into account the attenuation due to the unmatched impedance between liquid and solid interface.

In this work we use the normal-mode method to deduce a tensorial transfer function for the Lamb waves, and its corresponding Green function. The tensorial transfer function determines the impulse response (temporal and spatial) for longitudinal and transverse Lamb waves displacements. Thus, it applies to any kind of generating source. This transfer function depends only on the plate material and its thick-

ness. It is easy to manipulate and avoids laborious calculations for each type of source.

We compare our theoretical results with experimental measurements made using wedge transducers. We obtain the dispersion curves for the Lamb waves, and we show that the amplitude of each mode in the normal-mode development agrees reasonably with theoretical predictions.

I. THEORY

A. Fundamental relationships

Localized perturbations produced on the plane surface of an elastic plate of thickness $2d$, between $x=d$ and $x=-d$, generate Lamb waves modes which propagate in the yz plane (Fig. 1).

If we assume that perturbations are y -axis independent, and we are interested in the waves traveling towards $z>0$, then the displacements field of the n th propagating Lamb mode, with angular frequency ω , is [omitting the $\exp(-i\omega t)$ factor]^{5,6}

$$U_j^{(n)}(x, z; \omega) = \exp(ik_n(\omega)z) W_j^{(n)}(x; \omega) \quad (1)$$

Here, $k_n(\omega)$ is the wave vector of the n th mode at the frequency ω , and $W_j^{(n)}$ is the j component of the displacement amplitude of the n th mode.

Dispersion relations $k_n(\omega)$ are obtained from the solutions of the Rayleigh-Lamb equations,⁴

$$(k_n^2 + s_n^2)^2 \cosh(q_n d) \sinh(s_n d) - 4k_n^2 q_n s_n \sinh(q_n d) \cosh(s_n d) = 0, \quad (2)$$

for symmetrical modes, and

$$(k_n^2 + s_n^2)^2 \cosh(s_n d) \sinh(q_n d) - 4k_n^2 q_n s_n \sinh(s_n d) \cosh(q_n d) = 0, \quad (3)$$

for antisymmetrical modes, with

$$q_n = \sqrt{k_n^2 - (\omega/c_L)^2}, \quad s_n = \sqrt{k_n^2 - (\omega/c_T)^2}. \quad (4)$$

Here, c_L and c_T are, respectively, the velocities of the longitudinal and transversal bulk waves in the material, and $2d$ is the plate thickness.

The dispersion curves in Fig. 2 for the first symmetrical (solid lines) and antisymmetrical (dashed lines) modes can be deduced from the numerical solutions of Eqs. (2) and (3). The calculation is performed for aluminum ($c_L = 6420$ m/s, $c_T = 3040$ m/s). It is plotted in the dimensionless variables kd against $\omega d/c_T$, where k is the wave number of the propagating Lamb mode, d is half the thickness of the plate, ω is the angular frequency, and c_T is the bulk shear wave velocity in aluminum.

The components $W_j^{(n)}$ of the displacement amplitudes of symmetrical and antisymmetrical modes are⁴

$$W_x^{(n)}(x; \omega) = \begin{cases} A q_n \left(\frac{\sinh(q_n x)}{\sinh(q_n d)} - \frac{2k_n^2}{k_n^2 + s_n^2} \frac{\sinh(s_n x)}{\sinh(s_n d)} \right) & (n \text{ symmetric}) \\ B q_n \left(\frac{\cosh(q_n x)}{\cosh(q_n d)} - \frac{2k_n^2}{k_n^2 + s_n^2} \frac{\cosh(s_n x)}{\cosh(s_n d)} \right) & (n \text{ antisymmetric}) \end{cases}, \quad (5)$$

$$W_z^{(n)}(x; \omega) = \begin{cases} -A k_n \left(\frac{\cosh(q_n x)}{\sinh(q_n d)} - \frac{2q_n s_n}{k_n^2 + s_n^2} \frac{\cosh(s_n x)}{\sinh(s_n d)} \right) & (n \text{ symmetric}) \\ -B k_n \left(\frac{\sinh(q_n x)}{\cosh(q_n d)} - \frac{2q_n s_n}{k_n^2 + s_n^2} \frac{\sinh(s_n x)}{\cosh(s_n d)} \right) & (n \text{ antisymmetric}) \end{cases}.$$

Here, A and B are arbitrary constants.

B. Transfer and Green functions by normal-modes development

According to the normal-modes theory,^{6,7} when a source of harmonic excitation of frequency ω is applied over a localized zone of the plate surface, the perturbed displacement amplitude field can be developed as a sum of normal modes,

$$U_j(x, z; \omega) = \sum_n A_n(z; \omega) W_j^{(n)}(x; \omega). \quad (6)$$

$A_n(z; \omega)$ denotes the amplitude of the n th mode of frequency ω at the point z . Equation (6) involves the summation over all propagating and nonpropagating modes.

We will consider the waves traveling in the $z > 0$ direction excited by a harmonic stress $T_{xk}(z; \omega) \exp(-i\omega t)$ ($k = x, z$) localized in a given region of the plate surface $x = d$. By using the reciprocity relations and orthogonality

condition of eigenmodes^{6,7} and considering only propagating modes, we obtain the following equation for the amplitudes $A_n(z; \omega)$:⁸

$$\frac{dA_n}{dz} - ik_n A_n = \frac{i\omega W_k^{(n)*}(d; \omega) T_{xk}(z; \omega)}{4P_n(\omega)}, \quad (7)$$

$P_n(\omega)$ being the average flow power (per unit length across the y axis) of the n th mode at frequency ω . That is,⁸

$$P_n(\omega) = \frac{1}{2} \omega^2 \rho V_G^{(n)}(\omega) \int_{-d}^d (|W_x^{(n)}(x; \omega)|^2 + |W_z^{(n)}(x; \omega)|^2) dx, \quad (8)$$

where ρ is the mass density and $V_G^{(n)}(\omega)$ is the group velocity of the n th mode at frequency ω .

Tensorial notation is used in Eq. (7) indicating the summation over the repeated subscript k .

Integration of Eq. (7) gives

$$A_n(z; \omega) = \frac{i\omega}{4P_n(\omega)} \exp[ik_n(\omega)z] W_k^{(n)*}(d; \omega) \times \int_0^z T_{xk}(\xi; \omega) \exp(-ik_n \xi) d\xi, \quad (9)$$

where the initial condition $A_n(0; \omega) = 0$ is chosen, as usual, for waves propagating in the $z > 0$ direction.

We shall assume that the excitation $T_{xk}(z; \omega)$ only exists in the source region. This means that we neglect the radiation towards the surrounding medium when the Lamb wave travels along the plate.

For transducer-wedge excitation, the stress $T_{xk}(z; \omega)$ must take into account not only the exciting incident wave, but also the reflected waves in the wedge material as Jia had done.⁸ In the experimental results, the stress $T_{xk}(z; \omega)$ arising from the exciting waves inside the wedge is directly measured on the wedge surface. Afterwards, this amount is put directly into Eq. (7) for purposes of theoretical comparison. However, as long as the normal stress $T_{xk}(z; \omega)$ on the free surface of the wedge is slightly different from the actual stress at the interface between the wedge and the plate surface (with liquid coupling), our theoretical calculations in the case of the wedge-transducer Lamb waves generation are only approximated.

By assuming that the excitation stress $T_{xk}(z; \omega)$ becomes zero outside a finite area, we can perform integration of Eq. (9) along the whole z axis. The result is

$$A_n(z; \omega) = \frac{i\omega}{4P_n(\omega)} W_k^{(n)*}(d; \omega) \exp[ik_n(\omega)z] \times \int_{-\infty}^{\infty} T_{xk}(\xi; \omega) \exp[-ik_n(\omega)\xi] d\xi. \quad (10)$$

By substituting in (10) the inverse spatial Fourier transform of the stress,

$$\tilde{T}_{xk}(k, \omega) = \frac{1}{2\pi} \int_{-\infty}^{\infty} T_{xk}(\xi; \omega) \exp(-ik\xi) d\xi, \quad (11)$$

the amplitude of the excitation of the n th mode becomes

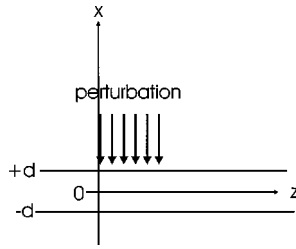


FIG. 1. Plate of thickness $2d$. A surface perturbation around $z=0$ is applied to excite Lamb waves.

$$A_n(z; \omega) = \frac{i\pi\omega}{2P_n(\omega)} W_k^{(n)*}(d; \omega) \tilde{T}_{xk}(k_n, \omega) \exp[ik_n(\omega)z]. \quad (12)$$

By referring to Eq. (6), we obtain the displacement for the n th mode by multiplying (12) and $W_j^{(n)}(x; \omega)$. That is

$$U_j^{(n)}(x; z; \omega) = \frac{i\pi\omega}{2P_n(\omega)} W_j^{(n)}(x; \omega) W_k^{(n)*}(d; \omega) \times \tilde{T}_{xk}(k_n, \omega) \exp[ik_n(\omega)z]. \quad (13)$$

Equation (13) can be written as

$$U_j^{(n)}(x; z; \omega) = \tilde{U}_j^{(n)}(x; k_n, \omega) \exp[ik_n(\omega)z], \quad (14)$$

$\tilde{U}_j^{(n)}$ being the displacement amplitude component of the n th mode according to the j direction. That is, according to Eqs. (13) and (14),

$$\tilde{U}_j^{(n)}(x; k_n, \omega) = \frac{i\pi\omega}{2P_n(\omega)} W_j^{(n)}(x; \omega) W_k^{(n)*}(d; \omega) \tilde{T}_{xk}(k_n, \omega). \quad (15)$$

Equation (15) relates the double Fourier transform $\tilde{T}_{xk}(k, \omega)$ of an arbitrary exciting stress $T_{xk}(z, t)$ on the plate surface to the displacement amplitude of the n th mode. The wave number k follows the dispersion curve $k = k_n(\omega)$. That means that a tensorial transfer function for Lamb waves excitation should now be expressed,

$$H_{jk}^{(n)}(x; \omega) = \frac{i\pi\omega}{2P_n(\omega)} W_j^{(n)}(x; \omega) W_k^{(n)*}(d; \omega). \quad (16)$$

According to (16), (15), and (14), Eq. (6) should now be expressed under the following form:

$$U_j(x; z; \omega) = \sum_n H_{jk}^{(n)}(x; \omega) \tilde{T}_{xk}(k_n, \omega) \exp[ik_n(\omega)z]. \quad (17)$$

The transient behavior can be obtained by using the relation,

$$u_j(x; z, t) = \int_{-\infty}^{+\infty} U_j(x; z; \omega) \exp(-i\omega t) d\omega. \quad (18)$$

By substituting Eq. (17) in (18), it becomes

$$u_j(x; z, t) = \int_{-\infty}^{+\infty} \sum_n H_{jk}^{(n)}(x; \omega) \tilde{T}_{xk}(k_n, \omega) \times \exp[ik_n(\omega)z] \exp(-i\omega t) d\omega. \quad (19)$$

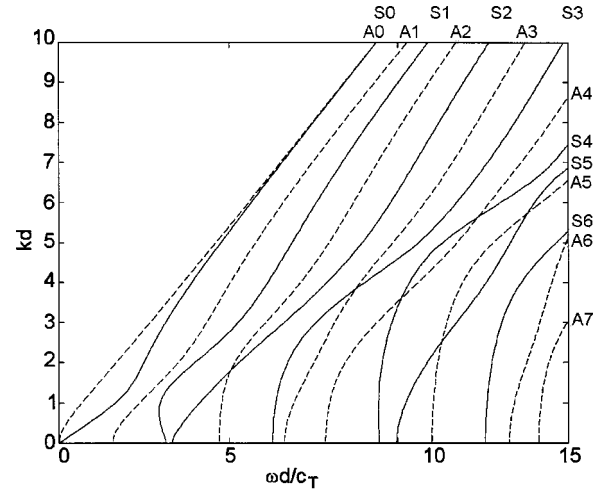


FIG. 2. Dispersion curves for symmetrical (S) and antisymmetrical (A) Lamb modes.

The stress amplitude for a given frequency $T_{jk}(\xi; \omega)$ in Eq. (11) is the inverse Fourier transform of the transient exciting stress $t_{jk}(z, t)$ on the plate surface. This is,

$$T_{xk}(\xi; \omega) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} t_{xk}(\xi, \tau) \exp(i\omega\tau) d\tau. \quad (20)$$

Then, Eq. (11) can be written as

$$\tilde{T}_{xk}(k, \omega) = \left(\frac{1}{2\pi} \right)^2 \int \int t_{xk}(\xi, \tau) \exp(-ik\xi) \exp(i\omega t) d\xi d\tau. \quad (21)$$

By substituting (21) in (19), the transient displacement component according to the j direction at any point x, z in the plate is

$$u_j(x; z, t) = \left(\frac{1}{2\pi} \right)^2 \int \int t_{xk}(\xi, \tau) \int_{-\infty}^{+\infty} \sum_n H_{jk}^{(n)}(x; \omega) \times \exp[ik_n(\omega)(z - \xi)] \exp[-i\omega(t - \tau)] d\omega d\xi d\tau. \quad (22)$$

This relation is a double convolution product in space and time,

$$u_j(x; z, t) = \int \int G_{jk}(x; z - \xi, t - \tau) t_{xk}(\xi, \tau) d\xi d\tau, \quad (23)$$

between the exciting stress t_{xk} and the tensorial Green function

$$G_{jk}(x; z, t) = \left(\frac{1}{2\pi} \right)^2 \int_{-\infty}^{+\infty} \sum_n H_{jk}^{(n)}(x; \omega) \exp[ik_n(\omega)z] \times \exp(-i\omega t) d\omega. \quad (24)$$

Because of Lamb waves detection systems, we are generally interested in the plate surface displacements. That means that in all equations we set $x=d$. However, in this case, we will omit the x variable in the transfer function and will rewrite Eq. (16) as

$$H_{jk}^{(n)}(\omega) = \frac{i\pi\omega}{2P_n(\omega)} W_j^{(n)}(d; \omega) W_k^{(n)*}(d; \omega), \quad (25)$$

and Eq. (15) as

$$\tilde{U}_j^{(n)}(k_n, \omega) = H_{jk}^{(n)}(\omega) \tilde{T}_{xk}(k_n, \omega). \quad (26)$$

Equation (26) can be experimentally evaluated. First, the transfer function (25) for each mode is theoretically evaluated by numerical computation with expressions (5) and (8).

Second, the displacement response according to the j -direction $u_j(z, t)$ can be measured at the plate surface ($x = d$), as a function of position and time.

Following Alleyne and Cawley,⁹ if we perform a numerical double-Fourier transform of the displacement we obtain the two-dimensional function $\tilde{U}_j(k, \omega)$. Since the wave is a superposition of discrete Lamb modes, only the values $\tilde{U}_j(k, \omega)$ such that $k = k_n(\omega)$ do not vanish. That means that the different excited modes are split in the (k, ω) plane, forming the dispersion curves. We can then measure the function $\tilde{U}_j^{(n)}(k_n, \omega)$ for every frequency of each mode through its dispersion curve.

Finally, we can also measure the exciting stress $t_{xk}(z, t)$ according to the k direction on the plate surface. We perform its two-dimensional Fourier transform to obtain the function $\tilde{T}_{xk}(k, \omega)$, and we must evaluate it over the curve $k = k_n(\omega)$ in order to use it in Eq. (26).

The modulus of the tensorial transfer function in Eq. (26) gives the excitation efficiency of the vibrations according to the j direction for the n th Lamb mode as a function of the frequency ω , when a stress in the k direction is applied.

C. Particular case of normal surface displacement detection and normal excitation stress: Comparison with Viktorov results

The most useful detection methods (interferometers or liquid-coupled transducers) can only detect normal displacements. Therefore, we are only interested in the transversal displacements of the Lamb waves on the plate surface (x direction in the axis system of Fig. 1). In this case, we put $j = x$ in Eqs. (25) and (26).

Moreover, excitations of Lamb waves are generally achieved using wedge transducers with a liquid coupling between the wedge and the plate. Then, only normal stresses exist and the important component of the tensorial transfer function in Eq. (25) is the following, with $k = x$,

$$H_{xx}^{(n)}(\omega) = \frac{i\pi\omega}{2P_n(\omega)} |W_x^{(n)}(d; \omega)|^2. \quad (27)$$

By substituting Eq. (27) in (26), we have the Fourier transform of the normal displacement amplitude of the n th Lamb mode, in terms of frequency,

$$\tilde{U}_x^{(n)}(k_n, \omega) = H_{xx}^{(n)}(\omega) \tilde{T}_{xx}(k_n, \omega). \quad (28)$$

By using the Rayleigh–Lamb equations (2), (3), and (4), the Viktorov expressions for the displacements (5), and the flux power expression (8), we have performed the numerical calculation of the transfer function (27) for aluminum ($c_L = 6420$ m/s, $c_T = 3040$ m/s, $\rho = 2700$ kg/m³).

By using the theory of elastic potentials, Viktorov⁴ has performed calculations for Lamb waves amplitudes arising from a normal stress of the form

$$\begin{aligned} T_{xx} &= T_0 \exp[i(k_0 z - \omega t)] \quad \text{if } -b < z < b, \\ T_{xx} &= 0 \quad \text{otherwise.} \end{aligned} \quad (29)$$

Viktorov results for the normal displacements at the plate surface are

$$\begin{aligned} \tilde{U}_x^{(Sn)} &= \frac{iT_0}{\mu} \frac{\sin[(k_0 - k_n)b]}{k_0 - k_n} \frac{q_n(s_n^2 - k_n^2) \sinh(s_n d) \sinh(q_n d)}{\Delta'_S}, \\ \tilde{U}_x^{(An)} &= \frac{iT_0}{\mu} \frac{\sin[(k_0 - k_n)b]}{k_0 - k_n} \\ &\quad \times \frac{q_n(s_n^2 - k_n^2) \cosh(s_n d) \cosh(q_n d)}{\Delta'_A}, \end{aligned} \quad (30)$$

for the symmetrical (Sn) and antisymmetrical (An) modes, respectively. The amounts q_n and s_n have been defined in Eq. (4), where μ is the Lamé constant, and Δ'_S , Δ'_A are the derivatives with respect to k of the Rayleigh–Lamb characteristic determinants,

$$\Delta_S = (k^2 + s^2)^2 \cosh(qd) \sinh(sd) - 4k^2 qs \sinh(qd) \cosh(sd) \quad (31)$$

$$\Delta_A = (k^2 + s^2) \cosh(sd) \sinh(qd) - 4k^2 qs \sinh(sd) \cosh(qd),$$

evaluated in $k = k_n$.

By using excitation (29) in our equation (28), we obtain the relation

$$\tilde{U}_x^{(n)} = \frac{T_0}{\pi} \frac{\sin[(k_0 - k_n)b]}{k_0 - k_n} H_{xx}^{(n)}. \quad (32)$$

Comparison between (30) and (32) shows us that it must be

$$\begin{aligned} H_{xx}^{(n)} &= \frac{i\pi}{\mu} \frac{q_n(s_n^2 - k_n^2) \sinh(s_n d) \sinh(q_n d)}{\Delta'_S} \\ &\quad \text{for symmetrical modes,} \\ H_{xx}^{(n)} &= \frac{i\pi}{\mu} \frac{q_n(s_n^2 - k_n^2) \cosh(s_n d) \cosh(q_n d)}{\Delta'_A} \\ &\quad \text{for antisymmetrical modes.} \end{aligned} \quad (33)$$

We have calculated the first members of (33) with normal-mode development theory—Eq. (27). Independently, we have calculated the right members of (33) emerging from the potential theory. The comparison is plotted in Fig. 3 for symmetrical modes and in Fig. 4 for the antisymmetrical ones. Solid lines represent the transfer function from Eq. (27). The “+” plots represent the potential theory calculations. The agreement is right.

By regarding Figs. 3 and 4, it can be observed that the theory predicts a strong attenuation in the transfer function for certain modes at certain frequencies. This will be shown in experimental results.

II. EXPERIMENTAL RESULTS

Figure 5 shows the experimental setup used for exciting and detecting Lamb waves propagating inside a thin aluminum plate.

The angle θ can be chosen in order to reinforce some Lamb modes. If a plane monochromatic wave is generated

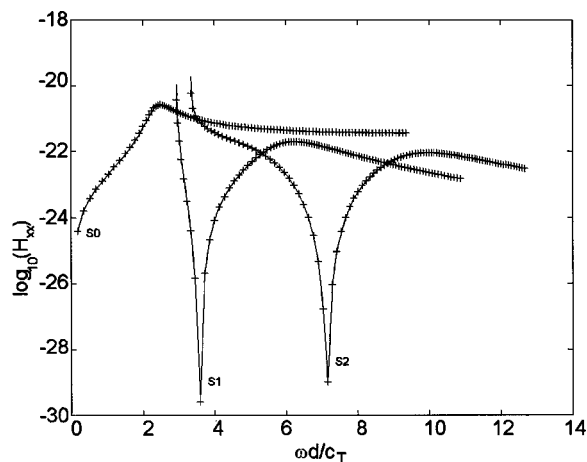


FIG. 3. Transfer function for modes S0, S1, and S2 deduced from normal modes development (solid lines) and Viktorov theory (+ points).

inside the wedge material by a plane transducer of infinite width, the Snell law for refraction states that the phase velocity v_φ of the expected Lamb waves propagating in the z direction satisfies the relation

$$v_\varphi = \frac{v_p}{\sin \theta}, \quad (34)$$

with v_p being the longitudinal velocity in Plexiglas wedge material. Each Lamb mode n has a different phase velocity $v_\varphi^{(n)}$ at a given angular frequency ω . If we choose an incidence angle $\theta^{(n)}$ that satisfies (34), then we can only excite this n th mode.

Actually, plane waves are not achievable because of diffraction effects. Furthermore, we only use pulsed acoustic waves instead of monochromatic ones. We then have finite frequency and wave-number bandwidths. All we can do is choose the angle θ in order to center our excitation in some finite region of Fig. 2.

The pulsed source we used was a Sonic FTS Mark IV ultrasonic tester. The detector was an "NP1000" PVDF needle hydrophone. The hydrophone response at -2 dB goes from 1 to 20 MHz.

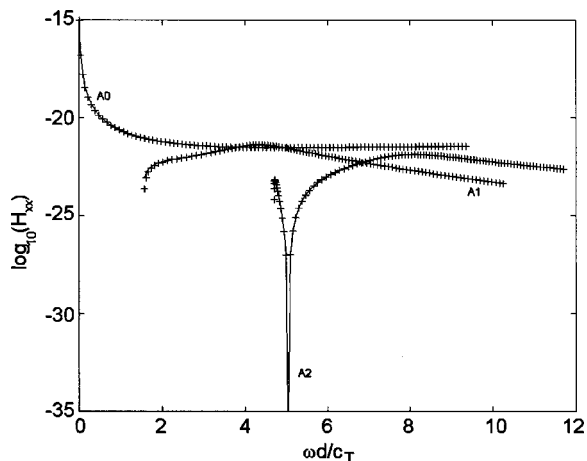


FIG. 4. Transfer function for modes A0, A1, and A2 deduced from normal modes development (solid lines) and Viktorov theory (+ points).

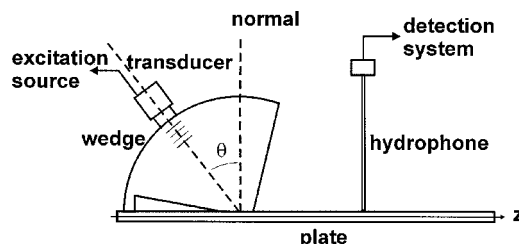


FIG. 5. Experimental setup to generate and detect Lamb waves propagating inside a plate.

The first experiment was performed with an aluminum plate 3.0 mm thick. The transducer was an Aerotech E18975 with a central frequency of 2.25 MHz, a bandwidth frequency of 1.5 MHz, and a diameter of 1.0 in. The incidence angle θ in Fig. 5 was 30° .

The hydrophone picked up a signal of 500 points of 50- μ s time duration. A scan of 1-mm step through 108-mm length in the z direction toward the source was made. As an example, Fig. 6 shows a signal measured in the middle of the path. Notice that the signal generated by the hydrophone is proportional to the pressure. Through the liquid coupling, this pressure corresponds to the normal stress generated by the Lamb waves at the plate surface which is due to the condition of continuity of the normal forces at the solid-liquid interface.

The square modulus of the double-Fourier transform (in space and time) of the signals measured by the hydrophone, corresponds to the normal stress intensity array I of the signals. Figure 7 shows such a resulting array I in a gray-scale representation. These arrays correspond to the experimental Lamb waves dispersion curves.^{9,10}

In order to identify the various Lamb modes excited, Fig. 8 shows the superposition of the experimental data with the theoretical dispersion curves (solid lines) for the symmetrical modes S0, S1, S2 and antisymmetrical modes A1, A2 for a 3-mm aluminum plate.

We observe in Fig. 7 that the S1 Lamb mode almost disappears in the neighborhood of 1.2 MHz. The same phenomenon also occurs for the S2 Lamb mode in the neighborhood of 2.3 MHz. Figure 9 shows the square modulus of the transfer function $|H_{xx}|^2$ for both modes for the 3-mm alumi-

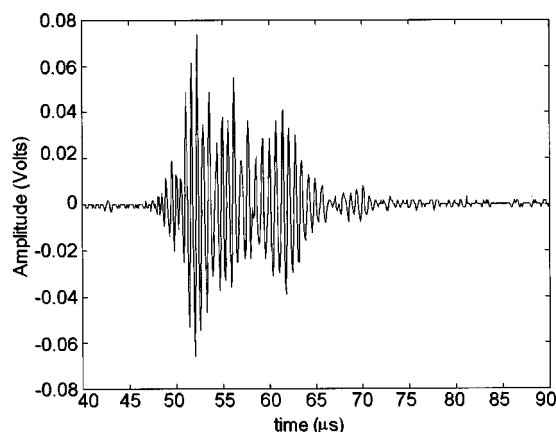


FIG. 6. Example of signal detected over the plate (far from the source).

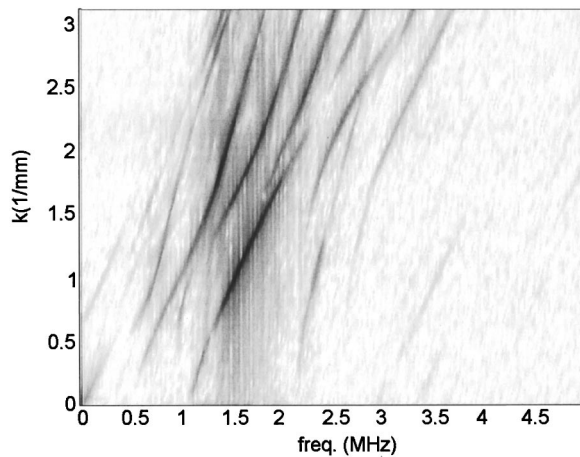


FIG. 7. Experimental dispersion curves of Lamb modes for a 3-mm thickness aluminum plate.

num plate. We observe a strong attenuation at such frequencies.

In order to make a quantitative evaluation of these observations we must involve Eq. (28).

By scanning the signal directly over the free base of the wedge, we can obtain an approximated evaluation of the normal excitation stress $T_{xx}(z, t)$, which is the actual stress that occurs when the wedge is applied onto the aluminum plate.

Thereafter we performed a double-Fourier transform of the signals measured and we obtained the phase diagram (k, ω) of the normal excitation stress. The intensity distribution $|T_{xx}(k, \omega)|^2$ is shown in Fig. 10.

To use Eq. (28) we need to extract the particular values of $|T_{xx}(k, \omega)|^2$ for $k = k_n(\omega)$. This is achieved by reading the intensity values $|T_{xx}(k, \omega)|^2$ along the dispersion curves of modes S1, S2, which are also shown in Fig. 10.

The normal excitation stress intensity measured along the dispersion curves $k = k_n(\omega)$ are shown in Fig. 11(a) and (b) for modes S1 and S2, respectively. Both plots are normalized to the maximum value of the normal excitation stress intensity of mode S1.

In Fig. 11(a) it can be seen that the normal excitation stress intensity of mode S1 at 1.2-MHz and 1.8-MHz fre-

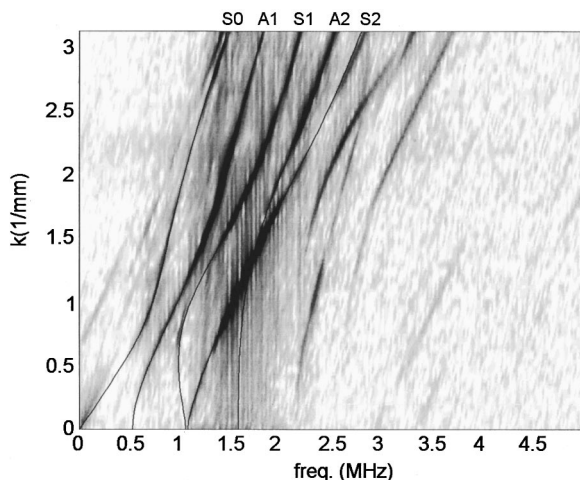


FIG. 8. Identification of experimental Lamb modes shown in Fig. 7.

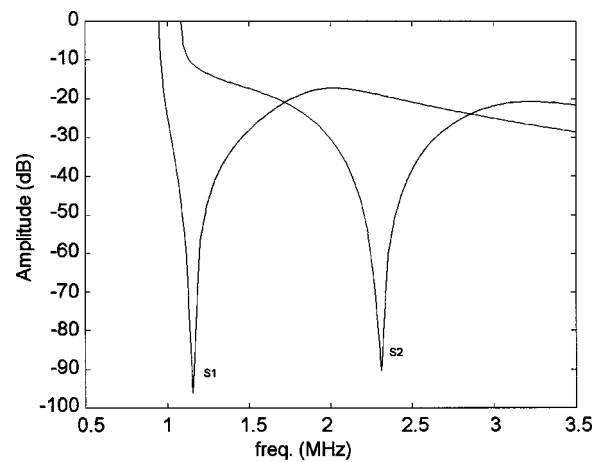


FIG. 9. Transfer function $|H_{xx}|^2$ for Lamb modes S1 and S2 in a 3-mm thickness aluminum plate.

quencies are roughly the same (approximately -7 dB). Nevertheless, Fig. 7 shows that mode S1 is strongly excited at 1.8 MHz while at 1.2 MHz it almost disappears.

The same phenomenon occurs for the Lamb mode S2 [Fig. 11(b)]. At 1.3-, 2.4-, and 2.6-MHz frequencies, the normal excitation stress intensities of mode S2 are roughly the same (approximately -17 dB). However, it is observed in Fig. 7 that mode S2 is well excited at 1.3 and 2.6 MHz, disappearing between 2.3 and 2.4 MHz.

Those behaviors are due to the attenuation at such frequencies of the transfer functions plotted in Fig. 9.

As Eq. (28) states, if we perform the product $|H_{xx}^{(n)}(\omega)|^2 |\tilde{T}_{xx}(k_n, \omega)|^2$ we obtain the square modulus of the displacement response $|\tilde{U}_x^{(n)}(k_n, \omega)|^2$ for the n th mode in terms of frequencies.

In order to compare the theoretical displacement response with the results obtained from the square modulus of the measured normal stress intensities shown in Fig. 7, we must consider that the hydrophone signals are proportional to the pressure field—i.e., normal stress through liquid coupling. By assuming that the hydrophone is very close to the

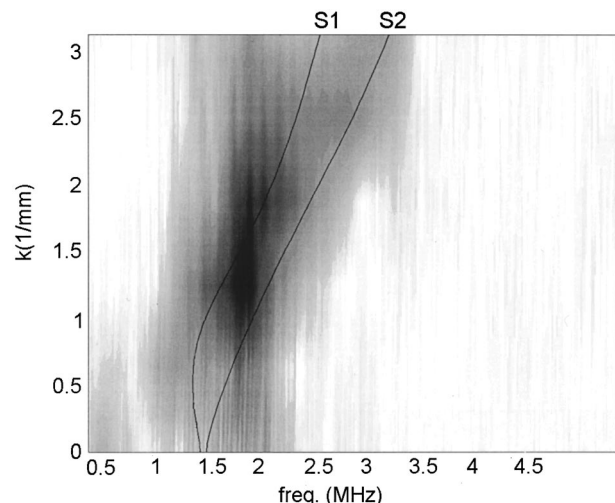


FIG. 10. Square modulus of normal excitation stress intensity on the base of the wedge.

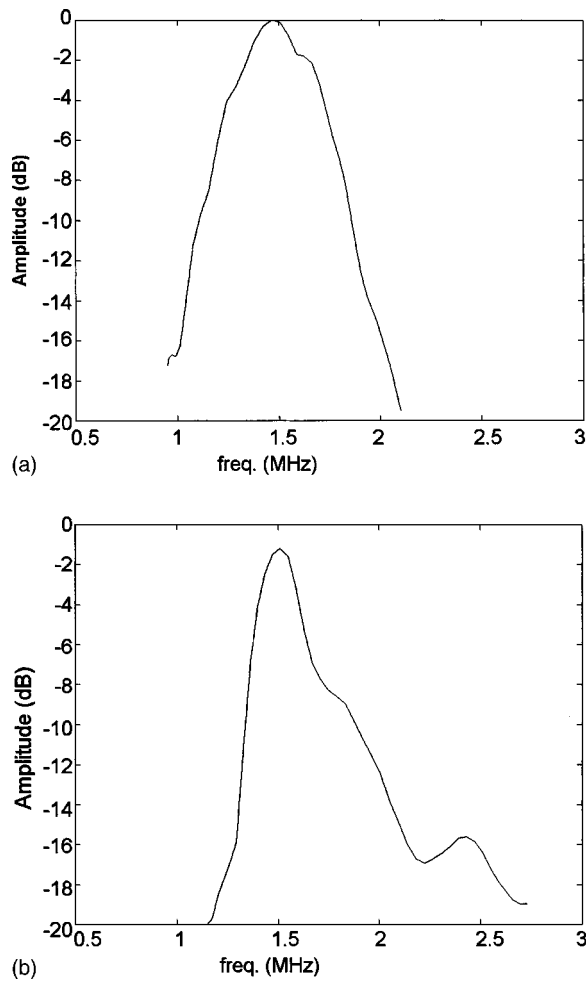


FIG. 11. (a) Square modulus of normal excitation stress intensity generated by Lamb mode S1. (b) Square modulus of normal excitation stress intensity generated by Lamb mode S2 with a 2.25-MHz central frequency transducer and 30° incidence angle.

plate surface, we can assume that there is a plane wave traveling from the plate surface to the hydrophone. The amplitude of the pressure is then proportional to the amplitude of the fluid velocity, which is the same as the amplitude of the normal velocity response of the Lamb waves at the plate surface. The relationship between the displacement amplitude $\tilde{U}(\omega)$ and the pressure amplitude response $\tilde{R}_{xx}(\omega)$ is then $\tilde{R}_{xx}(\omega) \propto \omega \tilde{U}(\omega)$. In order to compare the normal stress excitation $\tilde{T}_{xx}(k_n, \omega)$ and the pressure response $\tilde{R}_{xx}(\omega)$, Eq. (28) can be written as

$$\tilde{R}_{xx}^{(n)}(k_n, \omega) \propto \omega H_{xx}^{(n)}(\omega) \tilde{T}_{xx}(k_n, \omega). \quad (35)$$

The resulting normal stress intensity calculations $|\tilde{R}_{xx}(k_n, \omega)|^2$ arising from Eq. (35) are plotted in solid lines in Fig. 12(a) for mode S1 and Fig. 12(b) for mode S2.

The experimental normal stress intensity responses for S1 and S2 modes shown in Fig. 7 and restricted to curves $k = k_n(\omega)$, are respectively, plotted in Fig. 12(a) and (b) using “+” signs.

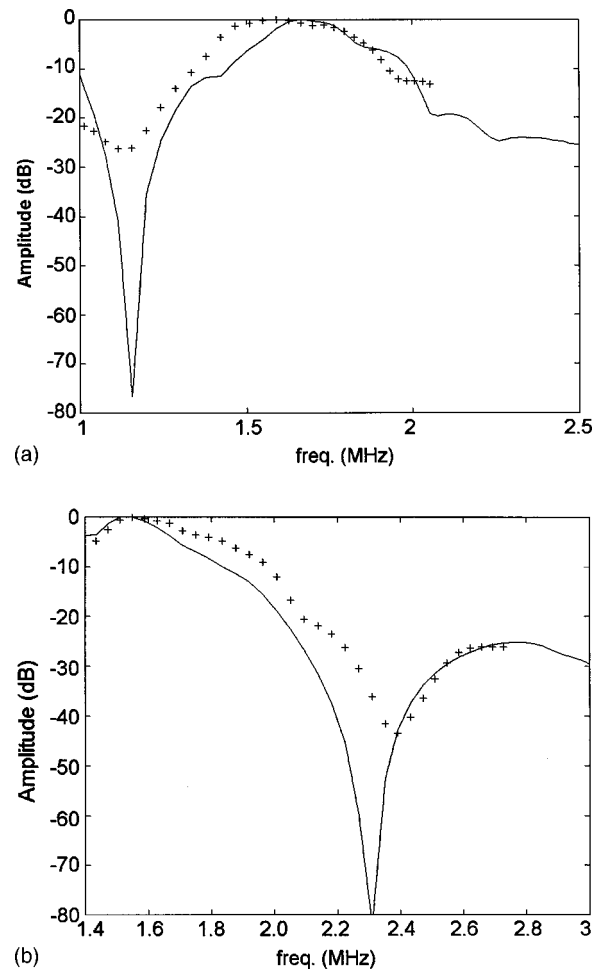


FIG. 12. (a) Comparison between theoretical (solid line) and experimental (+ points) square modulus of normal excitation stress intensities generated by mode S1 in a 3-mm thickness aluminum plate. (b) Comparison between theoretical (solid line) and experimental (+ points) square modulus of normal excitation stress intensities generated by mode S2 in a 3-mm thickness aluminum plate.

The agreement is relatively good. It would be better if the plate thickness were exactly 3.0 mm in its whole extension and the stress measured over the wedge base were the same as the stress that really occurs when the wedge is applied onto the surface.

In a second experiment, we chose a 4.0-mm aluminum plate, a transducer with a central frequency of 2.5 MHz, a bandwidth frequency of 1.5 MHz, a diameter of 1.0 in., and a 25° incidence angle (θ in Fig. 5). With this new arrangement, we tried to place the central region of the strong excitation just on the theoretical strong attenuation region of mode S2.

The transfer function $|H_{xx}^{(n)}(\omega)|^2$ for mode S2 for this new plate width is plotted in Fig. 13. We see that the attenuation peak is now around 1.7 MHz.

The square modulus of the normal excitation stress $|\tilde{T}_{xx}(k_n, \omega)|^2$ measured over the base of the wedge according to the dispersion curve of mode S2 is plotted in Fig. 14. We notice that the strong excitation region in Fig. 14 matches the attenuation peak shown in Fig. 13.

Here again we measure, using the hydrophone in the z direction through 108-mm length toward the source by 1-mm

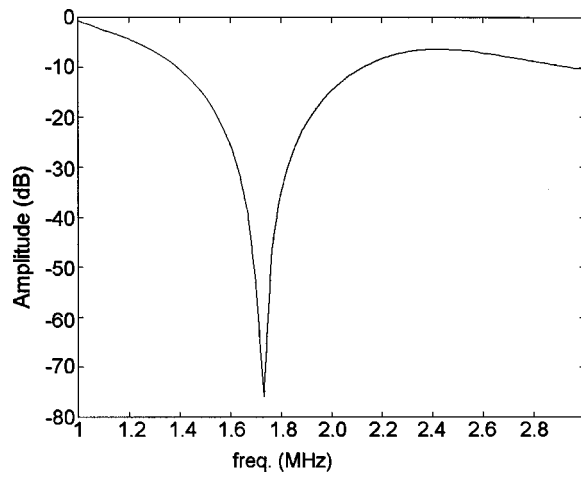


FIG. 13. Transfer function $|H_{xx}|^2$ for mode S2 in a 4-mm thickness aluminum plate.

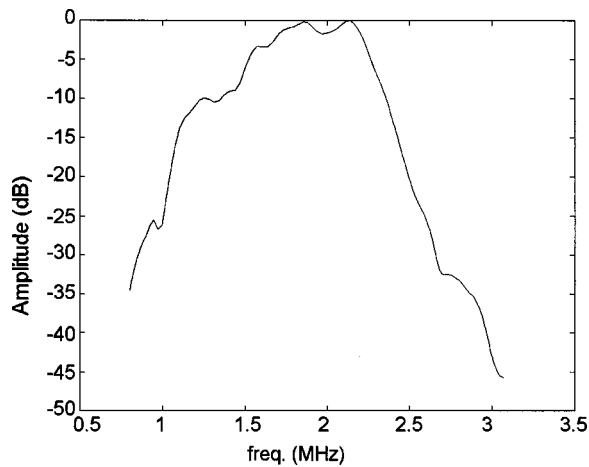


FIG. 14. Square modulus of normal excitation stress intensity generated by Lamb mode S2 with a 2.5-MHz central frequency transducer and 25° incidence angle.

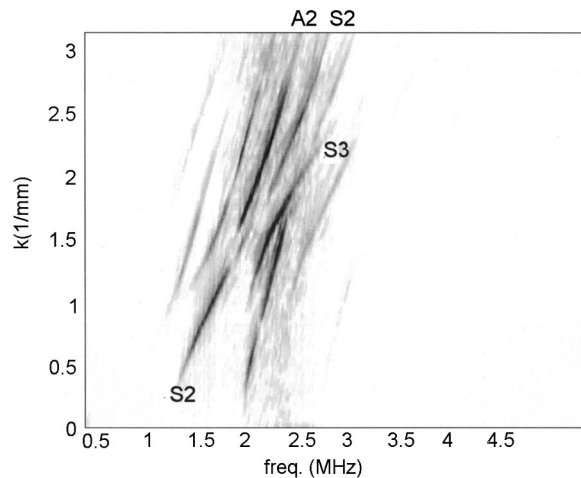


FIG. 15. Experimental dispersion curves of Lamb modes for a 4-mm thickness aluminum plate.

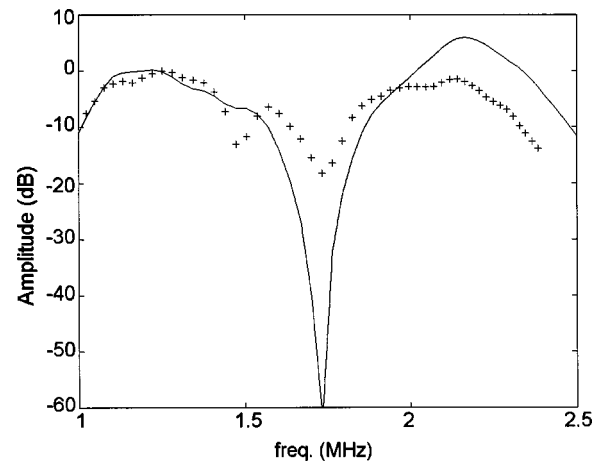


FIG. 16. Comparison between theoretical (solid line) and experimental (+ points) square modulus of normal excitation stress intensity generated by mode S3 in a 4-mm thickness aluminum plate.

steps the normal stresses generated by the Lamb waves that propagate inside the plate. Performing a double-Fourier transform of the measured signals, we obtain intensities of the normal stress generated by the Lamb waves $|\tilde{R}_{xx}(k_n, \omega)|^2$ (Fig. 15). We can observe that, despite the strong excitation around 1.7 MHz, the Lamb mode S2 is not present in such a region.

Figure 16 compares the experimental responses $|\tilde{R}_{xx}(k_n, \omega)|^2$ according to the dispersion curves $k = k_n(\omega)$ for mode S2 (+ points) with theoretical calculations arising from the second member of Eq. (35) (solid line). We see again that the agreement is relatively good.

III. CONCLUSIONS

We have deduced transfer and Green functions for Lamb waves generation, based on the normal-mode theory. The theoretical results completely agreed with a particular case analyzed by Viktorov. There was also reasonable agreement between two experiments performed using the wedge-transducer method and theory. The tensorial function has some advantages:

- It is easy to manipulate. Once the transfer function has been calculated, it remains the same for any kind of Lamb wave surface generation.
- As it is a tensorial function, it is not restricted to wedge-transducer generation with liquid coupling (only normal stress). It prevents Lamb waves generated by shear stress or tangential dilatations on the plate, allowing to determine the tangential excitations that exist in the Lamb wave. In the case of wedge-transducer generation, our numerical results are only approximate. This happens because the stresses assumed to exist at the wedge base differ from those measured (when the wedge is applied at the plate surface) because of reflections.
- It allows the detection of mode conversion due to defects in the plate. The theoretically calculated modes can be compared with the measured displacements or stresses on the plate.

- ¹D. N. Alleyne and P. Cawley, "The interaction of Lamb waves with defects," *IEEE Trans. Ultrason. Ferroelectr. Freq. Control* **39**, 381–397 (1992).
- ²S. I. Rokhlin, "Lamb wave interaction with lap-shear adhesive joint: Theory and experiment," *J. Acoust. Soc. Am.* **89**, 2758–2765 (1991).
- ³T. L. Mansfield, "Lamb wave inspection of aluminum sheet," *Mater. Eval.* **33**, 535–541 (1990).
- ⁴I. A. Viktorov, *Rayleigh and Lamb Waves* (Plenum, New York, 1967).
- ⁵J. J. Ditri and J. L. Rose, "Excitation of guided waves in generally anisotropic layers using finite sources," *Trans. ASME* **61**, 330–338 (1994).
- ⁶B. A. Auld, *Acoustic Fields and Waves in Solids*, Vol. II (Krieger, Malabar, Florida, 1990).
- ⁷G. S. Kino, *Acoustic Waves: Devices, Imaging, and Analog Signal Processing* (Prentice-Hall, Englewood Cliffs, NJ, 1987).
- ⁸X. Jia, "Modal analysis of Lamb wave generation in elastic plates by liquid wedge transducers," *J. Acoust. Soc. Am.* **101**, 834–842 (1997).
- ⁹D. N. Alleyne and P. A. Cawley, "A two-dimensional Fourier transform method for the measurement of propagating multimode signals," *J. Acoust. Soc. Am.* **89**, 1159–1168 (1991).
- ¹⁰R. K. Ing and M. Fink, "Time reversed Lamb waves," *IEEE Trans. Ultrason. Ferroelectr. Freq. Control* **45**(4), 1032–1043 (1998).

Pseudostatic corrections for the forced vibroacoustic response of a structure-cavity system

Michel Tournour^{a)} and Noureddine Atalla^{b)}

G.A.U.S., Mechanical Engineering Department, Université de Sherbrooke, Sherbrooke, Québec
J1K 2R1, Canada

(Received 18 January 1999; revised 6 April 1999; accepted 7 June 1999)

The analysis of the coupled vibroacoustic behavior of a structure–cavity system is often performed using the *in vacuo* structure modes and the rigid cavity modes. Unfortunately, the use of such a modal basis can result in a poor convergence when the high-frequency modes of one of the two subsystems are coupled to the low-frequency modes of the other subsystem. This problem is made critical by the lack of a reliable criterion for selecting the number of kept modes for each subsystem. This paper shows that the effect of modal truncation is critical for the convergence of the method and that the convergence can be greatly improved using pseudostatic corrections for both the structure and the cavity. The theory behind the proposed technique is presented together with two generic problems exhibiting strong coupling: (i) an elastic cylindrical cavity filled with air; and (ii) an elastic plate coupled to a rectangular cavity filled with water. © 2000 Acoustical Society of America. [S0001-4966(99)03009-X]

PACS numbers: 43.20.Tb, 43.40.Rj [CBB]

INTRODUCTION

The direct resolution of a forced structure–cavity system obtained from a finite element approach is time consuming. To alleviate this problem, modal approaches are often used. Unfortunately, in the classical displacement–pressure (u, p) formulation, the coupled eigen problem is large, full, and not symmetric. Thus, computation time may be significant.

Alternatives to the (u, p) nonsymmetric approach have been developed. Symmetrization techniques, such as those proposed by Irons,¹ Felippa,² and Kanarachos and Antoniadis³ are time consuming. Another alternative, leading directly to symmetric matrices, consists of using the field displacement as the fundamental unknown.^{4–6} For three-dimensional problems, this approach results in three degrees of freedom per node rather than one for the pressure-based approach. Thus, this approach leads to a larger system of equations. Furthermore, this formulation is complicated by the fact that the rotational-free condition for the displacement in the fluid domain must be imposed as an extra kinematic condition.⁷ Another alternative has been developed which uses the velocity potential as the unknown for the fluid.^{8,9} However, because of the relationship between the pressure and the velocity potential, this system is of a gyroscopic type which renders its solution expensive. Another alternative leading to a symmetric eigenvalue problem uses both the pressure and displacement potential as unknowns in the fluid.¹⁰ After discretization, this formulation leads to a larger but symmetric system of equations which reduces the computation time for the coupled modes.

However, the widely used technique in structure–cavity interaction problems is based on the use of the *in vacuo*

structure modes and the rigid cavity modes. Unfortunately, there is no reliable criterion for choosing the number of kept modes for each subsystem. Since it is impossible to determine, *a priori*, which modes are coupled together, the modal superposition method can lead to significant errors and may require several iterations before reaching convergence.^{11,12} Bouhioi¹³ and Tournour *et al.*¹⁴ used the *in vacuo* structure modes and the rigid cavity modes to obtain the coupled structure–cavity modes. This class of methods also suffers from the modal truncation of the two subsystems and can also lead to significant errors for strongly coupled problems.

The use of pseudostatic correction is classically used to improve the accuracy of the modal superposition technique for elastic structures.^{15,16} For a coupled structure–cavity system, even if this technique may improve the convergence of the structure indicators (e.g., displacement field, quadratic velocity), it does not guarantee the convergence of the fluid indicators (e.g., quadratic pressure). The use of pseudostatic corrections for bounded fluid volumes has been also considered in a recent book by Ohayon and Soize.¹⁷ It is shown in the present paper that using pseudostatic corrections for both the fluid and the structure subsystems improves significantly the convergence of the structure and fluid indicators. The pseudostatic corrections allow for a simple and good approximation for the dropped modes without increasing the dimension of the governing system of equations.

In the following section, the use of the uncoupled modal basis is first recalled. Based on two generic case studies depicting strong structure–cavity coupling, it is shown that such a technique can lead to important errors.

In Sec. II, a pseudostatic corrections technique is presented for both subsystems. It is shown that the presented technique improves significantly the convergence of the uncoupled modal superposition approach.

^{a)}Currently at LMS Numerical Technologies, Interleuvenlaan 70, 3001 Leuven, Belgium.

^{b)}Electronic mail: noureddine.atalla@gme.usherb.ca

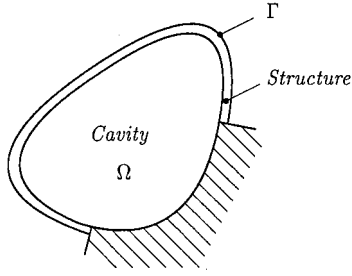


FIG. 1. The internal coupled fluid–structure problem.

I. THE MODAL SYNTHESIS METHOD

A. The problem

The discretization of the governing equations of a undamped structure–cavity system leads to the following system of equations, Fig. 1:^{10,18}

$$\begin{bmatrix} \mathbf{K} - \omega^2 \mathbf{M} & \mathbf{C} \\ \mathbf{C}^T & \frac{1}{\omega^2} \mathbf{H} - \mathbf{Q} \end{bmatrix} \begin{Bmatrix} \mathbf{u} \\ \mathbf{p} \end{Bmatrix} = \begin{Bmatrix} \mathbf{F} \\ \frac{1}{\omega^2} \mathbf{S} \end{Bmatrix}, \quad (1)$$

where \mathbf{K} and \mathbf{M} are the structure stiffness matrix and mass matrix, respectively, \mathbf{H} and \mathbf{Q} are the cavity kinetic energy matrix and compressibility matrix, respectively, and \mathbf{C} is the coupling matrix. The mass density and the sound velocity of the fluid are implicitly taken into account in the fluid matrices. Vectors \mathbf{u} , \mathbf{p} , \mathbf{F} , and \mathbf{S} are the structure displacement vector, the cavity pressure vector, the mechanical forces applied to the structure, and the acoustical sources in the cavity, respectively. Finally, scalar ω is the angular frequency.

In general, the system governed by Eq. (1) includes damping effects. Structural damping models have been retained for the present paper, but this choice is not restrictive and any other model may be selected. Using these models, the damping effects are taken into account using a complex stiffness matrix and a complex compressibility matrix. System (1) is then transformed into

$$\begin{bmatrix} \tilde{\mathbf{K}} - \omega^2 \mathbf{M} & \mathbf{C} \\ \mathbf{C}^T & \frac{1}{\omega^2} \mathbf{H} - \tilde{\mathbf{Q}} \end{bmatrix} \begin{Bmatrix} \mathbf{u} \\ \mathbf{p} \end{Bmatrix} = \begin{Bmatrix} \mathbf{F} \\ \frac{1}{\omega^2} \mathbf{S} \end{Bmatrix}, \quad (2)$$

where $\tilde{\mathbf{K}}$ and $\tilde{\mathbf{Q}}$ are the complex stiffness matrix and complex compressibility matrix, respectively. In the following, the tilde indicates a complex matrix, its imaginary part being a result of our choice for the damping model.

B. The uncoupled basis

The direct resolution of system (2) is time consuming. The undamped *in vacuo* structure modes and the undamped rigid cavity modes are commonly used to perform a modal reduction and alleviate this problem. Using this approach, the structure displacement and the cavity pressure are written in terms of the system uncoupled modes

$$\mathbf{u} = \Phi_s \mathbf{q}_s, \quad (3)$$

$$\mathbf{p} = \Phi_c \mathbf{q}_c, \quad (4)$$

where Φ_s and Φ_c are the undamped (real) kept—i.e., retained—modes of the structure and cavity, respectively; and where \mathbf{q}_s and \mathbf{q}_c are the complex generalized coordinates of the structure and cavity, respectively. Using Eqs. (3) and (4) together with classical normalizations with respect to the mass matrix \mathbf{M} and compressibility matrix \mathbf{Q} , system (2) transforms to the following system:

$$\begin{bmatrix} \tilde{\Lambda}_s - \omega^2 \mathbf{I} & \mathbf{C}_m \\ \mathbf{C}_m^T & \frac{1}{\omega^2} \Lambda_c - \tilde{\mathbf{I}} \end{bmatrix} \begin{Bmatrix} \mathbf{q}_s \\ \mathbf{q}_c \end{Bmatrix} = \begin{Bmatrix} \mathbf{F}_g \\ \frac{1}{\omega^2} \mathbf{S}_g \end{Bmatrix}, \quad (5)$$

where $\tilde{\Lambda}_s$ is the generalized complex stiffness matrix, Λ_c is the generalized kinetic energy matrix, $\tilde{\mathbf{I}}$ is the generalized complex compressibility matrix, \mathbf{C}_m is the modal coupling matrix, \mathbf{F}_g is the generalized force vector, and \mathbf{S}_g is the generalized acoustic source vector. The real part of matrix $\tilde{\mathbf{I}}$ is the identity matrix. For a modal damping model, matrices $\tilde{\Lambda}_s$, Λ_c , and $\tilde{\mathbf{I}}$ are diagonal. Once Eq. (5) is solved, Eqs. (3) and (4) are used to calculate the vibroacoustic indicators. In this paper, two vibroacoustic indicators are used to measure the performance of the presented methods: the mean-square normal velocity and the mean-square pressure. The mean-square normal velocity is defined as

$$\langle \overline{V^2} \rangle = \frac{1}{2\Gamma} \int_{\Gamma} |V_n|^2 d\Gamma, \quad (6)$$

where Γ is the structure surface and V_n is the normal velocity. Similarly, the mean-square pressure is defined as

$$\langle \overline{p^2} \rangle = \frac{1}{2\Omega} \int_{\Omega} |p|^2 d\Omega, \quad (7)$$

where Ω is the cavity volume.

Finally, for a spectral analysis, a commonly used rule of thumb is to retain all modes up to $1.5 \times f_{\max}$ where f_{\max} is the highest frequency of interest. This widely used criterion is not reliable. Two examples are presented in the following section to show its limitations.

C. Performance of the method

In order to test the performance of the method, two generic examples are presented. The first example is a plate-backed cavity filled with water. The second example is an elastic cylindrical cavity filled with air and closed with two rigid caps. For both cases, the mean-square normal velocity and the mean-square pressure obtained using the modal superposition are compared to the results obtained using a direct approach.

1. Rectangular cavity filled with water

A plate-backed cavity is considered. The plate is simply supported and has the following dimensions: 35 cm long, 29 cm wide, and 0.15 cm thick. The cavity has the following dimensions: 35 cm long, 29 cm wide, and 14 cm deep. The plate is made from aluminum with the following mechanical properties: Young's modulus $E = 0.72 \times 10^{11}$ N/m², mass

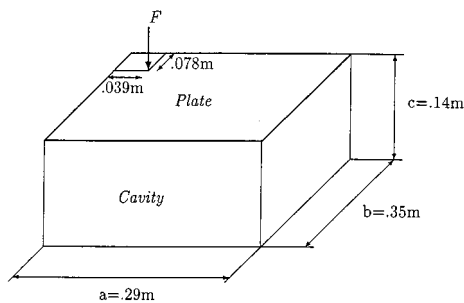


FIG. 2. The plate-backed cavity system.

density $\rho_s = 2700 \text{ kg/m}^3$, Poisson's ratio $\nu = 0.3$, and damping factor $\eta_s = 0.01$. The cavity is filled with water having the following properties: speed of sound $c = 1500 \text{ m/s}$, mass density $\rho_c = 1000 \text{ kg/m}^3$, and damping factor $\eta_c = 0.01$. The system is excited using a normal unit point load defined in Fig. 2.

The plate is discretized using (9×8) 8-node quadrilateral Mindlin elements and the cavity is discretized using $(9 \times 8 \times 4)$ 20-node brick elements.

Figure 3(a) shows the plate mean-square normal velocity obtained using two different approaches. The first approach consists of solving Eq. (2) directly. The second approach consists of solving Eq. (5) using 50 plate modes and 50 cavity modes. More details about the modal bases are given in Table I. The frequency range of interest is 10–600 Hz. It can be observed that, although the highest structural-kept mode has a natural frequency of 2800 Hz and the highest cavity-kept mode has a natural frequency of 12 000 Hz, the modal approach has not yet converged. Indeed, since the modal coupling is proportional to the square root of the bulk modulus of the fluid, the presence of a heavy fluid results in a very strong coupling between the plate modes and the cavity modes. This is also true for the coupling between the lower plate modes and the higher cavity modes, usually low but drastically increased by the high bulk modulus of the fluid. To reach convergence, the modal superposition method must include an extremely high number of cavity modes, making the method impractical.

Figure 3(b) shows the mean-square pressure in the cav-

ity using the two approaches. Once again, it can be observed that the modal approach has not yet converged.

2. Cylindrical cavity filled with air

A steel cylinder with rigid caps clamped at its ends is considered here. The cylinder has the following dimensions: 101 cm long, 18.256 cm radius, and 0.1219 cm thick. The steel has the following properties: Young's modulus $E = 2.1 \times 10^{11} \text{ N/m}^2$, mass density $\rho_s = 7800 \text{ kg/m}^3$, Poisson's ratio $\nu = 0.3$, and damping factor $\eta_s = 0.0006$. The cavity is filled with air having the following properties: speed of sound $c = 340 \text{ m/s}$, mass density $\rho_c = 1.21 \text{ kg/m}^3$, and damping factor $\eta_c = 0.0005$. The system is excited using a normal unit point load defined in Fig. 4.

The cylinder is discretized using 28 4-node quadrilateral Mindlin elements along the perimeter and 16 elements along the length. The cavity is discretized using 6-node wedge elements with 28 nodes along the perimeter, 17 nodes along the length, and 15 nodes along the diameter.

Figure 5(a) shows the cylinder mean-square normal velocity obtained using the two different approaches defined previously. Forty-two cylinder modes and eighty cavity modes are used in the modal approach. More details about the modal bases are given in Table I. There is full agreement between the two techniques, proving the ability of the kept modes in synthesizing correctly the structural response of the system.

Figure 5(b) shows the mean-square pressure in the cavity using the two approaches. It can be observed that the modal approach has not yet converged and that the contribution of several modes have been clearly missed by the modal technique. It can also be observed that the cavity-dominated modes that were not present in the mean-square velocity appear in the mean-square pressure (e.g., rigid cavity mode and first elastic cavity mode around 168 Hz).

Figure 6 shows the maximum coupling between the cavity modes and the kept cylinder modes (obtained from matrix C_m). It can be observed that the first cavity modes are almost orthogonal to the kept cylinder modes. Therefore, they cannot contribute to the mean-square pressure in the modal approach, thus leading to the failure of this method. Boily and Charron¹² studied the same cavity, but with free boundary

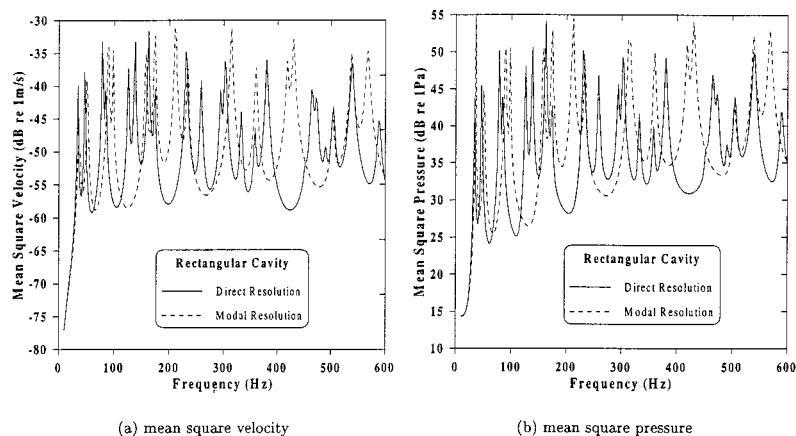


FIG. 3. Response of the plate-backed cavity system using the modal approach.

TABLE I. Description of the modal bases.

	Structure			Cavity		
	Number of dof	Number of modes	Maximum frequency	Number of dof	Number of modes	Maximum frequency
Rectangular cavity	901	50	2800 Hz	1615	50	12 000 Hz
Cylindrical cavity	2520	42	760 Hz	3468	80	1700 Hz

conditions. They performed a convergence study of the mean quadratic pressure around the first elastic cavity mode. They showed that even using 572 structural modes (more than one-quarter of the total number of degrees of freedom (DOF) of their structural model), the modal superposition technique did not reach convergence. They conclude that almost every structural modes should be retained in the modal basis, rendering the modal solution ineffective.

The problem at hand seems to be more complex. It is due to the theoretical formulation of the problem. For instance, in this example, since the rigid cavity mode and several elastic modes are almost orthogonal to the structure mode, they can never be captured using the modal synthesis technique as classically used. However, from a physical viewpoint, these modes should contribute to the response, at least statically.

II. THE PROPOSED TECHNIQUE

A. Pseudostatic corrections

From the previous examples, it seems obvious that the dropped modes need to be taken into account. To assess their contributions, a technique based on pseudostatic corrections of the modal bases is proposed. The pressure and the displacements in terms of kept and dropped modes are given by

$$\mathbf{p} = \Phi_c \mathbf{q}_c + \Phi_c^d \mathbf{q}_c^d = \Phi_c \mathbf{q}_c + \mathbf{p}_d, \quad (8)$$

$$\mathbf{u} = \Phi_s \mathbf{q}_s + \Phi_s^d \mathbf{q}_s^d = \Phi_s \mathbf{q}_s + \mathbf{u}_d, \quad (9)$$

where the superscript d stands for the dropped modes. Using these equations and the modes' orthogonality properties, system (2) leads to the following system for the dropped modes:

$$\tilde{\Lambda}_s^d \mathbf{q}_s^d - \omega^2 \mathbf{q}_s^d = \Phi_s^{dT} (\mathbf{F} - \mathbf{Cp}), \quad (10)$$

$$\Lambda_c^d \mathbf{q}_c^d - \omega^2 \tilde{\Gamma}^d \mathbf{q}_c^d = \Phi_c^{dT} (\mathbf{S} - \omega^2 \mathbf{C}^T \mathbf{u}). \quad (11)$$

A pseudostatic correction for the dropped modes is sought. Consequently, the following approximations are used:

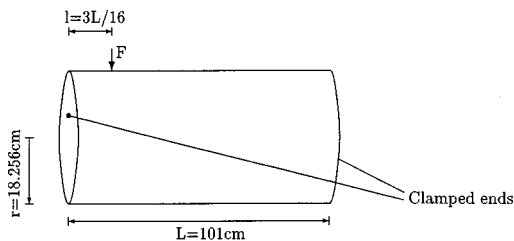


FIG. 4. The elastic cylinder-cavity system.

$$\tilde{\Lambda}_s^d - \omega^2 \mathbf{I} \approx \tilde{\Lambda}_s^d, \quad (12)$$

$$\Lambda_c^d - \omega^2 \tilde{\Gamma}^d \approx \Lambda_c^d, \quad (13)$$

and therefore Eqs. (10) and (11) reduce to

$$\tilde{\Lambda}_s^d \mathbf{q}_s^d \approx \Phi_s^{dT} (\mathbf{F} - \mathbf{Cp}), \quad (14)$$

$$\Lambda_c^d \mathbf{q}_c^d \approx \Phi_c^{dT} (\mathbf{S} - \omega^2 \mathbf{C}^T \mathbf{u}). \quad (15)$$

Equations (14) and (15) show that the kinetic energy of the dropped displacement modes and the compressibility energy of the dropped pressure modes are neglected. The contribution of the dropped modes is thus said to be pseudostatic.

In terms of nodal variables, the pseudostatic displacements and pseudostatic pressure are given by

$$\tilde{\mathbf{K}} \mathbf{u}^0 = \mathbf{F} - \mathbf{Cp}, \quad (16)$$

$$\mathbf{H} \mathbf{p}^0 = \mathbf{S} - \omega^2 \mathbf{C}^T \mathbf{u}. \quad (17)$$

Before looking at the effect of rigid modes, we first assume matrices $\tilde{\mathbf{K}}$ and \mathbf{H} to be positive definite, therefore

$$\mathbf{u}^0 = \tilde{\mathbf{K}}^{-1} (\mathbf{F} - \mathbf{Cp}), \quad (18)$$

$$\mathbf{p}^0 = \mathbf{H}^{-1} (\mathbf{S} - \omega^2 \mathbf{C}^T \mathbf{u}). \quad (19)$$

Furthermore, the pseudostatic displacements can be expanded in terms of kept modes and dropped modes

$$\mathbf{u}^0 = \mathbf{u}_k^0 + \mathbf{u}_d, \quad (20)$$

with

$$\mathbf{u}_k^0 = \Phi_s \tilde{\Lambda}_s^{-1} \Phi_s^T (\mathbf{F} - \mathbf{Cp}). \quad (21)$$

Similarly, the pseudostatic pressure can be expanded in terms of kept modes and dropped modes

$$\mathbf{p}^0 = \mathbf{p}_k^0 + \mathbf{p}_d, \quad (22)$$

with

$$\mathbf{p}_k^0 = \Phi_c \Lambda_c^{-1} \Phi_c^T (\mathbf{S} - \omega^2 \mathbf{C}^T \mathbf{u}). \quad (23)$$

Using Eqs. (18)–(23), new expressions for the displacement and the pressure due to the dropped modes are obtained

$$\mathbf{u}_d = (\tilde{\mathbf{K}}^{-1} - \Phi_s \tilde{\Lambda}_s^{-1} \Phi_s^T) (\mathbf{F} - \mathbf{Cp}), \quad (24)$$

$$\mathbf{p}_d = (\mathbf{H}^{-1} - \Phi_c \Lambda_c^{-1} \Phi_c^T) (\mathbf{S} - \omega^2 \mathbf{C}^T \mathbf{u}). \quad (25)$$

Using pseudostatic corrections of the first order only, i.e., approximation (3) is used for the evaluation of the displacement \mathbf{u} in the right-hand side of Eq. (25) and approximation (4) is used for the evaluation of the pressure \mathbf{p} in the

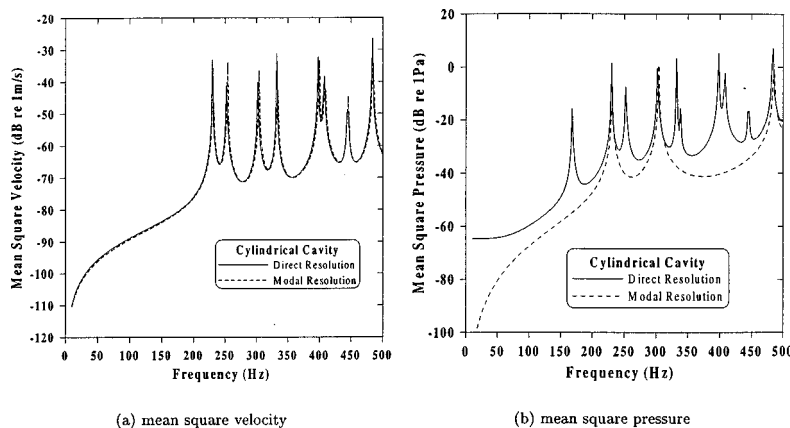


FIG. 5. Response of the cylinder-cavity system using the modal approach.

right-hand side of equation (24), the final expressions for the displacements and the pressure due to the dropped modes are given by

$$\mathbf{u}_d = (\tilde{\mathbf{K}}^{-1} - \Phi_s \tilde{\Lambda}_s^{-1} \Phi_s^T) (\mathbf{F} - \mathbf{C} \Phi_c \mathbf{q}_c), \quad (26)$$

$$\mathbf{p}_d = (\mathbf{H}^{-1} - \Phi_c \Lambda_c^{-1} \Phi_c^T) (\mathbf{S} - \omega^2 \mathbf{C}^T \Phi_s \mathbf{q}_s). \quad (27)$$

Combining Eqs. (8), (9), (26), and (27) with Eq. (2), and using the modes' orthogonality properties, the new system for the coupled structure is obtained

$$\begin{bmatrix} \tilde{\Lambda}_s - \omega^2 \mathbf{I} - \omega^2 \mathbf{M}_{\text{am}} & \mathbf{C}_m \\ \mathbf{C}_m^T & \frac{1}{\omega^2} \Lambda_c - \tilde{\mathbf{I}} - \tilde{\mathbf{Q}}_{\text{am}} \end{bmatrix} \begin{Bmatrix} \mathbf{q}_s \\ \mathbf{q}_c \end{Bmatrix} = \begin{Bmatrix} \mathbf{F}_g - \mathbf{G}_{\text{cm}} \mathbf{S} \\ \frac{1}{\omega^2} \mathbf{S}_g - \tilde{\mathbf{G}}_{\text{sm}} \mathbf{F} \end{Bmatrix}, \quad (28)$$

where \mathbf{M}_{am} and $\mathbf{G}_{\text{cm}} \mathbf{S}$ are the added mass and loading due to the inertia of the dropped cavity modes, respectively and where $\tilde{\mathbf{Q}}_{\text{am}}$ and $\tilde{\mathbf{G}}_{\text{sm}} \mathbf{F}$ are the added compressibility and loading due to the stiffness of the dropped structure modes, respectively.

The modal residual added mass matrix is given by

$$\mathbf{M}_{\text{am}} = \Phi_s^T \mathbf{C} \mathbf{H}_{\text{res}}^{-1} \mathbf{C}^T \Phi_s, \quad (29)$$

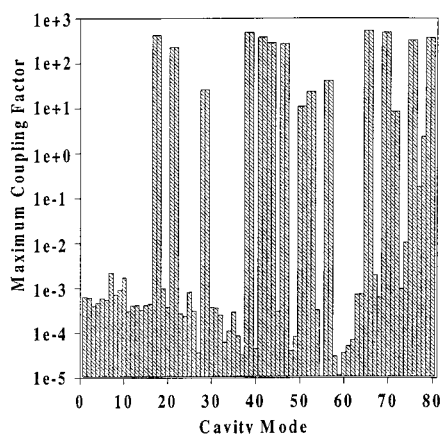


FIG. 6. Maximum coupling between the cavity modes and the kept cylinder modes.

with

$$\mathbf{H}_{\text{res}}^{-1} = \mathbf{H}^{-1} - \Phi_c \Lambda_c^{-1} \Phi_c^T. \quad (30)$$

Similarly, the modal residual added compressibility matrix is given by

$$\tilde{\mathbf{Q}}_{\text{am}} = \Phi_c^T \mathbf{C}^T \tilde{\mathbf{K}}_{\text{res}}^{-1} \mathbf{C} \Phi_c, \quad (31)$$

with

$$\tilde{\mathbf{K}}_{\text{res}}^{-1} = \tilde{\mathbf{K}}^{-1} - \Phi_s \tilde{\Lambda}_s^{-1} \Phi_s^T. \quad (32)$$

Finally, matrices $\tilde{\mathbf{G}}_{\text{sm}}$ and \mathbf{G}_{cm} are given by the expressions

$$\tilde{\mathbf{G}}_{\text{sm}} = \Phi_c^T \mathbf{C}^T \tilde{\mathbf{K}}_{\text{res}}^{-1}, \quad (33)$$

$$\mathbf{G}_{\text{cm}} = \Phi_s^T \mathbf{C} \mathbf{H}_{\text{res}}^{-1}. \quad (34)$$

The solution of Eq. (28) leads to the structure displacement and pressure in the cavity using the two expressions

$$\mathbf{u} = \Phi_s \mathbf{q}_s + \mathbf{K}_{\text{res}}^{-1} \mathbf{F} - \tilde{\mathbf{G}}_{\text{sm}}^T \mathbf{q}_c, \quad (35)$$

$$\mathbf{p} = \Phi_c \mathbf{q}_c + \mathbf{H}_{\text{res}}^{-1} \mathbf{S} - \omega^2 \mathbf{G}_{\text{cm}}^T \mathbf{q}_s. \quad (36)$$

B. Presence of rigid body modes

In general, matrix \mathbf{H} is not positive definite (in the case of presence of a rigid cavity mode). The inversion of matrix \mathbf{H} consists then of the inversion of a constrained matrix \mathbf{H}_{con} and in the use of a filter \mathbf{A}_c to eliminate the rigid cavity mode¹⁹

$$\mathbf{A}_c = \mathbf{I} - \mathbf{Q} \phi_{\text{cr}} \phi_{\text{cr}}^T, \quad (37)$$

where ϕ_{cr} is the rigid cavity mode. Consequently, the residual matrix $\mathbf{H}_{\text{res}}^{-1}$ is defined as

$$\mathbf{H}_{\text{res}}^{-1} = \mathbf{A}_c^T \mathbf{H}_{\text{con}}^{-1} \mathbf{A}_c - \Phi_{\text{ce}} \Lambda_{\text{ce}}^{-1} \Phi_{\text{ce}}^T, \quad (38)$$

where the subscripts e stand for the elastic modes. Combining this expression with Eq. (29), a new expression for the modal residual added mass is obtained

$$\mathbf{M}_{\text{am}} = \mathbf{D}_{\text{cm}}^T \mathbf{H}_{\text{con}}^{-1} \mathbf{D}_{\text{cm}} - \mathbf{C}_m \Lambda_c^{-1} \mathbf{C}_m^T, \quad (39)$$

with

$$\mathbf{D}_{\text{cm}} = \mathbf{A}_c \mathbf{C}^T \Phi_s, \quad (40)$$

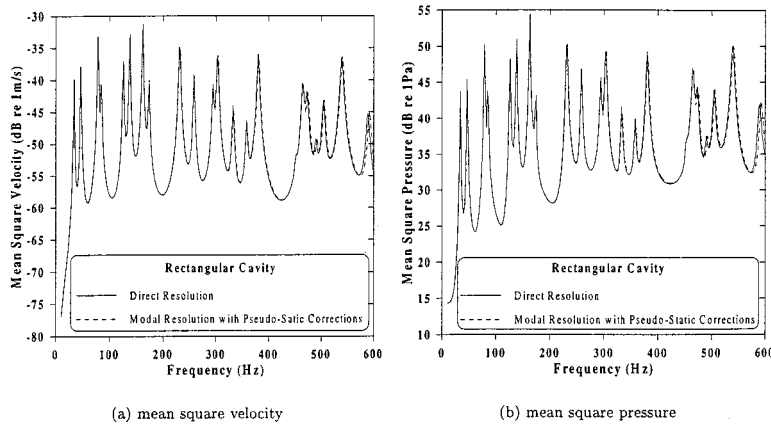


FIG. 7. Response of the plate-backed cavity system using the modal approach with pseudostatic corrections.

and where only the elastic cavity modes are taken into account in the correction term $\mathbf{C}_m \mathbf{\Lambda}_c^{-1} \mathbf{C}_m^T$.

Similarly, matrix $\tilde{\mathbf{K}}$ may be non-positive definite. The inversion of matrix $\tilde{\mathbf{K}}$ consists then of the inversion of a constrained matrix $\tilde{\mathbf{K}}_{\text{con}}$ and in the use of a filter \mathbf{A}_s to eliminate the rigid body modes¹⁹

$$\mathbf{A}_s = \mathbf{I} - \mathbf{M} \Phi_{\text{sr}} \Phi_{\text{sr}}^T, \quad (41)$$

where matrix Φ_{sr} contains the structure rigid body modes. Consequently, the residual flexibility $\tilde{\mathbf{K}}_{\text{res}}^{-1}$ is defined as

$$\tilde{\mathbf{K}}_{\text{res}}^{-1} = \mathbf{A}_s^T \tilde{\mathbf{K}}_{\text{con}}^{-1} \mathbf{A}_s - \Phi_{\text{se}} \tilde{\mathbf{\Lambda}}_{\text{se}}^{-1} \Phi_{\text{se}}^T. \quad (42)$$

Combining this expression with Eq. (31), a new expression for the modal residual added compressibility matrix is obtained

$$\tilde{\mathbf{Q}}_{\text{am}} = \mathbf{D}_{\text{sm}}^T \tilde{\mathbf{K}}_{\text{con}}^{-1} \mathbf{D}_{\text{sm}} - \mathbf{C}_m^T \tilde{\mathbf{\Lambda}}_s^{-1} \mathbf{C}_m, \quad (43)$$

with

$$\mathbf{D}_{\text{sm}} = \mathbf{A}_s \mathbf{C} \Phi_c, \quad (44)$$

and where only the elastic structure modes are taken into account in the correction term $\mathbf{C}_m^T \tilde{\mathbf{\Lambda}}_s^{-1} \mathbf{C}_m$.

C. Performance of the proposed method

In order to show the performance of the proposed method, the latter is used to solve the two examples studied

in Sec. I C. The same number of modes are used. The results obtained are compared to the results obtained using a direct resolution approach.

1. Rectangular cavity filled with water

Figure 7(a) shows the plate mean-square velocity obtained using the direct approach and the proposed technique. There is full agreement between the two methods. Comparing Figs. 3(a) and 7(a), it can be observed that the use of the pseudostatic corrections significantly improves the convergence of the modal approach. Indeed, the modal superposition fails to represent the fluid loading while the pseudostatic corrections allow for an accurate representation.

Figure 7(b) shows the cavity mean-square pressure obtained using the direct approach and the proposed technique. Excellent agreement is found. The use of the pseudostatic corrections solves the convergence problem of the classical modal approach [compare Figs. 3(b) and 7(b)].

In order to better understand the relative contribution of each term of the pseudostatic corrections, the response of the system is calculated using the modal approach with three different corrections:

- (i) with all the corrections (\mathbf{M}_{am} , \mathbf{G}_{cm} , $\tilde{\mathbf{Q}}_{\text{am}}$, and $\tilde{\mathbf{G}}_{\text{sm}}$),
- (ii) with the corrections for the dropped fluid modes only (\mathbf{M}_{am} and \mathbf{G}_{cm}),
- (iii) with the added mass only (\mathbf{M}_{am}).

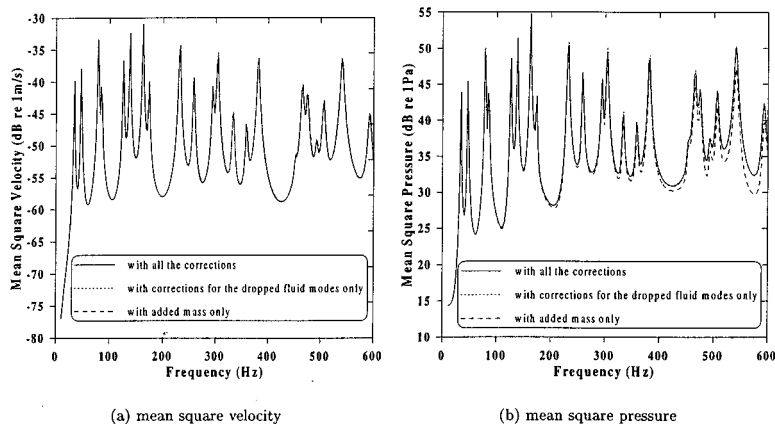


FIG. 8. Response of the plate-backed cavity system using the modal approach with different pseudostatic corrections.

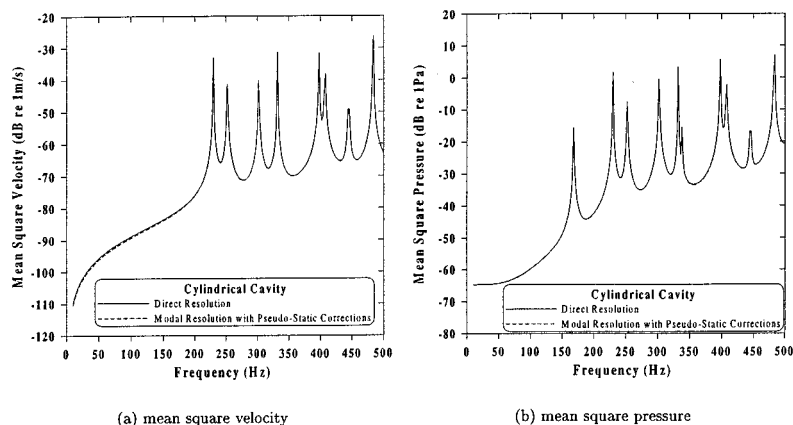


FIG. 9. Response of the cylinder-cavity system using the modal approach with pseudostatic corrections.

The mean-square velocity of the plate and the mean-square pressure in the cavity are given in Fig. 8(a) and 8(b), respectively. Figure 8(a) shows that only the added mass \mathbf{M}_{am} is necessary to correctly represent the fluid loading on the structure. On the other hand, Fig. 8(b) shows that the correction term \mathbf{G}_{cm} should be used in Eq. (36) to correctly assess the pressure in the cavity. Also, these results explain why the classical use of pseudostatic correction for the structure does not solve the problem.

2. Cylindrical cavity filled with air

Figure 9(a) and (b) show the cylinder mean-square velocity and the cavity mean-square pressure obtained using the direct approach and the proposed technique. There is full agreement between the two methods. Comparing Figs. 9(b) and 5(b), it can be observed that the use of the pseudostatic corrections eliminates the convergence problem of the modal approach. Indeed, as mentioned in Sec. IC 2, the kept cylinder modes are almost orthogonal to some cavity modes. This results in the nonparticipation of those modes. The pseudostatic corrections allow for a better representation of the coupling between the cavity and the cylinder.

Finally, as for the plate-backed cavity system, the response of the system is calculated using the modal approach with three different corrections:

- (i) with all the corrections (\mathbf{M}_{am} , \mathbf{G}_{cm} , $\tilde{\mathbf{Q}}_{am}$, and $\tilde{\mathbf{G}}_{sm}$),

- (ii) with the corrections for the dropped structure modes only ($\tilde{\mathbf{Q}}_{am}$ and $\tilde{\mathbf{G}}_{sm}$),
- (iii) with the added compressibility only ($\tilde{\mathbf{Q}}_{am}$).

The mean-square velocity of the cylinder and the mean-square pressure in the cavity are given in Fig. 10(a) and (b), respectively. Figure 10(b) shows the importance at low frequencies of the $\tilde{\mathbf{G}}_{sm}\mathbf{F}$ term in the right-hand side of Eq. (28) to correctly assess the contribution of the first two cavity modes. It also shows that using only a pseudostatic correction for the structure is not sufficient to accurately calculate the pressure in the cavity. Consequently, it demonstrates the importance of the corrections for the dropped cavity modes.

3. Computational performance of the method

In order to show the performance of the proposed method in terms of computation time, Table II shows the computation time of the direct approach, the modal approach, and the modal approach with pseudostatic corrections. For the modal approaches, the computation time includes the cost of modes calculation. It can be noted that, when correctly implemented, the added computation time of the correction terms is insignificant. Actually, the proposed technique can result in important time savings since fewer modes are necessary to represent the subsystems. More importantly, the convergence is guaranteed.

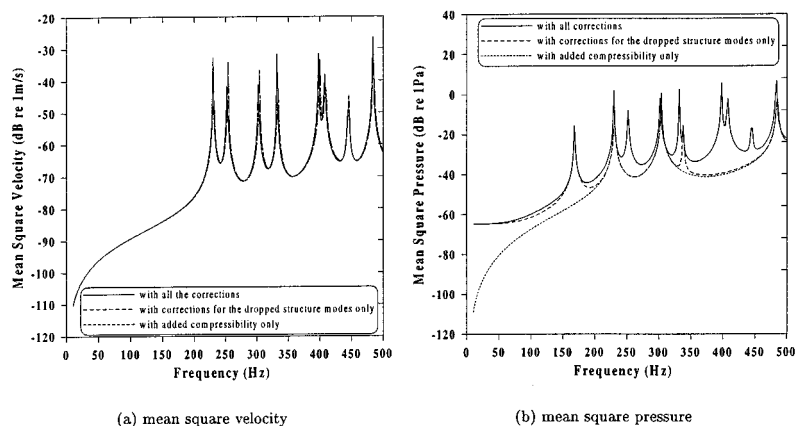


FIG. 10. Response of the cylinder-cavity system using the modal approach with different pseudostatic corrections.

TABLE II. Calculation times.

	Direct	Modal	Modal+Corrections
Plate	58 mn 37s	2 mn 35 s	2 mn 40 s
	100%	4.4%	4.6%
Cylinder	1h 47 mn	9 mn 42 s	10 mn 17 s
	100%	9.0%	9.5%

III. CONCLUSION

Modal synthesis techniques based on the finite element method are widely used in the structure–cavity interaction problems. This technique allows for a significant reduction of the total number of degrees of freedom and results in large time savings. Unfortunately, the method suffers from the lack of a reliable criterion for selecting the number of kept modes for each subsystem. Moreover, the use of a large number of modes, therefore a very strict criterion, does not guarantee the convergence of the method.

It has been shown that the use of pseudostatic corrections for both the structure and the cavity greatly improves the accuracy of the modal approach for the forced vibroacoustic response of a coupled structure–cavity system. Furthermore, it is expected that the method results in significant time savings since fewer modes are necessary to reach convergence.

ACKNOWLEDGMENTS

This work was funded by l'I.R.S.S.T.: the Institut de Recherche en Santé et Sécurité au Travail, Québec, Canada.

¹B. M. Irons, "Role of part inversion in fluid–structure problems with mixed variables," *AIAA J.* **7**(3), 568 (1970).

²C. A. Felippa, "Symmetrization of the contained compressible-fluid vibration eigenproblem," *Comm. Appl. Num. Meth.* **1**, 241–247 (1985).

³A. Karanachos and I. Antoniadis, "Symmetric variational principles and modal methods in fluid-structure interaction problems," *J. Sound Vib.* **121**(1), 77–107 (1988).

⁴M. A. Hamdi, "Étude des vibrations élasto-acoustique par une méthode d'éléments finis aux déplacements," Ph.D. thesis, Université de Technologie de Compiègne, Compiègne, France, 1978.

⁵A. J. Kalinowsky, "Fluid structure interaction," in *Shock and Vibration Computer Programs: Reviews and Summaries, SVM-10*, edited by W. Pilkey and B. Pilkey (The Shock and Vibration Information Center, Naval Research Laboratory, Washington, DC, 1975).

⁶L. Kiefing and G. C. Feng, "Fluid-structure finite element vibrational analysis," *AIAA J.* **14**(2), 199–203 (1976).

⁷H. C. Chen and R. L. Taylor, "Vibration analysis of fluid–solid systems using a finite element displacement formulation," *Int. J. Numer. Methods Eng.* **29**, 683–698 (1990).

⁸G. C. Everstine, "A symmetric potential formulation for fluid–structure interactions," *J. Sound Vib.* **107**, 121–129 (1981).

⁹L. G. Olson and K. J. Bathe, "Analysis of fluid–structure. A direct symmetric coupled formulation based on the fluid velocity potential," *Comput. Struct.* **21**, 21–32 (1985).

¹⁰H. J.-P. Morand and R. Ohayon, *Fluid Structure Interaction* (Wiley, New York, 1995).

¹¹F. Charron, R. Panneton, Y. Champoux, J.-M. Guérin, and S. Boily, "Analytical, numerical and experimental study of the vibro-acoustic behavior of a stiffened shell structure," *Proceedings of the 1995 ASME Design Engineering Technical Conference*, B(1), September 1995.

¹²S. Boily and F. Charron, "The vibroacoustic response of a cylindrical shell structure with viscoelastic and poroelastic materials," *Appl. Acoust.* **58**(2), 131–152 (1999).

¹³H. Bouhioui, "Étude vibroacoustique d'un montage en double paroi de verre," Ph.D. thesis, Université de Technologie de Compiègne, Compiègne, France, October 1993.

¹⁴M. Tournour and N. Atalla, "Vibroacoustic behavior of an elastic box using state-of-the-art fem–bem approaches," *Noise Control Eng. J.* **46**(3), 83–90 (1998).

¹⁵M. Petyt, *Introduction to Finite Element Vibration Analysis* (Cambridge University Press, Cambridge, 1990).

¹⁶O. E. Hansteen and K. Bell, "On the accuracy of mode superposition analysis in structural dynamics," *Int. J. Earthquake Eng. Struct. Dynamics* **7**, 405–411 (1979).

¹⁷R. Ohayon and C. Soize, *Structural Acoustics and Vibration* (Academic, San Diego, 1998).

¹⁸N. Atalla and R. J. Bernhard, "Review of numerical solutions for low-frequency structural-acoustic problems," *Appl. Acoust.* **43**, 271–294 (1994).

¹⁹M. G rardin and D. Rixen, *Mechanical Vibrations: Theory and Applications to Structural Dynamics* (Wiley, New York, 1997).

Calculation of acoustic radiation using equivalent-sphere methods

Laurent Bouchet and Thierry Loyau

Institut National de Recherche et de Sécurité (INRS), B.P. 27-54501 Vandoeuvre, France

Nacer Hamzaoui and Claude Boisson

Laboratoire Vibrations-Acoustique, INSA de Lyon, 69261 Villeurbanne Cedex, France

(Received 1 July 1998; accepted for publication 31 January 2000)

Among the methods generally used to solve a problem in the domain of acoustic radiation, the equivalent sources method offers an interesting alternative. It consists in replacing the vibrating surface with a distribution of acoustic sources placed inside the structure. The contribution of each source is determined in such a way that the acoustic field radiated by these sources verifies the same boundary conditions on the structure. The number of unknowns in the problem is no longer directly linked to the number of mesh points on the structure, as with boundary elements methods, but to the number of equivalent sources employed in the model. The equivalent source method is therefore of major interest if the acoustic radiation of the structure can be approximated with a sufficiently low number of sources. This paper proposes its application when the equivalent source is a sphere. In this case, the number of unknowns is equal to the number of modes. In contrast to the one-point multipole, the sphere has a surface surrounding a closed volume to express the boundary conditions. Although sphere/multipole equivalence has been demonstrated, the surface of the sphere allows normalization of the functions used, leading to stabilization of the system to be resolved. First, the main acoustic radiation characteristics of a sphere and of the linear system verified by the modal coefficients of an equivalent sphere are presented. The different parameters of the model are then studied: position and radius of the equivalent sphere, truncation of the series, and influence of the spatial sampling (mesh). In the same vein, a second approach is presented. It consists of making each point of the structure correspond to a point of the sphere, and the vibrating field at the surface of the sphere is deduced from that of the structure by simple geometric projection. Results can be obtained very quickly as no matrix inversion is required. The accuracy of the results depends on the distance between the sphere and the structure. Finally, an experimental validation that uses both methods is presented and shows interesting results when the structure is closed, and when its shape is not too far removed from a sphere. © 2000 Acoustical Society of America.

[S0001-4966(00)01805-1]

PACS numbers: 43.20.Tb, 43.40.Rj [DEC]

INTRODUCTION

Predicting the noise radiated by a vibrating structure from calculated or measured vibratory data is a common problem in the field of vibro-acoustics. It is of interest to numerous industrialists who, at the design stage, can predict, analyze and possibly correct the harmful sound emissions produced by a machine. The software packages currently available employ calculation techniques that require complex numerical calculations which are difficult to apply. Any method that tends to reduce computer system requirements is therefore of major interest.

A closed structure S surrounded by a homogeneous, compressible and infinite fluid is assumed. At each time t , the acoustic pressure radiated at any point M outside of S is noted $P(M, t)$. A time harmonic dependence in the form $P(M, t) = P(M)e^{j\omega t}$ is assumed. The structure is subject to internal harmonic excitation which, at any point M_S of its surface S , generates a normal vibratory velocity field V^S . This velocity field is assumed known for every excitation frequency. Standard linear acoustic hypotheses then allow the acoustic pressure radiated by the structure to be defined

as the solutions of an external Neumann problem described by the following three equations:

(i) the Helmholtz equation at any point M outside of the structure:

$$(\Delta + k^2)P(M) = 0, \quad (1a)$$

where k is the acoustic wave number associated with the fluid medium ($k = \omega/c_0$) and c_0 is the speed of sound;

(ii) the boundary conditions of the structure:

$$\frac{\partial P(M_S)}{\partial n_S} = -j\omega\rho_0 V^S(M_S) \quad \forall M_S \in S, \quad (1b)$$

where n_S is the outward normal of the structure at point M_S ;

(iii) the Sommerfeld radiation condition when M is infinitely far from S :

$$\lim_{r=\|OM\| \rightarrow +\infty} r \left(\frac{\partial P(M, f)}{\partial r} + jkP(M, f) \right) = 0. \quad (1c)$$

From a mathematical point of view, this is a well-formulated problem that requires a single solution for every frequency.

It can be resolved by an analytical method that consists in seeking the solution in the form of a series of functions. In our case, this means the wave functions associated to eigenvectors of the operator $\Delta + k^2$ that verify the Sommerfeld radiation condition. These functions constitute the acoustic radiation modes of certain ideal structures. Morse and Feshbach have listed the 11 coordinate systems that can three-dimensionally determine these inherent functions by applying the variable separation method.¹ In practice, emphasis is placed on coordinate systems that, when one of the coordinates is held constant, generate surfaces with a simple geometry that can be closed (spheres or spheroids) or open (infinite plate and cylinder of infinite length).

As it is known how to express in an analytical form the acoustic radiation modes associated to a given structure, the differential system is resolved using simplified numerical methods.

This is the case with finite element methods that transform the differential system into a linear system by medium propagation discretization.² Mainly used to compute the vibration of a source, software packages based on these methods are already available to process interior acoustic problems. Exterior problems require a more complex approach. To limit the domain to be discretized, it is necessary to impose conditions equivalent to the Sommerfeld radiation condition^{3,4} on a surface surrounding the vibrating structure. When this control surface is a sphere, series development can also be employed on the spherical wave functions to calculate the acoustic pressure radiated outside the sphere.⁵ An interesting variant consists in defining infinite elements beyond the vibrating surface with a decay model respecting the Sommerfeld condition.^{6,7} Although the matrices encountered are generally sparse and symmetrical, the three-dimensional mesh required makes the method very costly in terms of calculation time, particularly when the vibrating surface is large compared to the acoustic wavelength.

The three-dimensional mesh imposed by this finite element method would appear even more unsuited to our problem given that the propagation medium is assumed homogeneous. The Green's theorem indeed allows the problem to be shifted from its differential form (1) to an integral form known as the Helmholtz integral equation. The acoustic field radiated then depends only on the surface pressure and the velocity of the structure, and on choosing a propagation function that respects the Sommerfeld radiation condition (e.g., the Green's function for unbounded medium). This method uses finite boundary elements and resolves a squared linear system whose size is directly linked to the number of mesh points on the surface of the structure. Whatever the excitation frequency the structure is submitted to, the existence of a solution can be demonstrated. Indirect methods also form part of the finite boundary element methods. They consist in looking for a distribution of one-point simple sources on the structure of the surface, called a layer potential, whose amplitude is determined from the boundary conditions.

The well-documented problem of critical or irregular frequencies is characterized from finite boundary element methods.^{8,9} It is linked to the choice of Green's function,¹⁰

and is expressed by numerical problems at the eigenfrequencies of the internal, homogeneous, Dirichlet-type differential system associated to the problem. For the direct methods, various solutions have been proposed⁹⁻¹⁷ which employ an overdetermined linear system. For the indirect methods, a single solution for every frequency is obtained by combining the potentials of single and double layers.¹⁸ Several integrals are then singular, and must be calculated either with the main Cauchy value or with the Hadamard finite part. The numerical calculation is, of course, more complex.

The equivalent sources method offers an interesting alternative, and consists in replacing the vibrating surface with a distribution of acoustic sources placed inside the structure. The contribution of each source is determined so that the acoustic field radiated by these sources verifies the same boundary conditions on the structure.

Depending on the authors, this method is found under the name "superposition method" (Ref. 19), "auxiliary source method" (Ref. 20), "equivalent source method" (Ref. 21), or even more recently "source simulation method" (Ref. 22). They have been developed for single monopolar- or dipolar-type equivalent sources. In contrast to the boundary element methods, the integrals that must be calculated are no longer singular. Furthermore, the number of unknowns in the problem is no longer directly linked to the number of mesh points on the structure, but to the number of equivalent sources employed in the model. The equivalent source method is therefore of major interest if the acoustic radiation of the structure can be approximated with a sufficiently low number of sources. Hwang and Chang²³ have shown that when the monopoles and dipoles are distributed on a surface inside the structure, only one combination similar to that used in the indirect methods ensures the existence and the uniqueness of a solution for every solution. Here, the problem of critical frequencies is linked to the surface that supports the sources. It can also be prevented by placing the sources on a single surface where the critical frequencies can be known (a sphere for instance). How to determine an optimal configuration (number and position of sources) has not yet been resolved.^{24,25} Jeans and Mathews have demonstrated that in the simple case of a sphere and of a spheroid, a linear system to be solved becomes ill-conditioned when the sources are placed far from the vibrating surface.²⁶ In contrast, the integrals become singular again if the sources are too close to the surface. The use of position optimization techniques for each source leads to a prohibitive calculation time, the number of sources generally remaining too high.

The use of equivalent sources more complex than monopoles and dipoles allows the number of sources to be limited. In 1988, Cremer and Wang²⁷ proposed the use of acoustic radiation modes of ideal sources defined by analytical methods to approximate the radiation modes of a structure. Indeed, the numerical techniques presented primarily employ the Green's function for unbounded medium as the propagation function. Part of the structural geometry information is then contained in wave functions more complex than the Green's function in free space. A reduction in the number of unknowns in the problem can therefore be expected.

The multipole is a natural extension of the single sources already encountered. It is a one-point source that can be defined from Taylor's expansion of the Green's function in free space²⁸ or from spherical wave functions.²⁹ The equivalence of these two representations has been demonstrated.²⁹ In both cases, the multipole is defined by an infinite series of wave functions. It can therefore be considered as a series of sources placed at the same point. Each source is associated with a function of the series, and the coefficients of the series express the participation of each source. Heckl³⁰ and Ochmann³¹ applied the weighted residue method to determine the coefficients of the series minimizing the difference between the calculated quadratic velocity and that known on the surface of the structure. It has been shown that the matrix system formed does not present the problem of critical frequencies.⁹ In 1964, Williams *et al.*³² applied this method to model the acoustic radiation of a cylinder with a finite length. The convergence velocity of the series therefore clearly depends on the ratio of the radius to the length of the cylinder. The choice of spheroidal wave functions was proposed. The numerical calculation of these functions is, nevertheless, more complex than that of spherical wave functions and, to our knowledge, no work has been published that employs them in the domain of equivalent sources. When the geometry of the vibrating structure is far from that of a sphere, Ochmann prefers to use several multipolar sources.³³ The coefficients of each of the series constitute new unknowns in the problem. To determine the optimal position of the multipoles, the coordinates of their centers are also considered as unknowns. A nonlinear system is then resolved using the least squares method, applying the Levenberg-Marquardt method,³³ but the calculation times rise considerably.

Using equivalent sources more complex than monopoles and dipoles therefore constitutes an interesting alternative to the boundary finite element methods generally used in the domain of acoustic radiation. It has not yet been fully covered, and this paper proposes its application when the equivalent source is a sphere. In Sec. I, following an overview of the main acoustic radiation characteristics of a sphere, we describe the linear system verified by the modal coefficients of an equivalent sphere. The different parameters of the model are then studied in Sec. II: position and radius of the equivalent sphere, and influence of the spatial sampling (mesh).

In contrast to the one-point multipole, the sphere has a surface to express the boundary conditions. Although sphere/multipole equivalence has been demonstrated,²⁹ the surface of the sphere allows normalization of the functions used, leading to stabilization of the system to be resolved. The surface of the sphere was also used by Hickling and Marin³⁴ to determine the noisiest parts of a diesel engine. They replaced this engine with an equivalent sphere, identified each part of the engine by a part of the sphere, and compared the acoustic radiation of each of the spherical caps defined according to the different models of single vibrating fields (pulsing or oscillating sphere cap, random distribution proposed by Morse and Ingard²⁹). The differences in geometry between the sphere and the engine were considered negli-

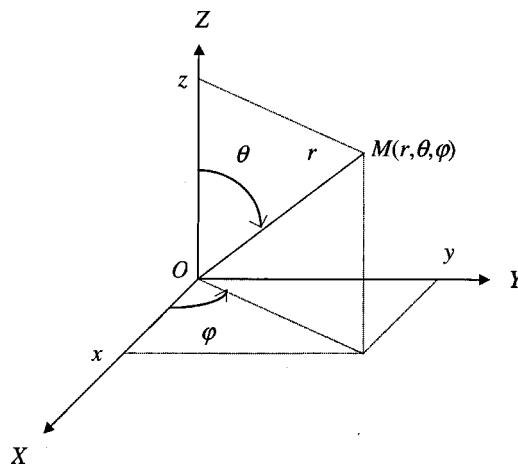


FIG. 1. Notations—spherical and cartesian coordinates.

gible at the low-frequency range considered, and the results are encouraging. In the same vein, in Sec. III we present a second approach which consists in making each point of the structure correspond to a point of the sphere, the vibrating field at the surface of the sphere being deduced from that of the structure by simple geometric projection. Results can be obtained very quickly as no matrix inversion is required. The accuracy of the results depends on the distance between the sphere and the structure. Finally, an experimental validation that uses both methods is presented in Sec. IV.

I. THEORY

A. Calculation of the acoustic radiation of a sphere

Let us consider the case of a vibrating sphere S_a of radius a centered at the origin O of a spherical coordinates system (O, r, θ, φ) (see Fig. 1). By applying the variable separation method, the acoustic pressure radiated by the sphere S_a is put into the form¹

$$P(M) = P(r, \theta, \varphi) = \sum_{n=0}^{\infty} \sum_{m=0}^n \sum_{\sigma=0}^1 V_{nm\sigma}^a Z_n(ka) F_{nm\sigma}^a(r, \theta, \varphi), \quad (2)$$

where the acoustic radiation mode of S_a associated with the triplet (n, m, σ) is

$$F_{nm\sigma}^a(r, \theta, \varphi) = \bar{Y}_{nm}^{\sigma}(\theta, \varphi) \frac{h_n(kr)}{h_n(ka)}. \quad (3a)$$

This depends on the normalized spherical harmonic function \bar{Y}_{nm}^{σ} defined by

$$\bar{Y}_{nm}^{\sigma}(\theta, \varphi) = \frac{1}{N_{nm}} \begin{cases} \cos(m\varphi) P_{nm}(\cos \theta) & \text{if } \sigma=1, \\ \sin(m\varphi) P_{nm}(\cos \theta) & \text{if } \sigma=0, \end{cases} \quad (3b)$$

where P_{nm} is the Legendre function (with $m \leq n$). A normalization coefficient N_{nm} is given by

$$N_{nm} = \sqrt{\frac{4\pi}{\epsilon_m(2n+1)} \frac{(n+m)!}{(n-m)!}}, \quad (3c)$$

where ϵ_m is the Neumann factor, which is 1 if m is equal to 0, otherwise 2.

The h_n function is the Hankel spherical function of order n defined by $h_n = j_n - jy_n$ where j_n and y_n are the Bessel and Neumann spherical functions of order n . The modal radiation impedance associated with mode $F_{nm\sigma}^a$ is given by

$$Z_n(ka) = -j\rho_0 c_0 \frac{h_n(ka)}{h_n'(ka)}; \quad (3d)$$

$V_{nm\sigma}^a$ is the complex modal coefficient defined from the vibration velocity V^a of sphere S_a by

$$V_{nm\sigma}^a = \int_{\theta=0}^{\pi} \int_{\varphi=0}^{2\pi} V^a(\theta, \varphi) \bar{Y}_{nm}^{\sigma}(\theta, \varphi) \sin \theta d\theta d\varphi. \quad (4)$$

The normalized spherical harmonic functions are a complete orthonormal set of basis functions on the surface of the sphere in which the vibration velocity is broken down into the form

$$V^a(\theta, \varphi) = \sum_{n=0}^{\infty} \sum_{m=0}^n \sum_{\sigma=0}^1 V_{nm\sigma}^a \bar{Y}_{nm}^{\sigma}(\theta, \varphi). \quad (5)$$

Each radiation mode is associated with a critical modal radiation frequency that depends only on the radius a of the sphere and the order n of the mode: $f_n = nc_0/2\pi a$. Modes of order n much greater than ka are therefore nonradiating modes of the sphere which, outside the near field of the sphere, contribute little to the acoustic radiation.³⁵ If the vibration velocity of the sphere does not allow significant nonradiating components, the series (2) can be truncated, retaining only the modes whose order n is less than or equal to the integer N_p defined by the expression

$$N_p(S_a, M) = [ka] + n_p(r), \quad (6)$$

where $[ka]$ is the closest integer to ka , and n_p is an integer that depends on the receiving point M ; the closer r is to a , the higher it is.

B. Application to the calculation of the acoustic radiation of a given structure

In the case of a given geometric structure, it is not known how to define a simple system of coordinates that allows the application of the variable separation method and the expression in analytical form of the radiation modes of the structure.

The principle of equivalent sources applied to the case of the sphere consists in employing the radiation modes of the sphere to calculate the acoustic radiation of the structure (Fig. 2). At any point M_S of the structure S the outward velocity field radiated by the sphere can be calculated by

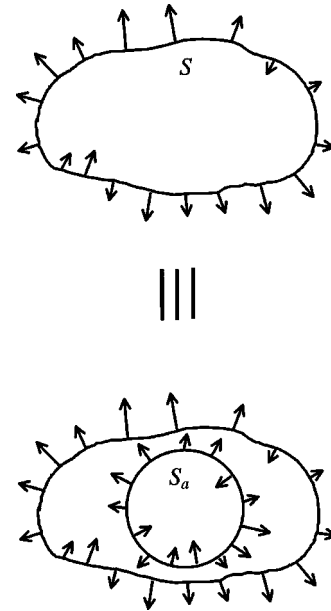


FIG. 2. Vibrating structure S and vibrating equivalent sphere S_a .

$$V^{\text{cal}}(M_S) = \frac{j}{k\rho_0 c_0} \overrightarrow{\text{grad}}(P(M_S)) \cdot \vec{n}. \quad (7)$$

Adopting a single-index notation, this is written

$$V^{\text{cal}}(M_S) = \sum_{j=1}^{\infty} x_j \frac{\partial \Psi_j(M_S)}{\partial n}, \quad (8)$$

where $x_j = V_{nm\sigma}^a$, and are the unknowns of the problem, and

$$\begin{aligned} \frac{\partial \Psi_j(M_S)}{\partial n} &= \bar{Y}_{nm}^{\sigma}(\theta, \varphi) \frac{h_n'(kr)}{h_n'(ka)} n_r \\ &+ \frac{\partial \bar{Y}_{nm}^{\sigma}(\theta, \varphi)}{kr \partial \theta} \frac{h_n(kr)}{h_n'(ka)} n_{\theta} \\ &+ \frac{\partial \bar{Y}_{nm}^{\sigma}(\theta, \varphi)}{kr \sin \theta \partial \varphi} \frac{h_n(kr)}{h_n'(ka)} n_{\varphi}. \end{aligned}$$

$\vec{n} = (n_r, n_{\theta}, n_{\varphi})$, and is the unit-length vector normal to S expressed in spherical coordinates. The indexes j and (n, m, σ) are linked by the relationship $j = n^2 + 2m + \sigma$ (see Table I).

If V_N^{cal} is limited to the N first terms of the series, then the acoustic radiation of the structure is correctly approached by that of the sphere, the difference between V_N^{cal} and V^S being minimum. The N first modal coefficients of the equivalent sphere that reduce the residual quadratic velocity to a minimum are therefore sought:

TABLE I. Links between j and (n, m, σ) .

j	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	...
n	0		1				2						3				4		...
m	0	0	1		0	1		2		0	1		3	2		3	0	1	...
σ	1	1	0	1	1	0	1	0	1	1	0	1	0	1	0	1	1	0	...

$$J_N(x) = \frac{1}{2} \int_S |\Delta V_N(M_S)|^2 dS, \quad (9)$$

where $\Delta V_N(M_S) = V^S(M_S) - V_N^{\text{cal}}(M_S)$. The modal coefficients therefore verify

$$\frac{\partial}{\partial x_j} [J_N(x)] = 0 \quad \text{for } 1 \leq j \leq N. \quad (10)$$

A simple mesh of the structure is therefore constructed where each of the N_S elements is defined by the coordinates of the center of the element, by the outward normal of the structure at this point, and by the elementary area associated to the element. Worthy of note is that if the vibrating mesh elements have the same elementary surface ΔS , the modal coefficients are independent of ΔS and directly minimize the difference between the velocity fields V^S and V_N^{cal} at the mesh points. It can be deduced that the x_j coefficients are solutions in the least-squares sense of the linear system:³⁶

$$A^* \cdot A \cdot x = A^* \cdot b, \quad (11)$$

$$(A^* A)_{ij} = \sum_{l=1}^{N_S} \frac{\partial \Psi_j}{\partial n}(M_l) \frac{\partial \Psi_i^*}{\partial n}(M_l) \quad \text{for } i \in [1, \dots, N] \quad \text{and } j \in [1, \dots, N], \quad (12)$$

$$(A^* b)_j = \sum_{l=1}^{N_S} V^S(M_l) \frac{\partial \Psi_j^*}{\partial n}(M_l) \quad \text{for } j \in [1, \dots, N], \quad (13)$$

where matrix A and vector b verify

$$Ax = b, \quad (14)$$

$$A_{ij} = \frac{\partial \Psi_j}{\partial n}(M_i) \quad \text{and} \quad b_i = V^S(M_i). \quad (15)$$

To apply the equivalent sphere method, different parameters must be determined, namely, the position and radius of the equivalent sphere, and the number of points of the mesh. These are studied in the following section.

II. NUMERICAL RESOLUTION AND ANALYSIS OF THE PARAMETERS

A. Numerical resolution method

Minimizing the residual quadratic velocity comes down to decomposing the vibration velocity of the structure into all the $\partial \Psi_j / \partial n$ functions. The linear system [(14) and (15)] will therefore be even more poorly conditioned as these functions are not orthonormalized on S ; in other words, as the form of S is further from a sphere. The system [(14) and (15)] is therefore resolved using the singular value decomposition method as this allows the unstable components of the measured velocity field of the structure to be filtered. This method is already being used in holography and, more generally, in all inverse problems. It consists in breaking down matrix A into the form

$$A = U^* W V, \quad (16)$$

where U is a column-orthogonal matrix of the same dimensions as A , V is a N^*N orthogonal matrix and W is a diagonal matrix whose diagonal elements are the singular values of matrix A . If matrix A^*A is nonsingular, then all the singular values of A are strictly positive, and the least-squares solution of the system is written

$$x = V W^{-1} U^* b = V \text{diag} \left(\frac{1}{w_1}, \dots, \frac{1}{w_N} \right) U^* b. \quad (17)$$

If U_i and V_i are denoted the i th vector-columns of matrices U and V , respectively, then they verify

$$A V_i = w_i U_i. \quad (18)$$

The SVD method therefore allows the explicit construction of a basis of orthonormal vectors (i.e., the columns of matrix U) from the radiation modes of the equivalent sphere (i.e., the columns of matrix A). The vector-columns of U therefore form a modal vibratory basis of the surface of S .³⁷ Furthermore, the vector-columns of matrix V define an orthonormal basis in which the modal coefficients of the equivalent sphere are decomposed. The singular value w_i is an amplification factor that characterizes the contribution of the different radiation modes of the equivalent sphere to the i th component of the modal vibratory basis of S . The contribution of this vibratory mode of the measured velocity field is therefore low if the amplitude of w_i is low (i.e., if it is much lower than the other singular values). In contrast, the vibratory modes associated with the high-magnitude singular values contribute significantly to the velocity field measured.

B. Radius and position of the equivalent sphere

The radius and position of the equivalent sphere are determined so as to obtain the most stable solution. This stability is measured by the condition number of matrix A , denoted $\mu(A)$ and defined by³⁶

$$\mu(A) = \frac{\max\{w_j, j=1, \dots, N\}}{\min\{w_j, j=1, \dots, N\}}. \quad (19)$$

It is therefore a real number greater than or equal to 1. The closer $\mu(A)$ is to 1 the better matrix A is conditioned. Two reasons of poor linear system conditioning can be distinguished. The first is linked to the nonorthogonal nature of the vector-columns of A . The $\partial \Psi_j / \partial n$ functions are indeed not orthogonal on nonspherical surfaces. Matrix A is therefore more poorly conditioned the further the structure is from a sphere. The second is associated to the norm difference between the vector-columns of A . Indeed, a column of A with an order of magnitude much lower than the other columns will be associated with a low-amplitude singular value. For the columns of A to be normalized, the $\partial \Psi_j / \partial n$ functions must be of the same order of magnitude on the surface of the structure. In addition, the acoustic pressure decreases with a/r for the radiating modes of S_a and with $(a/r)^{n+1}$ for the nonradiating modes.³⁸ The amplitude of functions $\partial \Psi_j / \partial n$ will therefore be of the same order of magnitude on the structure if the distance between the sphere and the structure is as low as possible. Figures 3 and 4 illustrate this behavior when the structure S is a sphere with a center O and radius R .

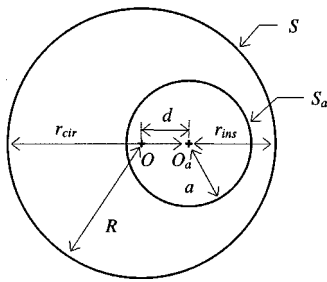


FIG. 3. Vibrating sphere and decentered equivalent sphere (r_{ins} =radius of the inscribed sphere, r_{cir} =radius of the circumscribed sphere).

A mesh of 900 points is defined on S . All the modes of the equivalent sphere whose order n is less than or equal to 5 are taken into account. The modes on sphere S can be expressed in terms of those on the equivalent sphere. For different positions of the equivalent sphere, the condition number $\mu(A)$ is represented as a function of the radius of the equivalent sphere. A minimum is obtained when the radius of the equivalent sphere is close to the radius of the sphere inscribed in the structure. This minimum is equal to 1 when both spheres have the same centers and increases the further the equivalent sphere is decentered.

More generally, the rapid decay in the amplitude of the nonradiating modes of the sphere with the distance between the two spheres explains the conditioning of the system (14). The compactness of the vibrating structure is defined as the ratio of the radius of the sphere surrounding the structure to the radius of the sphere surrounded by the structure. The position of the equivalent sphere that minimizes the condition number is therefore that giving a minimum compactness.

The equivalent sphere method has something in common with a “vibratory” holography method where the unknown velocity field on a “source” surface (the equivalent sphere S_a) is determined by retropropagation of the velocity field measured on a hologram (i.e., the vibrating surface of S).

Holography no longer conforms when the surface of S is

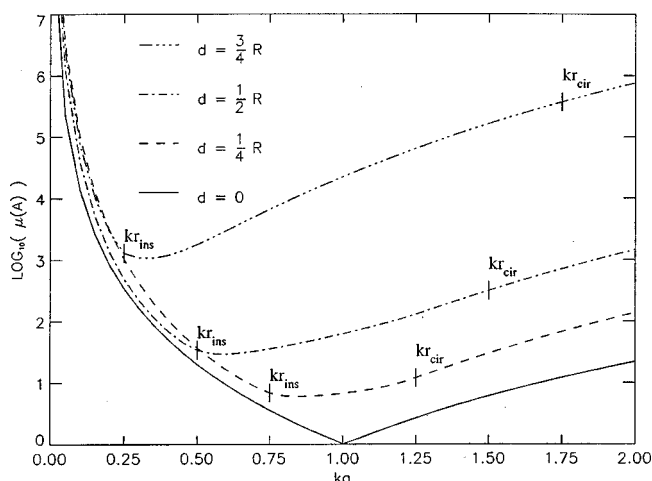


FIG. 4. Variation in the condition number of A as a function of the Helmholtz number for different positions of the equivalent sphere. The radii of the inscribed (r_{ins}) and circumscribed (r_{cir}) spheres have been shown for each position of the equivalent sphere.

not a sphere with the same center as the equivalent sphere. In reality, it is an inverse problem that worsens the conditioning the further the surface of the hologram is from the surface of the source. As in near-field holography, the amplitude of the nonradiating modes associated with the source surface can only be determined accurately if the hologram is close to this source; in other words, if the distance between the equivalent sphere and the structure is the shortest possible.

C. Spatial sampling of the vibratory field

The mesh of the structure must be sufficiently fine for the integrals C_{ij} and d_j to be calculated accurately. A criteria of a few mesh points between oscillations of the V^S and $\partial\Psi_j/\partial n$ functions can be considered sufficient. The variations in $\partial\Psi_j/\partial n$ are, however, difficult to predict. The number of oscillations of these functions increases rapidly with the index j . The spherical harmonic function \bar{Y}_{nm}^σ indeed allows $n+m$ nodal circles when $(\theta, \varphi) \in]0, \pi[\times]0, 2\pi[$ and its derivatives have similar properties. Functions $h'_n(kr_S)/h'_n(ka)$ and $h_n(kr_S)/h_n(ka)$ with large order n are associated with nonradiating modes of the equivalent sphere. Therefore, they have very different amplitudes at a point of the structure close to the sphere and one far from the sphere. This behavior increases the variations in amplitude of the $\partial\Psi_j/\partial n$ functions. The geometry of the structure is also taken into account by the outward normal at every point of the mesh. The high-radius curved surfaces are therefore characterized by rapid variations in components $(n_r, n_\theta, n_\varphi)$ of the normal. An accurate description of the $\partial\Psi_j/\partial n$ functions therefore requires a significant number of mesh points in the zones where the curve varies rapidly. If the structure has edges, the normal even becomes discontinuous. The vibrating surface is then no longer regular and the convergence of the method is no longer ensured from a theoretical point of view.

When the mesh of S is insufficiently dense, many $\partial\Psi_j/\partial n$ functions can no longer be distinguished. They then appear as linearly dependent, and provide an ill-conditioned A matrix. By refining the mesh, the orthogonal nature of $\partial\Psi_j/\partial n$ is better expressed, and the condition number of matrix A is reduced. The mesh is therefore sufficiently dense if it at least allows the contribution of all the $\partial\Psi_j/\partial n$ functions taken into account in the model to be distinguished. The condition number of A is an indicator of the spatial under-sampling of the structure.

The mesh of the surface of S can simply be interpreted as a sampling criteria of the $\partial\Psi_j/\partial n$ functions. Indeed, the relationship explicitly expresses the decomposition of the vibratory velocity V^S in all the $\partial\Psi_j/\partial n$ functions. These are therefore interpolation functions that allow reconstitution of the variations in V^S at any point of S from knowledge of V^S at a finite number of points of S . Between the mesh points, it is assumed that the calculated velocity allows for no significant variations. However, when higher-order $\partial\Psi_j/\partial n$ functions that allow rapid oscillations on S are under-sampled, the velocity radiated by the sphere is associated with a very low difference at the mesh points and significant oscillations between the mesh points (Fig. 5). The calculations of the

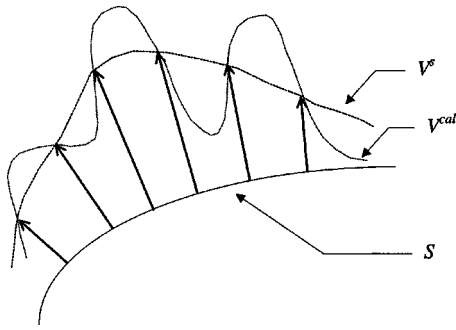


FIG. 5. Spatial undersampling of the $\partial\Psi_j/\partial n$ functions.

contribution of high n and/or m order modes depends greatly on the mesh and on the numerical integration method.

If a spatial over-sampling of the surface of S can be defined (i.e., a mesh containing two or three times as many points as the initial mesh), then the validity of the equivalent sphere model can be verified *a posteriori* by comparing the quadratic velocity calculated by the equivalent sphere from the initial mesh and the quadratic velocity calculated from the new mesh. If the initial mesh is not dense enough, the oscillations between the mesh points will be expressed by a difference in the two quadratic velocities.

D. Analysis of the quadratic error

The relative quadratic error is defined by

$$e_2 = \frac{\|b - Ax\|_2}{\|b\|_2}, \quad (20)$$

where $\|b\|_2$ is the Euclidian norm of vector b .

It can be considered that the equivalent sphere and the structure have similar boundary conditions when the quadratic error e_2 is of the same order of magnitude as the vibratory velocity measurement errors. Knowledge of this single criterion is, however, insufficient to guarantee that the acoustic radiation of the equivalent sphere is close to that of the structure. Indeed, when certain equivalent sphere modes are insufficiently sampled, the velocity field radiated by the sphere can be approximated correctly at the points of the mesh while allowing significant oscillations between mesh points. If these modes are radiating modes, the acoustic radiation of the equivalent sphere is then very different from that of the structure even though the quadratic error is close to zero.

The acoustic radiation of the structure can also be accurately approximated by the equivalent sphere without correct reconstruction of the vibratory deformation. Indeed, for the acoustic radiation to be correct, the modes of the sphere that contribute to the acoustic radiation of the structure simply have to converge rapidly. The quadratic error then characterizes the “nonradiating” components of the vibratory velocity that have not been reconstructed (i.e., those that vary more quickly than the acoustic wavelength). The quadratic error is, therefore, in general, not enough to qualify the quality of the model.

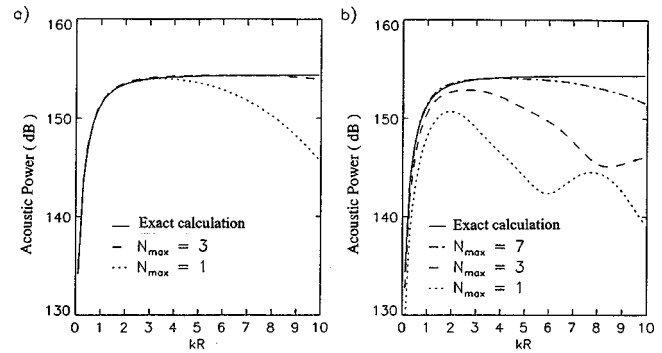


FIG. 6. Comparison of the sound power radiated by a pulsing sphere S with a radius R and two equivalent spheres: case (a) slightly decentered equivalent sphere S_a with center O_a ($a=0.626R$, $d=0.375R$), and case (b) highly decentered equivalent sphere S_a with center O_a ($a=0.25R$, $d=0.75R$). Vibratory mesh of the sphere=900 points, N_{\max} is the maximum order n of the equivalent sphere modes taken into account.

E. Numerical validations

1. Case of a sphere

First, the acoustic power radiated by a pulsing sphere is compared to the acoustic power radiated by a decentered equivalent sphere by applying the relationship

$$W = \frac{\rho_0 c_0}{2k^2} \sum_{n,m,\sigma} \left| \frac{V_{nm\sigma}^a}{h_n'(ka)} \right|^2. \quad (21)$$

Figure 6 compares the convergence of the calculation when the equivalent sphere is only slightly decentered (case a) or highly decentered (case b).

2. Case of a box

The integral method used to validate the results employed standard, commercially available software. It was applied to treat the simple case of a cube of dimensions L and center O [Fig. 7(a)] for which all the faces are rigid except face $X=L/2$ where

$$V(y,z) = V_0 \sin \frac{2\pi y}{L} \sin \frac{2\pi z}{L}$$

$$\text{for } (y,z) \in \left[-\frac{L}{2}, +\frac{L}{2} \right] \times \left[-\frac{L}{2}, +\frac{L}{2} \right],$$

$$\text{with } V_0 = 1.0 \text{ m/s}. \quad (22)$$

The regular mesh was 7 points by 7 points for each face of the cube. For a reduced frequency of $kL=10$, the directivity of the pressure field radiated at a distance $r=10L$ from the center of the cube is approximated correctly by the equivalent sphere method ($N_{\max}=10$) by halving the number of unknowns of the linear system [Fig. 7(b)].

Figure 8 shows that the results remain satisfactory with rapid convergence of the equivalent sphere method in another case of a vibrating box $(L_x, L_y, L_z) = (2L, L, L)$ in size. Here, all the plates vibrate, and the distance to the center of the cube was $r=10L$. The mesh of the cube was 264 points for the boundary element method and 1000 points for the equivalent sphere method (reduced frequency $kL=10$, $N_{\max}=10$, $V_0=1.0$ m/s).

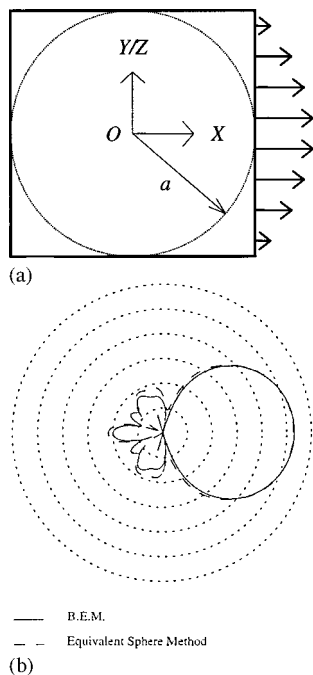


FIG. 7. (a) Cube of dimension L . (b) Directivity pattern, 10 dB between two consecutive circles. Directivity plane: $Z=0$.

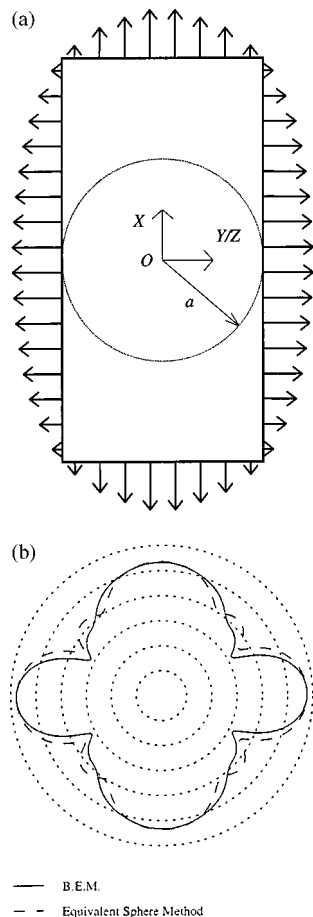


FIG. 8. (a) Box $(L_x, L_y, L_z) = (2L, L, L)$ in size. With mode (1,1), all faces vibrate in phase. (b) Directivity pattern, 10 dB between two consecutive circles. Directivity plane: $Z=0$.

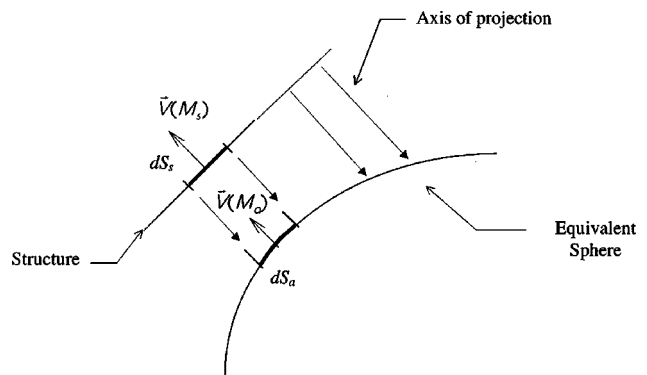


FIG. 9. Projection of the velocity field of the structure onto the sphere.

III. GEOMETRIC APPROACH

The concept developed in this section consists in directly determining the vibratory velocity field of the surface of the equivalent sphere by geometric projection of the velocity field of the structure, thus avoiding having to resolve an inverse problem. First, a mesh of the equivalent sphere is defined so that every point of the structure is associated with an image point on the sphere. Projection allows the vibratory velocity at each point of the equivalent sphere to be defined according to the vibratory velocity of its image on the structure. Two types of projection were studied, projection along the Oz axis (illustrated by Figs. 9 and 10) and projection along the axes which cross the center of the sphere. Second, the acoustic radiation of the equivalent sphere is determined by calculating the modal coefficients associated with the radiating modes of the sphere by means of Eq. (4), using the mesh of the equivalent sphere:

$$V_{nm\sigma}^a = \frac{1}{a^2} \int \int_{S_a} V^a(\theta, \varphi) \bar{Y}_{nm}^\sigma(\theta, \varphi) dS_a$$

$$= \frac{1}{a^2} \sum_{i=1}^{N_a} V^a(\theta_i, \varphi_i) \bar{Y}_{nm}^\sigma(\theta_i, \varphi_i) \Delta S_i. \quad (23)$$

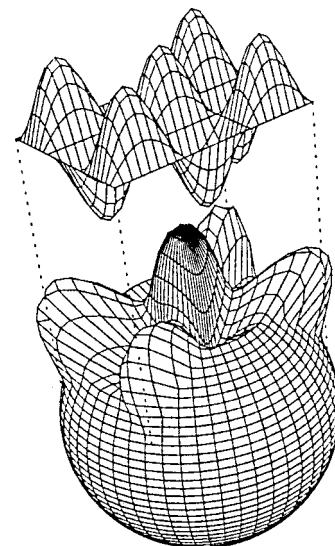


FIG. 10. Example of a geometric projection along a fixed axis.

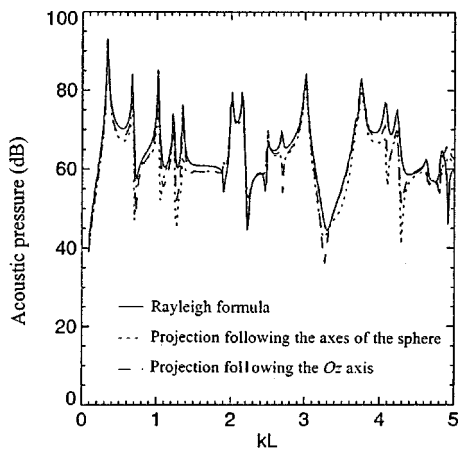


FIG. 11. Validation of the geometric approach. The Rayleigh formula is applied for a steel plate ($L_x; L_y$) = ($L; 0.7L$) in size and with a thickness of $e = L/40$. The mechanical characteristics are density: $\rho = 7800 \text{ kg/m}^3$, Young modulus: $E = 2.0 \times 10^{11} \text{ N/m}^2$, Poisson ratio: $\nu = 0.3$, and damping factor: $\eta = 0.01$. This plate, located in the plane $Z = 0$, is excited by a point force located at point $(x_0; y_0) = (-0.15L; 0.10L)$. The receiving point is near the plate $(x = 0.2L; y = 0.15L; z = 0.3L)$. The equivalent sphere (radius $a = 2L$) is tangent to the plate in its center.

The geometric approach was used to treat the case of a baffled plate (Fig. 11). The vibratory velocity is identified directly between the image points defined by the projection. Only one portion of the sphere is therefore associated with the vibrating plate. The velocity on the sphere is zero outside this portion and represents the presence of the baffle.

By increasing the radius of the equivalent sphere, the distance between the structure and the sphere can be kept as low as required for every point of the plate and of the baffle by choosing a sphere which cuts the plate tangentially at its center. The results obtained by this geometric approach were very close to those obtained by a standard analytical method. The errors remain negligible as long as the distance between the sphere and the structure is low compared to the acoustic wavelength, the characteristic length of the variations in the velocity field of the structure, and the directivity of the radiation modes of the sphere.

IV. EXPERIMENTAL VALIDATION

Both aforementioned methods were validated on a simple structure made up of a parallelepipedic box laid on the rigid reflecting floor of a semi-anechoic chamber (i.e., plane $\Pi: Z = 0$, Fig. 12). The $0.610 \times 0.410 \times 0.305 \text{ m}^3$ box

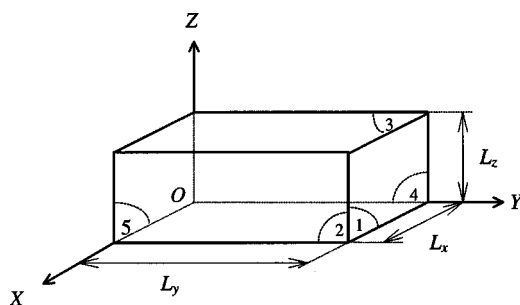


FIG. 12. The box and the number of sides.

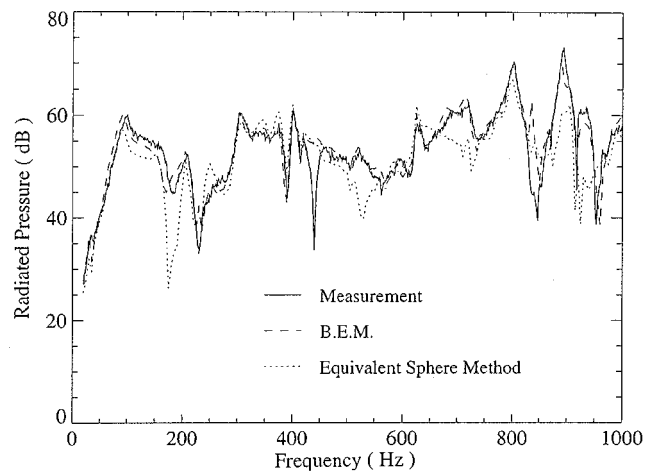


FIG. 13. Acoustic pressure radiated by the box at point $(x = -0.1 \text{ m}, y = 0.4 \text{ m}, z = 0.4 \text{ m})$ in dB (ref. $2 \times 10^{-5} \text{ Pa}$), calculated by several methods.

comprised five homogeneous sides with a constant thickness of 4 mm.

An electrodynamic shaker placed inside the box excites side 4 at coordinate point (0.0 m, 0.475 m, 0.10 m) with a wideband noise signal over the frequency range 20–1000 Hz. A regular mesh of 5 cm, 5 cm, and 4.5 cm was established on axes Ox , Oy , and Oz , respectively. This gave a 425-point measurement mesh (63 points on sides 1 and 5, 91 points on sides 2 and 4, and 117 points on side 3), which allowed a mesh of six points per flexural wavelength of the surface of the structure to be respected. The amplitude and phase of the vibratory velocity was measured at each of these points by a laser vibrometer mounted on a computer-controlled two-axis robot which allowed semi-automatic scanning of the different sides of the box. An impedance head inserted between the electrodynamic shaker and the structure was used to measure the reference signal. The autospectrum of the measured velocity yielded the vibratory velocity module, and the phase was obtained from the interspectrum phase between the measured velocity and the reference signal.

The image method is applied to take account of the reflecting ground (Π). S' is taken as the image surface of S with respect to plane Π , and Σ the fictitious closed surface defined by the merge of S and S' . The acoustic radiation of the box laid on the ground is identical to that of surface Σ when the velocity field on S' is symmetric to that of S with respect to plane Π .

The symmetry of Σ with respect to plane Π first leads to obtaining the minimum compactness of Σ when the center of the equivalent sphere belongs to plane Π . Second, if the axis Oz is chosen as normal to plane Π , the properties of symmetry verified by the $\partial \Psi_j / \partial n$ functions mean that some of these functions cannot contribute to the vibratory modal base,³⁹ and the number of unknowns can be reduced.

The equivalent sphere is chosen in the center of box Σ and is surrounded by the structure. Figure 13 compares the radiated acoustic pressure at a point close to the box obtained by measurement, by an integral method and by the equivalent sphere method. Resolving a system of 21 unknowns us-

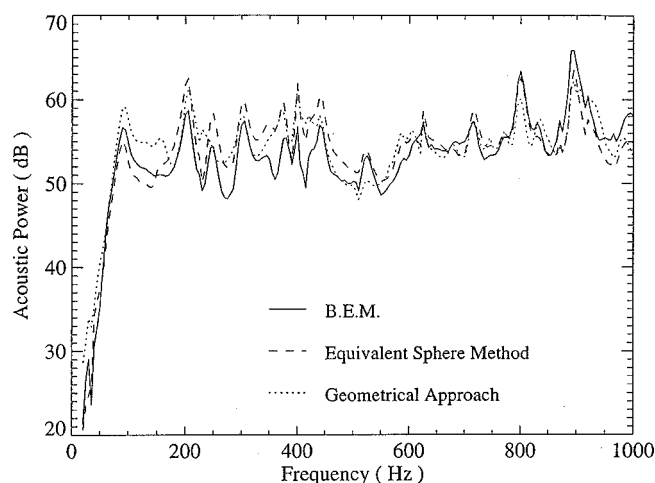


FIG. 14. Acoustic power radiated by the box in dB (ref. 10^{-12} W), calculated by several methods.

ing the equivalent sphere method (i.e., all the other modes n less than or equal to 5) gave similar results to the integral method with 425 unknowns. The discrepancies between the measurements and the BEM results are less than 3 dB. The discrepancies between the measurements and the equivalent sphere method results are higher and can reach 10 dB for certain frequencies. A clear trend is obtained in a very short calculation time.

Figure 14 provides a comparison of the acoustic power calculated by the different methods. In the geometric approach, the equivalent sphere has the same center as the box Σ , and the radius is 0.34 m. This is equal to the half-sum of the radii of the internal and external spheres. The discrepancies between the BEM and the equivalent sphere method results are less than 5 dB over the entire frequency range. The discrepancies between the results of the equivalent sphere method and the geometric approach are greater and can reach 10 dB, but the geometric approach gives a close approximation in a very short calculation time because it is based on a direct problem.

Figure 15 highlights a problem of insufficient sampling pointed out in Sec. II. The velocity field calculated with the first 50, 80, and 100 modes of the equivalent sphere on side 4 at 200 Hz is presented on a mesh containing three times more points than the measurement mesh. The oscillations that appear between the mesh points when the number of modes taken into account increases are linked to insufficient sampling of the high-order modes.

V. CONCLUSIONS

The principle of equivalent sources has been applied by replacing a vibrating structure with a sphere whose modal coefficients are calculated either by minimizing the residual quadratic velocity (equivalent sphere method itself) or by projecting the velocity field of the structure directly onto the sphere (geometric approach). In the first case, the radius and position of the sphere that provides the most stable model are obtained for the largest sphere surrounded by the structure. The number of modal coefficients to be determined to ap-

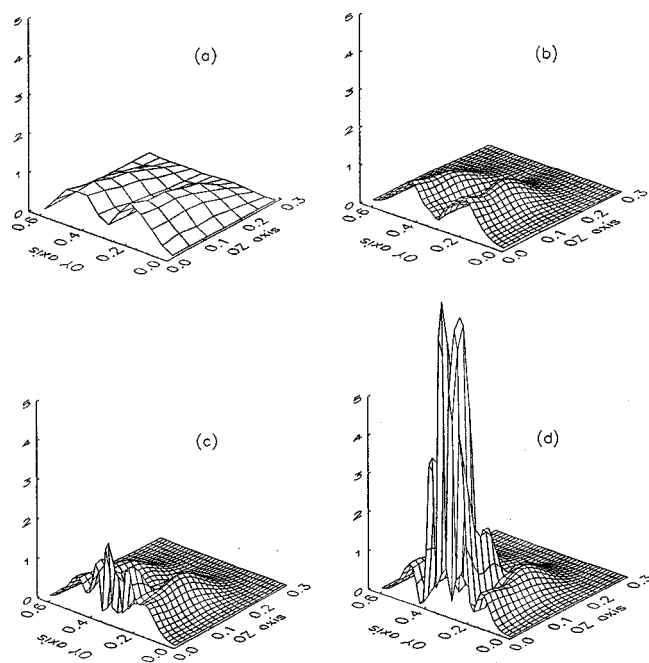


FIG. 15. Superimposition of the velocity modulus measured on side 4 (a) and generated by the (b) 50, (c) 80, and (d) 100 first modes of the equivalent sphere. The measurement mesh is (9×7) points for (a). The calculation mesh is (27×21) points for figures (b)–(d). The amplitude of the velocity field is normalized by the modulus of the highest measured velocity.

proximate the acoustic radiation of the structure is, generally, unknown, and depends on the geometry of the structure. In addition, it is limited by the density of the mesh. A simplified approach consists in projecting the velocity field of the structure onto the sphere. This provides quick results that are close to the exact solution when the distance between the structure and the sphere is low compared to the acoustic wavelength, the characteristic length of the field on the structure, and the distance between the directivity lobes of the radiating modes of the sphere.

For simple structures, comparing calculation with measurement has allowed validation of both approaches. The advantage of these approaches is the significant reduction in calculation time compared to boundary finite element methods. It should be borne in mind, however, that the expertise required to characterize the acoustic radiation of the equivalent sphere remains an important factor in employing these methods, and that other tools should be developed to determine the validity of the model.

¹ P. M. Morse and H. Feshbach, *Methods of Theoretical Physics* (McGraw-Hill, New York, 1953), Part II, p. 655.

² J. Assaad, J. N. Decarpigny, C. Bruneel, R. Bossut, and B. Hamonic, "Application of the finite element method to two-dimensional radiation problems," *J. Acoust. Soc. Am.* **94**, 562–573 (1993).

³ L. Ting and M. J. Miksis, "Exact boundary conditions for scattering problems," *J. Acoust. Soc. Am.* **80**, 1825–1827 (1986).

⁴ J. J. Shirron, "An iterative solution scheme for the exterior Neumann problem for $\Delta + \kappa^2$ using elements in bounded domain," *J. Acoust. Soc. Am.* **94**, 1121–1125 (1993).

⁵ J. T. Hunt *et al.*, "Finite element approach to acoustic radiation from elastic structures," *J. Acoust. Soc. Am.* **55**, 269–280 (1974).

⁶ D. S. Burnett, "A three dimensional acoustic infinite element based on a prolate spheroidal expansion," *J. Acoust. Soc. Am.* **96**, 2798–2816 (1994).

⁷ L. Cremers and K. R. Fyfe, "On the use of variable order infinite wave

- envelope elements for acoustic radiation and scattering," J. Acoust. Soc. Am. **97**, 2028–2040 (1995).
- ⁸ K. Brod, "On the uniqueness of solution for all wave numbers in acoustic radiation," J. Acoust. Soc. Am. **76**, 1238–1243 (1984).
- ⁹ H. A. Schenck, "Improved integral formulation for acoustic radiation problems," J. Acoust. Soc. Am. **44**, 41–58 (1967).
- ¹⁰ F. Ursell, "On the exterior problems of acoustics," Proc. Cambridge Philos. Soc. **74**, 117–125 (1973).
- ¹¹ L. G. Copley, "Fundamental results concerning integral representations in acoustic radiation," J. Acoust. Soc. Am. **44**, 28–32 (1968).
- ¹² A. J. Burton and G. F. Miller, "The application of integral equation methods to the numerical solution of some exterior boundary-value problems," Proc. R. Soc. London **323**, 201–210 (1971).
- ¹³ D. S. Jones, "Integrals equations for the exterior acoustic problem," Q. J. Mech. Appl. Math. **27** (Part 1), 129–142 (1974).
- ¹⁴ C. M. Piasczyk and J. M. Klosner, "Acoustic radiation from vibrating surfaces at characteristic frequencies," J. Acoust. Soc. Am. **75**, 363–375 (1984).
- ¹⁵ J. Ben Mariem and M. A. Hamdi, "A new boundary finite element method for fluid-structure interaction problems," Int. J. Numer. Methods Eng. **24**, 1251–1267 (1987).
- ¹⁶ B. Stupfel, A. Lavie, and J. N. Decarpigny, "Combined integral equation formulation and null-field method for the exterior acoustic problem," J. Acoust. Soc. Am. **83**, 927–941 (1988).
- ¹⁷ K. A. Cunefare, G. Koopman, and K. Brad, "A boundary element method for acoustic radiation valid for all wavenumbers," J. Acoust. Soc. Am. **85**, 39–48 (1989).
- ¹⁸ P. J. T. Filippi, "Layer potentials and acoustic diffraction," J. Sound Vib. **54**(4), 473–500 (1977).
- ¹⁹ G. H. Koopmann, L. Song, and J. B. Fahline, "A method for computing acoustic fields based on the principle of wave superposition," J. Acoust. Soc. Am. **86**, 2433–2438 (1989).
- ²⁰ Y. I. Bobrovnikskii and T. M. Tomilina, "Calculation of radiation from finite elastic bodies by the method of auxiliary sources," Sov. Phys. Acoust. **36**(4), 334–338 (1990).
- ²¹ T. Tomilina, "The equivalent sources approach to acoustical design of forced vibrating structures," Inter-Noise 93, Leuven (Belgium), 24–26 August 1993, pp. 1597–1600.
- ²² M. Ochmann, "The source simulation technique for acoustic radiation problems," Acustica **81**(6), 512–527 (1995).
- ²³ J. Y. Hwang and S. C. Chang, "A retracted boundary integral equation for exterior acoustic problem with unique solution for all wave numbers," J. Acoust. Soc. Am. **90**, 1167–1180 (1991).
- ²⁴ D. T. Wilton, I. C. Mathews, and R. A. Jeans, "A clarification of nonexistence problems with the superposition method," J. Acoust. Soc. Am. **94**, 1676–1680 (1993).
- ²⁵ W. Kropp and P. U. Svensson, "Time domain formulation of the method of equivalent sources," Acta Acust. (China) **3**, 67–73 (1995).
- ²⁶ R. Jeans and I. C. Mathews, "The wave superposition method as a robust technique for computing acoustic fields," J. Acoust. Soc. Am. **92**, 1156–1166 (1992).
- ²⁷ V. L. Cremer and M. Wang, "Die Synthese eines von einem beliebigen Körper in Luft erzeugten Feldes aus Kugelschallfeldern und deren Realisierung in Durchrechnung und Experiment," Acustica **65**(2), 53–74 (1988).
- ²⁸ A. J. Kempton, "The ambiguity of acoustic sources—a possibility for active control," J. Sound Vib. **48**(4), 475–483 (1976).
- ²⁹ P. M. Morse and K. U. Ingard, *Theoretical Acoustics* (McGraw–Hill, New York, 1968), p. 347.
- ³⁰ M. Heckl, "Bemerkung zur Berechnung der Schallabstrahlung nach der Methode der Kugelfeldsynthese," Acustica **68**, 251–257 (1989).
- ³¹ M. Ochmann, "Die Multipolstrahlersynthese—ein effektives Verfahren zur Berechnung der Schallabstrahlung von schwingenden Strukturen beliebiger Oberflächengestalt," Acustica **72**, 233–246 (1990).
- ³² W. Williams, N. G. Parme, D. A. Moran, and C. Sherman, "Acoustic radiation from a finite cylinder," J. Acoust. Soc. Am. **36**, 2316–2322 (1964).
- ³³ M. Ochmann, "Calculation of the sound radiation from complex structures using multipole radiator synthesis with optimised source locations," in 2nd International Congress on Recent Developments in Air- and Structure-borne Sound and Vibration, Auburn (USA), 1992, edited by M. J. Crocker, pp. 1187–1194.
- ³⁴ R. Hickling and S. P. Marin, "Enhancement of the sound power of a component of a complex noise source by sound from other nearby components," J. Acoust. Soc. Am. **84**, 262–274 (1988).
- ³⁵ M. C. Junger and D. Feit, *Sound, Structures and Their Interaction*, 2nd ed. (MIT, Cambridge, 1986).
- ³⁶ P. G. Ciarlet, *Introduction à l'analyse numérique matricielle et à l'optimisation* (Masson, Paris, 1982).
- ³⁷ G. V. Borgiotti, A. Sarkissian, E. G. Williams, and L. Schuetz, "Conformal generalized near-field acoustic holography for axisymmetric geometries," J. Acoust. Soc. Am. **88**, 199–209 (1990).
- ³⁸ X. Abramovitz and I. A. Stegun (eds.), *Handbook of Mathematical Functions With Formulas, Graphs, and Mathematical Tables*, 9th ed. (Dover, New York, 1970), ISBN 0-486-61272-4.
- ³⁹ L. Bouchet, "Calcul du rayonnement acoustique de structures à partir de données vibratoires par une méthode de sphère équivalente," Thèse de doctorat, INSA de Lyon, France, 1996.

Mass, momentum, and energy transfer by the propagation of acoustic solitary waves

N. Sugimoto

Department of Mechanical Science, Graduate School of Engineering Science, University of Osaka, Toyonaka, Osaka 560-8531, Japan

(Received 14 August 1998; revised 15 January 2000; accepted 5 February 2000)

This paper considers the mass, momentum, and energy transfer accompanied by the propagation of the acoustic solitary wave in a gas-filled tube. As was demonstrated previously [J. Acoust. Soc. Am. **99**, 1971–1976 (1996); Phys. Rev. Lett. **83**, 4053–4056 (1999)], the propagation of the solitary waves is made possible by connecting a periodic array of Helmholtz resonators axially with the tube. The solitary wave can convey the mass, momentum, and energy steadily with a constant speed that is subsonic but nearly equal to the linear sound speed. It is emphasized that the quantities transferred are of *first order* in magnitude. Formulating the basic equations in the conservation form, the total amount of the mass, momentum, and energy transferred is obtained by using the solitary-wave solutions. It has the upper bounds determined by the limiting solitary wave, which are proportional to the size of the resonator and the inverse of its natural frequency. In evaluating the energy transfer, a clear distinction should be made between the gas-dynamic energy and the acoustic energy. The former, which contains the first-order quantity whereas the latter begins with the quadratic ones, is to be used to determine the energy transfer in the form of heat (internal energy) associated with the mass transfer. © 2000 Acoustical Society of America.

[S0001-4966(00)02405-X]

PACS numbers: 43.25.Rq [MAB]

INTRODUCTION

It has recently been revealed, not only in theory but also in experiments, that acoustic solitary waves can propagate in a gas-filled and rigid tube by axially connecting with it a periodic array of Helmholtz resonators.^{1–3} The solitary wave is a compressive pulse localized both spatially and temporally and is propagated steadily with a constant speed slower than linear sound speed a_0 (i.e., subsonic), but faster than a threshold value $a_0/(1 + \kappa/2)$ where κ ($\ll 1$) is a small parameter to measure the size of the resonator. As the speed tends to the upper bound a_0 , the height of the solitary wave increases to approach the limiting value. The excess pressure in the tube corresponding to this height is given by $8\gamma\kappa/3(\gamma+1)$ relative to the pressure in equilibrium, γ being the ratio of specific heats. The solitary wave in this case is called the “limiting solitary wave.” On the other hand, as the speed tends to threshold, the height decreases and then the solitary wave approaches the Korteweg–de Vries soliton asymptotically.

It is a remarkable property that the solitary waves can be steadily propagated without any change of form. Although no dissipative effects are assumed, their propagation should be compared with that of nonlinear acoustic pulses in a tube without an array of resonators. It is usually the case that they evolve into shocks and decay quickly by nonlinearity even in the lossless limit. But when the array is connected, there arises a weak but pure dispersion,⁴ which can now compete with the nonlinearity to yield steady propagation. This persistent propagation implies a steady transfer of physical quantities. Because the solitary wave is a compressive pulse, an increase occurs in density and temperature of the gas, obeying the adiabatic relation. At the same time, the gas

moves in phase with the pressure pulse in the direction of propagation, although the speed of the gas itself is much smaller than the wave speed. Thus mass, momentum, and energy transfer can take place. The purpose of this paper is to examine how much of these quantities can be transferred by the propagation of the solitary wave.

In evaluating the energy transfer, one should distinguish the “gas-dynamic energy” as a fluid and the acoustic energy. The gas-dynamic energy consists of the internal (thermal) energy and the kinetic energy. It involves the first-order quantity, which does not vanish but remains when the solitary wave is integrated over the entire space or time. By contrast, the acoustic energy is defined as the sum of the potential energy stored by the excess pressure and the kinetic energy, and both are basically quadratic quantities. The conservation equation of the acoustic energy can be derived legitimately from the equation of the gas-dynamic energy by using the conservation equation of mass to eliminate the first-order terms. But use of the acoustic energy has a meaning only if a mean of the first-order quantity vanishes and then a quadratic relation, although small, is required. Since the integral of the first-order quantity does not vanish in the present case, it is the gas-dynamic energy that should be used in calculating the lowest energy transfer.

In what follows, the basic equations are recapitulated in Sec. I and then rewritten in the conservation form in Sec. II. The solitary-wave solutions are briefly summarized in Sec. III for use to evaluate the mass, momentum, and energy transfer. Section IV is devoted to the explicit calculations of the quantities transferred and to some discussion.

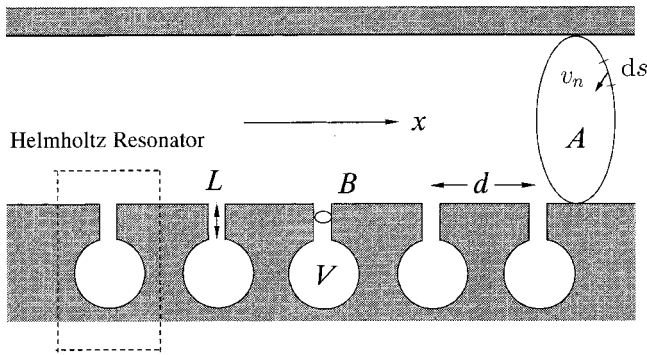


FIG. 1. A tube with a periodic array of Helmholtz resonators where A and B denote, respectively, the cross-sectional area of the tube and the throat, and d is the axial spacing between the neighboring resonators; V is the volume of the cavity, and L is the length of the throat, x being the axial coordinate along the tube. The integral in Eq. (1) is taken along the periphery of the cross-section of the tube, ds being a line element.

I. BASIC EQUATIONS

We start with reconsidering the basic equations used already in the previous paper, which will be referred to as Ref. 1. Figure 1 illustrates a tube to which identical Helmholtz resonators are connected in array with equal axial spacing. Since propagation of the solitary wave is a steady phenomenon, we ignore all dissipative effects due to viscosity and heat conduction. No account is taken of the boundary layer which will develop on the tube wall. The phenomenon is one-dimensional (plane) and uniform over the cross-section of the tube. Although the array of resonators gives rise to three-dimensional disturbances, the deviation from the plane wave remains small if the resonator is small in the sense that a cavity's volume V is small enough compared with a tube's volume per axial spacing d between neighboring resonators. This smallness is measured by a parameter κ defined by V/Ad , A being a cross-sectional area of the tube. When the inner surface of the tube is uniformly lined with many resonators, κ is taken as NV/A where N represents the number density of the resonators per unit axial length. In any case, we make the continuum approximation for the distribution of the resonators. It enables us to average their effects over axial length, on the basis of the assumption that the axial spacing is much smaller than a typical acoustic wavelength.

In formulating the problem, deviations from the plane wave are assumed to be so small that the equations averaged over the cross-section of the tube may effectively be used. We retain only terms of the first-order deviation and ignore all higher-order ones. Then the equation of continuity is given by

$$\frac{\partial \rho}{\partial t} + \frac{\partial}{\partial x}(\rho u) = \frac{1}{A} \oint \rho v_n ds, \quad (1)$$

where $\rho(x, t)$ and $u(x, t)$ represent, respectively, the mean density and axial velocity of the gas over the cross-section of the tube, which depend on the axial coordinate x and time t . The right-hand side represents the mass flux from the resonators to the tube or vice versa, where v_n denotes the velocity component normal inward to the inner surface of the tube and the integral is taken along the periphery of the cross-section at x (see Fig. 1). In the present context, v_n vanishes

on the inner surface except at an orifice where the tube opens to the resonator's throat. The quantity v_n gives rise to the small deviation from the plane wave. The right-hand side is small enough in comparison with the terms on the left-hand side and only the lowest term in v_n will be retained, whereas the nonlinear terms associated with variations in the tube are taken into account.

When the throat of the resonator is connected normal to the axis of the tube, the equation of motion in the axial direction is kept intact as the Euler equation as follows:

$$\frac{\partial u}{\partial t} + u \frac{\partial u}{\partial x} = -\frac{1}{\rho} \frac{\partial p}{\partial x}, \quad (2)$$

where $p(x, t)$ denotes the mean pressure in the tube. Since all dissipative effects are ignored and both tube and resonators are assumed to be thermally insulated from the outside, no heat is generated or conducted in the gas. Then the equation of energy is given in terms of the mean entropy $S(x, t)$ per unit mass as follows:

$$\frac{\partial S}{\partial t} + u \frac{\partial S}{\partial x} = 0. \quad (3)$$

The gas is assumed to obey the equation of state for an ideal gas, which is averaged over the cross-section as follows:

$$\frac{p}{p_0} = \frac{\rho T}{\rho_0 T_0}, \quad (4)$$

with $p_0 = R\rho_0 T_0$, where $T(x, t)$ denotes the mean temperature and R is a gas constant; the subscript 0 implies the value in equilibrium. Equation (3) states that the entropy of each gas particle is kept constant while in motion so that the gas is subjected to the adiabatic change. In this process, the pressure is expressed in terms of the density as follows:

$$\frac{p}{p_0} = \left(\frac{\rho}{\rho_0} \right)^\gamma, \quad (5)$$

with $\gamma = c_p/c_v$, where c_p and c_v are the specific heats at constant pressure and volume, respectively, and all the quantities in equilibrium are assumed constant everywhere. Equations (1) to (5) are the basic equations for the gas in the tube.

For the gas in the resonator, we consider equations separately in the cavity and in the throat. Assuming no motion of the gas in the cavity, use is only made of the equation for the conservation of mass. Balancing the rate of change of the mass in the cavity with the mass flux into it, the conservation of mass is given by

$$V \frac{\partial \rho_c}{\partial t} = Bq, \quad (6)$$

where $\rho_c(x, t)$ and $q(x, t)$ denote, respectively, the mean density in the cavity and the mean mass flux density averaged over the cross-section of the throat having the area B . The mass flux density q is assumed to be uniform along the throat because a typical acoustic wavelength is much longer than the length of the throat so that the gas in the throat may be regarded as being incompressible. Quantity q also represents the momentum density so the momentum balance of the gas in the throat is governed by

$$L \frac{\partial q}{\partial t} = -p_c + p, \quad (7)$$

where L is the length of the throat and $p_c(x, t)$ is the mean pressure in the cavity. These two equations may appear to be valid even in treating the nonlinear response of the resonator.¹ But we assume the response is linear because v_n is small enough and a typical length of the resonator is much smaller than the wavelength. Then the left-hand side of Eq. (6) may be approximated to be $(V/a_0^2) \partial p_c / \partial t$ to the lowest order in the adiabatic process. Using this and eliminating q from Eqs. (6) and (7), we obtain the linear equation for the resonator's response:

$$\frac{1}{\omega_0^2} \frac{\partial^2 p'_c}{\partial t^2} + p'_c = p', \quad (8)$$

where $\omega_0 (= \sqrt{Ba_0^2/LV})$ is the natural angular frequency of the resonator and $p' (= p - p_0)$ and $p'_c (= p_c - p_0)$ are the excess pressures; the prime is used henceforth to denote an excess quantity from the equilibrium. When the end corrections are made to the throat, L is lengthened by $0.82r$ on each end, r being the radius of the throat.¹

II. CONSERVATION FORM

We now rewrite the equations of continuity, motion, and energy presented in the preceding section into the conservation form.

A. Conservation of mass

Denoting the right-hand side of Eq. (1) by $-\rho\sigma$, it is given by

$$\frac{1}{A} \oint \rho v_n ds \equiv -\rho\sigma = -\frac{Bq}{Ad}. \quad (9)$$

Here the last term arises because, multiplying by d both the numerator and denominator on the first term, $d(ds)$ gives the area element on the inner surface of the tube where v_n vanishes except at the orifices. Using Eq. (6) and the parameter κ , Eq. (1) can be rewritten into the conservation form:

$$\frac{\partial}{\partial t}(\rho + \kappa\rho_c) + \frac{\partial}{\partial x}(\rho u) = 0. \quad (10)$$

This states that the conservation of the mass of the gas holds if the mass in the cavity is taken into account. Note that the temporal variation of the mass in the throat is negligible because the mass flux over the cross section of the throat is assumed to be uniform along its axis.

B. Conservation of momentum

Next we consider the conservation of the momentum in the axial direction of the tube. Multiplying Eq. (2) by ρ , and Eq. (1) by u , respectively, and adding them, it follows that

$$\frac{\partial}{\partial t}(\rho u) + \frac{\partial}{\partial x}(\rho u^2 + p) = -\rho\sigma u. \quad (11)$$

Since the gas in the resonator has no momentum in the x direction and the gas flown into (or out of) the resonator

loses (or gains) the axial momentum, the conservation of the momentum does not hold. In other words, the resonators play the role of the source or sink of the momentum for the gas in the tube, which appears on the right-hand side of Eq. (11). But if unidirectional propagation of weakly nonlinear waves is concerned just as in the case of the solitary waves, Eq. (11) can also be recast into the conservation form.

The right-hand side of Eq. (11) may be rewritten by using Eqs. (6) and (9) as

$$-\rho\sigma u = -\frac{V}{Ad} \frac{\partial p_c}{\partial t} u = -\frac{\kappa}{a_0^2} \frac{\partial p'_c}{\partial t} u. \quad (12)$$

Here u is related to p' by $u = p'/\rho_0 a_0$ to the lowest order (see Ref. 1, p. 64). Using this and Eq. (8),

$$-\rho\sigma u = -\kappa \frac{\partial m'_c}{\partial t}, \quad (13)$$

with m'_c defined by

$$m'_c = \frac{1}{2\rho_0 a_0^3} \left[\frac{1}{\omega_0^2} \left(\frac{\partial p'_c}{\partial t} \right)^2 + p'^2_c \right]. \quad (14)$$

As will be shown later, the momentum density m'_c is related to the energy density of the resonator. Thus Eq. (11) is expressed in the conservation form as

$$\frac{\partial}{\partial t}(\rho u + \kappa m'_c) + \frac{\partial}{\partial x}(\rho u^2 + p) = 0, \quad (15)$$

where $\rho u + \kappa m'_c$ is the momentum density, while $\rho u^2 + p$ is the momentum flux density. But it should be emphasized again that Eq. (15) holds only if the unidirectional propagation of weakly nonlinear waves is assumed and the source term is evaluated to the lowest order.

C. Conservation of energy

1. Equation for the total energy

In order to put Eq. (3) in the conservation form, we eliminate S by using the thermodynamic relation for the gas in motion:

$$T \frac{DS}{Dt} = \frac{DU}{Dt} + p \frac{D}{Dt} \left(\frac{1}{\rho} \right) = 0, \quad (16)$$

where D/Dt signifies the operator defined by $\partial/\partial t + u\partial/\partial x$, and $U(x, t)$ denotes the internal energy per unit mass of the gas. Using Eqs. (1) and (9), Eq. (16) is rewritten as

$$\frac{\partial U}{\partial t} + u \frac{\partial U}{\partial x} = -\frac{p}{\rho} \left(\frac{\partial u}{\partial x} + \sigma \right). \quad (17)$$

Multiplying this by ρ , and Eq. (1) by U , respectively, addition of both equations yields

$$\frac{\partial}{\partial t}(\rho U) + \frac{\partial}{\partial x}(\rho U u) = -p \frac{\partial u}{\partial x} - (\rho U + p)\sigma. \quad (18)$$

Besides the internal energy, the gas also has kinetic energy $u^2/2$ per unit mass. To write down the equation for the total energy E as the sum of the internal and kinetic energies, i.e., $E = U + u^2/2$, we multiply Eqs. (1) and (2) by $u^2/2$ and

ρu , respectively, to add both equations. Adding the resultant equation to Eq. (18), we have

$$\frac{\partial}{\partial t}(\rho E) + \frac{\partial}{\partial x}[(\rho E + p)u] = -(\rho E + p)\sigma. \quad (19)$$

Here it is to be noted that because the internal energy has an additive constant, we can specify only an excess quantity $U' (= U - U_0)$ from the equilibrium value U_0 arbitrarily chosen. But even if U' is used, the same form of Eq. (19) still holds, by virtue of Eq. (1), with E replaced by E' ($= U' + u^2/2$) as follows:

$$\frac{\partial}{\partial t}(\rho E') + \frac{\partial}{\partial x}[(\rho E' + p)u] = -(\rho E' + p)\sigma. \quad (20)$$

Next, we consider the energy equation for the gas in the resonator. Because there is no heat flux through the wall, the whole energy of the gas in the resonator consists, under the present assumptions, of the internal energy in the cavity and the kinetic energy in the throat. The rate of change of this energy should be balanced by the energy flown into the resonator per unit time through the orifice, plus the power done by the pressure force acting on the cross-section of the orifice. Since the velocity of the gas in the throat is assumed to be uniform over the cross-section, the conservation of the energy in the resonator is given by

$$\frac{\partial}{\partial t} \left(V \rho_c U_c + \frac{1}{2} B L \rho w^2 \right) = B(\rho E + p)w, \quad (21)$$

where $U_c(x, t)$ and $w(x, t)$ represent, respectively, the internal energy per unit mass of the gas in the cavity and the velocity of the gas in the throat. In this case as well, U_c and E may be replaced by their excess quantities U'_c and E' by virtue of Eq. (6) with $q = \rho w$ as

$$\frac{\partial}{\partial t} \left(V \rho_c U'_c + \frac{1}{2} B L \rho w^2 \right) = B(\rho E' + p)w. \quad (22)$$

Using $\sigma = Bw/Ad$ from Eq. (9), we substitute the right-hand side of Eq. (20) for that of Eq. (22). Then we arrive at the equation for the conservation of the total energy of the gas not only in the tube but also in the resonators:

$$\frac{\partial}{\partial t} \left[\rho E' + \kappa \left(\rho_c U'_c + \frac{1}{2} \frac{B L}{V} \rho w^2 \right) \right] + \frac{\partial}{\partial x}[(\rho E' + p)u] = 0. \quad (23)$$

Here the terms within the first square brackets are the energy density, while the terms within the second brackets are the energy flux density.

2. Equation for the acoustic energy

While Eq. (23) is the equation for the “gas-dynamic” energy, it may be transformed into that for the acoustic energy familiar in acoustics. To do so, the explicit expression of U' is evaluated. In the adiabatic process, the increment of U , dU , is due solely to the work done on the gas, i.e., $-pd(1/\rho)$. With the excess pressure $p' = p - p_0$, the excess internal energy is integrated to yield

$$U' = -\frac{p_0}{\rho} + \frac{p_0}{\rho_0} - \int_{\rho_0}^{\rho} p' d\left(\frac{1}{\rho}\right). \quad (24)$$

The first two terms correspond to the work done for volume change under constant pressure in equilibrium and the third term represents the potential energy stored by the excess pressure. The internal energy $\rho U'$ per unit volume is given by

$$\rho U' = \frac{p_0}{\rho_0} \rho' + \rho \int_{\rho_0}^{\rho} \frac{p'}{\rho^2} d\rho. \quad (25)$$

In order to evaluate U' , we notice the fact that U' for the ideal gas is determined by the temperature only as $U' = c_v T'$ with $T' = T - T_0$. When Eq. (4) is employed, U' is immediately obtained by using $c_v = R/(\gamma - 1)$ as

$$U' = \frac{1}{\gamma - 1} \left(\frac{p}{\rho} - \frac{p_0}{\rho_0} \right). \quad (26)$$

Equating the right hand sides of Eqs. (24) and (26), the integral can easily be obtained as follows:

$$\rho \int_{\rho_0}^{\rho} \frac{p'}{\rho^2} d\rho = p_0 + \frac{p_0}{\gamma - 1} \left(\frac{p}{p_0} - \frac{\gamma \rho}{\rho_0} \right). \quad (27)$$

When p'/p_0 is small, Eq. (27) may be expanded by using the adiabatic relation (5) as follows:

$$\rho \int_{\rho_0}^{\rho} \frac{p'}{\rho^2} d\rho = \frac{1}{2\rho_0 a_0^2} p'^2 - \frac{(2\gamma - 1)}{6\rho_0^2 a_0^4} p'^3 + \dots, \quad (28)$$

with $a_0^2 = \gamma p_0/\rho_0 = \gamma R T_0$. In a similar fashion, $\rho_c U'_c$ is given as the expression (25) and the corresponding integral is expanded in the same form as the expression (28) with ρ and p' replaced simply by ρ_c and p'_c , respectively.

Let us now go back to Eq. (23). We use the relation (25) and rearrange $\rho E'$ into the sum

$$\rho E' = \frac{p_0}{\rho_0} \rho' + e', \quad (29)$$

where e' is defined by

$$e' = \rho \int_{\rho_0}^{\rho} \frac{p'}{\rho^2} d\rho + \frac{1}{2} \rho u^2. \quad (30)$$

Here e' starts with quadratic terms. Equation (23) can be expanded around equilibrium, on substitution of the relation (29); however, there arise the linear terms $p_0 \rho'/\rho_0$ from $\rho E'$, and $p_0 \rho'_c/\rho_0$ from $\rho_c U'_c$, while there arises $p_0 \rho u/\rho_0$ from $(\rho E' + p)u$. But these terms drop out automatically by virtue of Eq. (10). As for the contribution from the resonators, we have only to take account of the quadratic terms because they behave linearly. Thus Eq. (23) is reduced to the following form:

$$\frac{\partial}{\partial t}(e' + \kappa e'_c) + \frac{\partial j'}{\partial x} = 0, \quad (31)$$

with

$$e'_c = \frac{1}{2\rho_0 a_0^2} p_c'^2 + \frac{1}{2} \frac{BL}{V} \rho_0 w^2, \quad (32)$$

and

$$j' = (e' + p')u. \quad (33)$$

This is the acoustic energy equation which holds among quadratic as well as higher-order quantities. If only the quadratic terms are concerned, it is nothing but the well-known equation in the linear acoustics. According to the convention, e' (and also e'_c) may be called the acoustic energy density, while j' may be called the acoustic energy flux (or simply acoustic intensity in the linear case). When w in e'_c is replaced with $(V/B\rho_0 a_0^2)\partial p'_c/\partial t$ by Eq. (6), we find that e'_c is related to m'_c through $e'_c = a_0 m'_c$. It should be remarked that the acoustic energy density and the acoustic energy flux differ from the gas-dynamic ones defined in Eq. (23). The gas-dynamic quantities $\rho E'$ and $\rho_c U'_c$ contain the first-order terms $p_0 \rho'/\rho_0$ and $p_0 \rho'_c/\rho_0$. As usually is the case with harmonic waves in linear acoustics, the time average of a quantity vanishes. Then, only the quadratic quantities have physical significance. This is the reason the acoustic energy is used. But when the time average of a quantity does not vanish, as in the case with solitary waves, we should employ the gas-dynamic quantities given in Eq. (23) to discuss thermal energy. Yet the acoustic energy density and flux give the quadratic quantity to be conserved.

III. SOLITARY-WAVE SOLUTIONS

In this section, we present the explicit form of the solitary-wave solutions obtained in Ref. 2. Solitary waves exist as steady progressive-wave solutions to the following system of equations for $f(\theta, X)$ and $g(\theta, X)$:

$$\frac{\partial f}{\partial X} - f \frac{\partial f}{\partial \theta} = -K \frac{\partial g}{\partial \theta}, \quad (34)$$

$$\frac{\partial^2 g}{\partial \theta^2} + \Omega g = \Omega f. \quad (35)$$

With the small parameter ϵ ($0 < \epsilon \ll 1$) specifying the weakness of the nonlinearity, ϵf and ϵg [$f \sim g \sim O(1)$] represent the excess pressures $[(\gamma+1)/2\gamma]p'/p_0$ and $[(\gamma+1)/2\gamma]p'_c/p_0$ in the tube and in the cavity, respectively. The independent variables X and θ are, respectively, the far-field coordinate $\epsilon\omega x/a_0$ associated with the nonlinearity, and the retarded time $\omega(t-x/a_0)$ in a frame moving with the sound speed a_0 , where ω is a typical angular frequency. The parameters K and Ω designate the ratio of the smallness of the resonator, and the ratio of the resonator's natural angular frequency to the typical frequency. They are defined as follows:

$$K = \frac{\kappa}{2\epsilon} \quad \text{and} \quad \Omega = \left(\frac{\omega_0}{\omega}\right)^2. \quad (36)$$

In Ref. 2, the solutions have been obtained after effecting the replacement (4) (Ref. 2) to remove the parameters K and Ω from Eqs. (34) and (35). To avoid confusion in this

replacement, we use an overbar to indicate f and g in the case with $K = \Omega = 1$:

$$(f, g) = (K\bar{f}, K\bar{g}) \quad \text{and} \quad (X, \theta) = (\bar{X}/K\sqrt{\Omega}, \bar{\theta}/\sqrt{\Omega}). \quad (37)$$

The solitary-wave solutions are obtained by looking for the solutions to Eqs. (34) and (35) in the form of $\bar{f} = \bar{f}(\bar{\zeta})$ and $\bar{g} = \bar{g}(\bar{\zeta})$ where $\bar{\zeta} = \bar{\theta} - s\bar{X}$, s being a parameter. The variable $\bar{\zeta}$ is related to x and t through

$$\bar{\zeta} = \omega_0 \left[t - \frac{x}{a_0} \left(1 + \frac{\kappa s}{2} \right) \right]. \quad (38)$$

From this, the propagation speed of the solitary wave is given by $a_0(1 + \kappa s/2)^{-1} = a_0(1 - \kappa s/2 + O(\kappa^2))$. It was demonstrated that the solitary-wave solution can exist only if $0 < s < 1$. This can be expressed as follows:

$$4 \tan^{-1} \sqrt{\frac{f_+ - \bar{f}}{\bar{f} - f_-}} - \frac{2s}{\sqrt{-f_+ f_-}} \times \log \frac{[\sqrt{-f_- (f_+ - \bar{f})} - \sqrt{f_+ (\bar{f} - f_-)}]^2}{(f_+ - f_-)\bar{f}} = \pm \bar{\zeta}, \quad (39)$$

where f_{\pm} are given by

$$f_{\pm} = -2 \left(s - \frac{2}{3} \right) \pm \sqrt{-\frac{4}{3}s + \frac{16}{9}}, \quad (40)$$

with the sign \pm vertically ordered. For later use, we note that the solution (39) is obtained by integrating the following differential equation

$$\left(\frac{d\bar{f}}{d\bar{\zeta}} \right)^2 = \frac{\bar{f}^2(f_+ - \bar{f})(\bar{f} - f_-)}{4(\bar{f} + s)^2}, \quad (41)$$

together with such boundary conditions as the undisturbed state far ahead of propagation, $\bar{f} \rightarrow 0$ as $\bar{\zeta} \rightarrow -\infty$ (i.e., $x \rightarrow \infty$), and the peak at the origin, $d\bar{f}/d\bar{\zeta} = 0$ at $\bar{\zeta} = 0$. When f is solved, then g is available through the algebraic relation

$$\bar{g} = \frac{1}{2}\bar{f}^2 + s\bar{f}. \quad (42)$$

Although the solution (39) appears to have a very complicated form, \bar{f} graphically takes a simple waveform of a pulse symmetric with respect to the peak at $\bar{\zeta} = 0$ and decayed out exponentially as $|\bar{\zeta}|$ increases. The explicit waveforms of \bar{f} and \bar{g} are displayed for some values of s (see Fig. 3 of Ref. 2). The heights in \bar{f} and \bar{g} given, respectively, by f_+ and $g_+ (= f_+^2/2 + s f_+)$ increase monotonously as s decreases in the range $0 < s < 1$.

Here we refer to the limiting behavior of the solitary waves as $s \rightarrow 0$ and $s \rightarrow 1$. As s tends to vanish, \bar{f} and \bar{g} approach, respectively,

$$\bar{f} = \frac{8}{3} \cos^2 \left(\frac{\bar{\zeta}}{4} \right) \quad \text{and} \quad \bar{g} = \frac{32}{9} \cos^4 \left(\frac{\bar{\zeta}}{4} \right), \quad (43)$$

for $-\pi \leq \bar{\zeta} \leq \pi$ and $\bar{f} = \bar{g} = 0$ for $|\bar{\zeta}| > \pi$. Because these solutions lose the regularity at $|\bar{\zeta}| = \pi$, they cannot be achieved and must be regarded as the limiting solutions as $s \rightarrow 0$. Thus these solutions are called the limiting solitary

wave. On the other hand, as $s \rightarrow 1$, \bar{f} and \bar{g} approach the soliton solution of the Korteweg-de Vries equation asymptotically. They are given by

$$\bar{f} = \bar{g} = \alpha \operatorname{sech}^2 \sqrt{\frac{\alpha}{12}} \bar{\xi}, \quad (44)$$

with $s = 1 - \alpha/3$ ($0 < \alpha \ll 1$) and $\bar{\xi} = \bar{\theta} - \bar{X} + \alpha \bar{X}/3$. It follows from the relation (42) that \bar{g} tends to \bar{f} as $s \rightarrow 1$ (i.e., $\alpha \rightarrow 0$).

Finally, we make the following points. Parameter ϵ has been introduced to designate the order of the excess pressures in the tube and in the cavity, referenced to p_0 , and used as the formal expansion parameter to derive the system of Eqs. (34) and (35). It is assumed small but left unspecified so far. By the definition (36) and the replacement (37), ϵf and ϵg turn out to be $\kappa \bar{f}/2$ and $\kappa \bar{g}/2$, respectively, where \bar{f} and \bar{g} are regarded as being of order unity. This suggests that the order of ϵ for the solitary waves is stipulated by the smallness of κ . In fact, the peak excess pressures $\Delta p'$ and $\Delta p'_c$ in the tube and in the cavity are given, respectively, by $\Delta p'/p_0 = \kappa \gamma f_+ / (\gamma + 1)$ and $\Delta p'_c/p_0 = \kappa \gamma g_+ / (\gamma + 1)$. Second, it is only within the lowest order of κ that the present theory can specify the excess pressures. In order to calculate them correctly up to the second order of κ , Eqs. (34) and (35) should be modified to include the higher-order terms of κ (or ϵ). This process would be straightforward but laborious.

Within the present lowest approximation, the other quantities are given in terms of \bar{f} and \bar{g} as follows:

$$\frac{p'}{p_0} = \frac{\gamma p'_c}{\rho_0} = \left(\frac{\gamma}{\gamma - 1} \right) \frac{T'}{T_0} = \frac{\gamma u}{a_0} = \left(\frac{\kappa \gamma}{\gamma + 1} \right) \bar{f}. \quad (45)$$

These lowest relations, except for the last but one, are simply the linear relations derived from the equation of state and the adiabatic relation. On the other hand, the quantities associated with the resonators are also calculated as follows:

$$\frac{p'_c}{p_0} = \frac{\gamma p'_c}{\rho_0} = \left(\frac{\kappa \gamma}{\gamma + 1} \right) \bar{g} \quad \text{and} \quad \frac{w}{a_0} = \frac{\kappa}{\gamma + 1} \left(\frac{a_0}{L \omega_0} \right) \frac{d\bar{g}}{d\bar{\xi}}. \quad (46)$$

Here the second relation is derived from the linearized version of Eq. (6): $V \partial \rho'_c / \partial t = B \rho_0 w$ with $\rho'_c = p'_c / a_0^2$, and the definition (38). The factor $a_0 / L \omega_0$ implies the ratio of a wavelength a_0 / ω_0 (divided by 2π) corresponding to the natural frequency of the resonator (not a wavelength of the acoustic waves concerned) to the throat's length L . It is assumed that this ratio is large but the small parameter κ makes $|w/a_0|$ much smaller than unity.

IV. CALCULATIONS OF MASS, MOMENTUM, AND ENERGY TRANSFER

With the explicit solutions in the preceding section, we now calculate the total amount of the mass, momentum, and energy transfer accompanied with the propagation of the solitary wave. Integrating Eqs. (10), (15), and (23) over the whole region of x ($-\infty < x < \infty$) and using the fact that all excess quantities vanish exponentially as $|x| \rightarrow \infty$, we have

$$\frac{d}{dt} \int_{-\infty}^{\infty} (\rho' + \kappa \rho'_c) dx = 0, \quad (47)$$

$$\frac{d}{dt} \int_{-\infty}^{\infty} (\rho u + \kappa m'_c) dx = 0, \quad (48)$$

and

$$\frac{d}{dt} \int_{-\infty}^{\infty} \left[\rho E' + \kappa \left(\rho_c U'_c + \frac{1}{2} \frac{BL}{V} \rho w^2 \right) \right] dx = 0. \quad (49)$$

Thus it is found that the integrals of the respective densities are conserved with respect to t .

In Eqs. (47) to (49), the terms associated with the resonators are smaller by the order of κ . To evaluate them correctly, ρ' , ρu , and $\rho E'$ should be specified up to the order of κ^2 inclusive. With the present approximation, however, we cannot evaluate the integrals up to this order and must be satisfied with the lowest-order transfer. This means the neglect of the terms with κ in the integrals.

By the lowest relations (45), the total excess mass is calculated by

$$\int_{-\infty}^{\infty} \rho' dx = \frac{\rho_0}{\gamma} \int_{-\infty}^{\infty} \frac{p'}{p_0} dx = \frac{\rho_0}{\gamma + 1} \int_{-\infty}^{\infty} \kappa \bar{f} d\bar{x}, \quad (50)$$

where t is held fixed. Changing the variable in the integral from x to $\bar{\xi}$, it follows that

$$\begin{aligned} \int_{-\infty}^{\infty} \kappa \bar{f} d\bar{x} &= \frac{\kappa a_0}{\omega_0 (1 + \kappa s/2)} \int_{-\infty}^{\infty} \bar{f} d\bar{\xi} \\ &= \frac{a_0}{\omega_0} \left[\kappa \int_{-\infty}^{\infty} \bar{f} d\bar{\xi} + O(\kappa^2) \right]. \end{aligned} \quad (51)$$

Introducing the definition,

$$\int_{-\infty}^{\infty} \bar{f} d\bar{\xi} \equiv I_1(s), \quad (52)$$

the total excess mass is written as

$$\int_{-\infty}^{\infty} \rho' dx = \frac{\kappa}{\gamma + 1} \left(\frac{\rho_0 a_0}{\omega_0} \right) I_1(s). \quad (53)$$

In a similar way, the total amount of the lowest momentum transfer is calculated by the following integral:

$$\int_{-\infty}^{\infty} \rho_0 u dx = \frac{\kappa}{\gamma + 1} \left(\frac{\rho_0 a_0^2}{\omega_0} \right) I_1(s). \quad (54)$$

It is found that since solitary waves are propagated with the speed close to the sound speed, the amount of the momentum transfer is given, to the lowest order, by that of the mass transfer times a_0 .

We next calculate the energy transfer in the form of the internal energy. The lowest term in the gas-dynamic energy density $\rho E'$ is given by $(p_0/\rho_0) \rho'$ in the expression (25). Because $(p_0/\rho_0) \rho' = p_0 T' / [(\gamma - 1) T_0] = c_v \rho_0 T'$ by the relations (45) and $R = c_v (\gamma - 1)$, the total amount of the lowest internal energy conveyed is given by

$$\int_{-\infty}^{\infty} \rho_0 c_v T' dx = \frac{\kappa}{\gamma + 1} \left(\frac{p_0 a_0}{\omega_0} \right) I_1(s). \quad (55)$$

Note that the quantities (53), (54), and (55) are given per unit cross-sectional area of the tube. Indeed they have been cal-

culated by integrating the respective densities at a time t over the entire region of x . Since the solitary waves are localized both spatially and temporally, we note that they can also be obtained by integrating the respective fluxes at a position x over t from $t = -\infty$ to $t = \infty$. For example, Eq. (53) is available by integrating the lowest mass flux density $\rho_0 u$ in Eq. (10) over t .

We calculate the acoustic energy of the solitary wave by integrating the energy density e' over x . Because $p' = \rho_0 a_0 u$ to the lowest order, the potential energy is equal to the kinetic energy and the total amount is given by

$$\int_{-\infty}^{\infty} e' dx = \int_{-\infty}^{\infty} \frac{p'^2}{\rho_0 a_0^2} dx. \quad (56)$$

To calculate the integral, we need the following integral of \bar{f}^2 defined by

$$\int_{-\infty}^{\infty} \bar{f}^2 d\bar{\xi} \equiv I_2(s). \quad (57)$$

Then the acoustic energy is given, to the lowest order, by

$$\int_{-\infty}^{\infty} e' dx = \frac{\kappa^2 \gamma}{(\gamma+1)^2} \left(\frac{\rho_0 a_0}{\omega_0} \right) I_2(s). \quad (58)$$

This is equal to the total amount of the acoustic energy flux which passes a location x .

Thus the problem is reduced to evaluation of the two integrals $I_1(s)$ and $I_2(s)$. With the solution (39), it appears to be very complicated to execute the integration. But we can achieve this easily if the differential equation (41) is used. Because \bar{f} is symmetric with respect to $\bar{\xi} = 0$, we have

$$I_1(s) = 2 \int_0^{\infty} \bar{f} d\bar{\xi}. \quad (59)$$

We then change the integral with respect to $\bar{\xi}$ to that with respect to \bar{f} by noting that $d\bar{f}/d\bar{\xi} < 0$ for $\bar{\xi} > 0$. In fact, it follows that

$$I_1(s) = 2 \int_{f_+}^0 \bar{f} \frac{d\bar{\xi}}{d\bar{f}} d\bar{f} = 4 \int_0^{f_+} \frac{(\bar{f} + s)}{\sqrt{(f_+ - \bar{f})(\bar{f} - f_-)}} d\bar{f}. \quad (60)$$

This integral can be executed analytically as

$$I_1(s) = 8\sqrt{s(1-s)} + \left(\frac{16}{3} - 4s \right) \cos^{-1} \frac{(-2+3s)}{\sqrt{4-3s}}, \quad (61)$$

where $\cos^{-1} z$ ($-1 \leq z \leq 1$) is defined to take the principal value between 0 and π . In a similar fashion, I_2 can also be integrated analytically as

$$I_2(s) = 16\sqrt{s(1-s)^3} + \frac{8}{3}(1-s)(4-3s) \cos^{-1} \frac{(-2+3s)}{\sqrt{4-3s}}. \quad (62)$$

Figure 2 shows the graph of $I_1(s)$ and $I_2(s)$ versus s where $0 < s < 1$. As s increases, both integrals decrease monotonically to vanish at $s = 1$. As s tends to vanish, $I_1(s)$ and $I_2(s)$ approach $16\pi/3$ and $32\pi/3$, respectively. The limiting values can also be calculated by using the solutions (43). As s tends to unity, on the other hand, $I_1(s)$ and $I_2(s)$ vanish as $12(1$

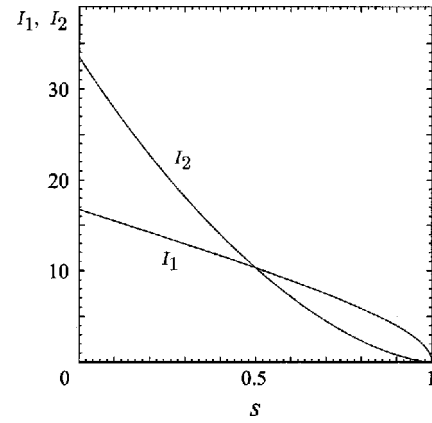


FIG. 2. Graph of $I_1(s)$ and $I_2(s)$ as a function of s ($0 < s < 1$) where as $s \rightarrow 0$, $I_1(s)$ and $I_2(s)$ tend to $16\pi/3$ and $32\pi/3$, respectively, while $s \rightarrow 1$, $I_1(s)$ and $I_2(s)$ vanish as $12(1-s)^{1/2}$ and $24(1-s)^{3/2}$, respectively.

$-s)^{1/2} + O(1-s)^{3/2}$ and $24(1-s)^{3/2} + O(1-s)^{5/2}$, respectively.

With $I_1(s)$ and $I_2(s)$ available, we now obtain explicitly the total amount of the mass, momentum, and energy transfer, respectively. Since $I_1(s)$ is bounded, the quantities (53) to (55) have the upper bounds, which are given, respectively, by

$$0 < \int_{-\infty}^{\infty} \rho' dx < \frac{16\pi\kappa}{3(\gamma+1)} \left(\frac{\rho_0 a_0}{\omega_0} \right), \quad (63)$$

$$0 < \int_{-\infty}^{\infty} \rho_0 u dx < \frac{16\pi\kappa}{3(\gamma+1)} \left(\frac{\rho_0 a_0^2}{\omega_0} \right), \quad (64)$$

and

$$0 < \int_{-\infty}^{\infty} \rho_0 c_v T' dx < \frac{16\pi\kappa}{3(\gamma+1)} \left(\frac{\rho_0 a_0}{\omega_0} \right). \quad (65)$$

Similarly, the acoustic energy is also bounded as follows:

$$0 < \int_{-\infty}^{\infty} e' dx < \frac{32\pi\kappa^2\gamma}{3(\gamma+1)^2} \left(\frac{\rho_0 a_0}{\omega_0} \right). \quad (66)$$

One solitary wave cannot convey mass, momentum, and energy greater than these limiting values. It is worth noting that the amount conveyed is increased as ω_0 becomes low.

Using these formulas, we calculate the upper bounds of the quantities transferred in the tube used for the experiments.³ For this tube, the parameter κ takes the value 0.197 and the natural frequency is 238 Hz with the end corrections. Then the upper bounds in the quantities (63) to (65) are given by the following values:

$$\frac{16\pi\kappa}{3(\gamma+1)} \left(\frac{\rho_0 a_0}{\omega_0} \right) = 3.8 \times 10^{-1} \text{ kg/m}^2, \quad (67)$$

$$\frac{16\pi\kappa}{3(\gamma+1)} \left(\frac{\rho_0 a_0^2}{\omega_0} \right) = 1.3 \times 10^2 \text{ N}\cdot\text{s/m}^2, \quad (68)$$

and

$$\frac{16\pi\kappa}{3(\gamma+1)} \left(\frac{\rho_0 a_0}{\omega_0} \right) = 3.1 \times 10 \text{ kJ/m}^2, \quad (69)$$

where $\rho_0 = 1.2 \text{ kg/m}^3$, $p_0 = 1.0 \times 10^5 \text{ Pa}$, $\gamma = 1.4$, and $a_0 = 340 \text{ m/s}$. On the other hand, the acoustic energy is smaller than that in the gas-dynamic one by $2\kappa\gamma/(\gamma+1)$. In the present case, the upper bound of the total amount of the acoustic energy is about 23% of that of the thermal energy and is given by 7.2 kJ/m^2 .

V. CONCLUSION

The mass, momentum, and energy transfer due to acoustic solitary waves propagating through a lossless and ideal gas in a tube with a periodic array of Helmholtz resonators have been examined. The transfer occurs steadily with a constant speed that is subsonic but nearly equal to the linear sound speed. The total amount of mass, momentum, and energy transfer has been calculated to the lowest order of the small parameter κ measuring the size of the resonator. It has upper bounds determined by the limiting solitary wave. It should be emphasized that the transfer takes place in the first-order quantity so the amount is proportional to κ . It may

also be worth noting that the total amount increases in proportion to the inverse of the natural angular frequency of the resonator ω_0 . Finally it is expected that the phenomenon of the steady transfer due to the solitary waves will be exploited in engineering devices such as heat pipes, heat engines, or heat pumps.

ACKNOWLEDGMENT

The author acknowledges the support by the Kajima Foundation, Tokyo, Japan.

¹N. Sugimoto, "Propagation of nonlinear acoustic waves in a tunnel with an array of Helmholtz resonators," *J. Fluid Mech.* **244**, 55–78 (1992).

²N. Sugimoto, "Acoustic solitary waves in a tunnel with an array of Helmholtz resonators," *J. Acoust. Soc. Am.* **99**, 1971–1976 (1996).

³N. Sugimoto, M. Masuda, J. Ohno, and D. Motoi, "Experimental demonstration of generation and propagation of acoustic solitary waves in an air-filled tube," *Phys. Rev. Lett.* **83**, 4053–4056 (1999).

⁴N. Sugimoto and T. Horioka, "Dispersion characteristics of sound waves in a tunnel with an array of Helmholtz resonators," *J. Acoust. Soc. Am.* **97**, 1446–1459 (1995).

Three-dimensional acoustic scattering from a penetrable layered cylindrical obstacle in a horizontally stratified ocean waveguide

Gerassimos A. Athanassoulis^{a)} and Aristides M. Prospathopoulos

*Department of Naval Architecture and Marine Engineering, National Technical University of Athens,
P.O. Box 64070, 15710 Zografos, Greece*

(Received 23 January 1998; revised 12 November 1999; accepted 10 December 1999)

In this work, a normal-mode solution is presented for the three-dimensional problem of acoustic scattering from a penetrable horizontally layered cylindrical obstacle in a shallow-water waveguide. The ocean environment around the obstacle is considered horizontally stratified and the bottom is assumed to be rigid. In the general case of depth-dependent sound-speed and density profiles (ssdp) the total acoustic field is calculated numerically, while an analytic solution is obtained in the special case of depth-independent ssdp. Numerical results concerning the transmission loss outside and inside single or double-layered cylindrical structures made of acoustic materials are given for a typical depth-dependent ocean environment. Comparisons with the cases of ideally soft and ideally hard cylindrical obstacles [J. Acoust. Soc. Am. **100**, 206–218 (1996)] are also made, illustrating the effect of acoustic properties of the obstacle. An important feature, which clearly emerges from the theoretical analysis and the numerical results, is the necessity of including evanescent modes in the calculations, in order to obtain physically meaningful and numerically accurate results. Furthermore, analytical expressions for the scattering cross section of a (penetrable or impenetrable) cylindrical obstacle are derived in terms of the expansion coefficients of the pressure field, and their behavior as frequency increases is numerically investigated. The solution presented in this paper, although addressing a special geometry, provides a means for handling strong discontinuities in both vertical and horizontal directions, and can serve as a benchmark solution to a problem for which no general numerical model exists, i.e., modeling acoustic scattering from a 3D obstacle in a 3D shallow-water waveguide. © 2000 Acoustical Society of America. [S0001-4966(00)01204-2]

PACS numbers: 43.30.Bp, 43.20.Fn, 43.20.Bi [SAC-B]

INTRODUCTION

A variety of methods has been proposed for the solution of three-dimensional (3D) waveguide propagation and scattering problems.^{1–3} However, efforts to benchmark the few existing 3D propagation models have been slowed by the sparsity of available analytic solutions providing a complete description of the field in all spatial coordinates.⁴ Especially, cases where backscattering is important (e.g., steep obstacles, seamounts, or islands) have been studied very little, and theoretical and computational analysis in these cases remains an area of active research. The presence of strong discontinuities of the acoustic parameters in the horizontal direction, as, e.g., a water–obstacle interface, renders most of the approximate techniques inapplicable. In these cases, methods based on the full-wave (Helmholtz) equation in conjunction with full-matching conditions seems to be the appropriate ones. Examples of this kind are the coupled-mode expansions, domain transformations and boundary-integral equation methods (BIEM). Relevant information can be found in the above mentioned Refs. 1–4 and Ref. 5. See also the more recent papers in Refs. 6–8.

In the present work, the 3D problem of acoustic scattering from a penetrable horizontally layered cylindrical obstacle in shallow water is studied by means of coupled-mode

expansions. A first attempt to treat the acoustic scattering from a penetrable cylindrical seamount using coupled modes has been undertaken by Evans,⁹ extending his stepwise coupled-mode method¹⁰ to three dimensions, without providing numerical results. Based on these works, Athanassoulis and Prospathopoulos¹¹ formulated a coupled-mode solution for the 3D problem of acoustic scattering from an impenetrable axisymmetric island in shallow water, also without giving numerical results. Working along the same lines, Taroudakis¹² presented first numerical results for a penetrable cylindrical seamount at a very low frequency ($f = 10$ Hz). There are, however, specific deficiencies in his approach making the accuracy of his results questionable and the extension of his approach to higher frequencies impossible. (See Sec. IV for a detailed discussion.) No reliable results had been provided until late 1997.¹³ In December 1997, Fawcett—who had earlier presented BIEM for 3D scattering in a 2D¹⁴ and an azimuthally symmetric¹⁵ ocean waveguide—presented a coupled mode method for computing the wave field scattered by 3D axisymmetric, fluid objects in a free space, a waveguide, or partially or fully buried in a basement, and gave results for the cases of a finite cylinder in free space and embedded between two half-spaces.¹⁶

The present work can be considered as a continuation of Refs. 5, 17, and 18, referring to the problem of 3D acoustic scattering of a source-generated acoustic field from an impenetrable cylindrical obstacle in a waveguide. In the present

^{a)}Electronic mail: mathan@central.ntua.gr

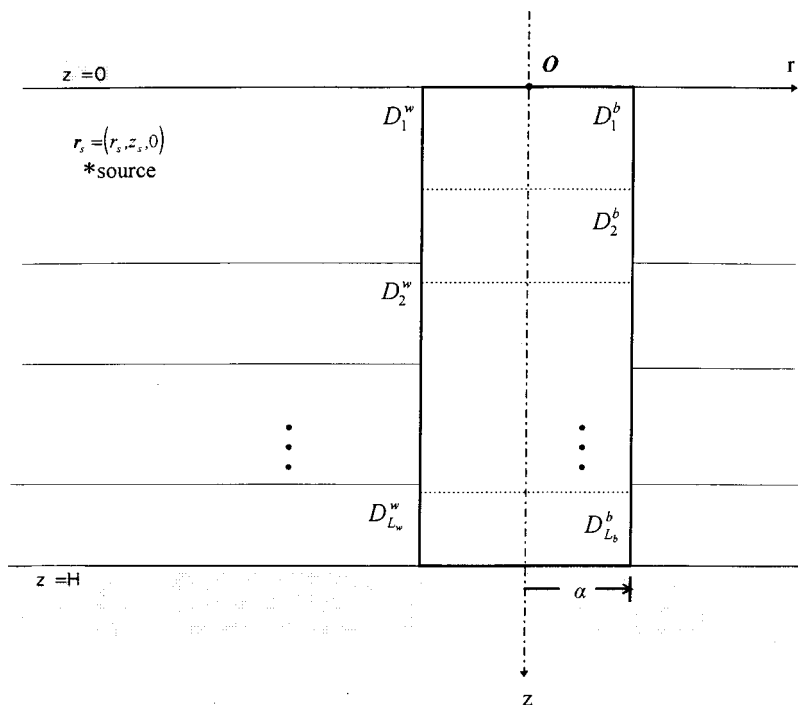


FIG. 1. The geometrical configuration of the studied environment.

paper, the case of a penetrable cylinder is considered, the emphasis being on obtaining reliable numerical results for an appreciable frequency range (ka up to about 400, where k is the wavenumber and a is the radius of the cylindrical inclusion). An important feature emerging from the study of this problem is the need to take into account a number of evanescent modes in order to obtain physically meaningful and numerically accurate results.

The main attribute of this work is that it provides a new benchmark solution to a 3D scattering problem that will be useful to model developers working on 3D and $N \times 2D$ underwater acoustic codes. Note that, although there is a number of underwater 3D propagation models available, no fully 3D, fully coupled scattering models exist, to the best of our knowledge.

We now summarize the organization of this article. In Sec. I the formulation of the problem is described. In Sec. II an infinite system of equations is derived for the unknown coefficients of the series expansions representing the pressure fields outside and inside the obstacle. In the case of a constant ssdp, the above system can be solved analytically, and the pressure field is given by exact, analytic formulas. Also, analytic formulas for the scattering cross section of an impenetrable or penetrable (layered) cylindrical obstacle are derived in terms of the expansion coefficients of the pressure field. In Sec. III numerical aspects concerning the truncation and convergence of the normal-mode series expansions are discussed. In the same section, results concerning the transmission loss outside and inside single- or double-layered cylindrical structures are presented, and comparisons with the ideally soft and ideally hard obstacle cases are made. Finally, the behavior of the scattering cross section of a cylindrical obstacle is numerically investigated versus frequency for various characteristic values of the environmental parameters. It is shown that, as frequency increases, the calculated scattering cross sections tend to the high-frequency limit ob-

tained by means of geometrical acoustics. The main features of the present work, conclusions, and discussion are given in Sec. IV.

I. FORMULATION OF THE PROBLEM

Let us consider the geometry of the environment illustrated in Fig. 1. A vertical cylinder of radius a , consisting of L_b horizontal acoustic layers, is surrounded by a horizontally stratified L_w -layered ocean of total constant depth H , which is confined between a rigid (perfectly reflecting) bottom and a pressure-release sea surface. The layers in the water and the cylinder may be of arbitrary width. The cylinder is extended from the bottom up to the sea surface, and its lateral (wetted) surface is acoustically penetrable. The geometric domains occupied by the water mass and the acoustic material in the cylinder are denoted by D^w and D^b , respectively. Somewhere in D^w a point source of unit strength is located, emitting monochromatic sound waves with angular frequency ω .

A cylindrical-polar coordinate system is introduced having as z axis the axis of the cylinder pointed downward, and origin at the intersection of z axis and the sea surface. The source point is denoted by $\mathbf{r}_s = (r_s, z_s, 0)$, and the generic field point is represented by $\mathbf{r} = (r, z, \theta)$. The acoustic properties of the water and the medium in the cylinder (pressure p , density ρ , and sound speed c) are denoted with a subscript w or b , respectively. The ssdps in both water and obstacle are assumed to be range independent, that is, $c_w = c_w(z)$, $\rho_w = \rho_w(z)$, and $c_b = c_b(z)$, $\rho_b = \rho_b(z)$.

The acoustic propagation/scattering problem in the above environment is modeled by the 3D Helmholtz equation, the waveguide boundary conditions (pressure release condition at the free surface and vanishing of the normal velocity at the rigid bottom), the matching conditions at the vertical and horizontal interfaces expressing the continuity of

pressure and normal velocity, and an appropriate (Sommerfeld) radiation condition expressing that the pressure field at infinity ($r \rightarrow \infty$) behaves like a system of cylindrical outgoing waves. The mathematical form of the aforementioned equations/conditions can be found in standard textbooks.¹ See also Ref. 5, where the same geometry is considered.

II. SERIES EXPANSION OF THE ACOUSTIC PRESSURE OUTSIDE AND INSIDE A CYLINDRICAL OBSTACLE

A. Derivation of an infinite system of equations

The problem of 3D acoustic scattering from an ideally soft or ideally hard cylindrical obstacle in a waveguide was treated in Ref. 5, by means of normal-mode series expansions [see Eqs. (22), (23) and (26), (27) of Ref. 5]. Subsequently, in Refs. 17 and 18, the series solution was improved by summing up explicitly the slowly convergent part. This analysis resulted in a decomposition into an incident field, p_i , and a scattered field, p_{sc}

$$p(r, z, \theta) = p_i + p_{sc} = \frac{i}{4\rho(z_s)} \sum_{n=1}^{\infty} g_n(z_s) H_0^{(1)}(\lambda_n R) g_n(z) - \frac{i}{4\rho(z_s)} \sum_{n=1}^{\infty} \sum_{m=0}^{\infty} e_m H_m^{(1)}(\lambda_n r_s) g_n(z_s) \times \Lambda_m(\lambda_n a) H_m^{(1)}(\lambda_n r) g_n(z) \cos m\theta, \quad (1)$$

where λ_n are the eigenvalues and $g_n(z)$ the normalized eigenfunctions of the vertical problem, $H_m^{(1)}(\cdot)$ are the Hankel functions of order m of the first kind, $e_0 = 1$, $e_m = 2$, $m > 0$, $R = (r_s^2 + r^2 - 2r_s r \cos \theta)^{1/2}$ is the distance between the source and the receiver, and the coefficients $\Lambda_m(\lambda_n a)$ are given by

$$\Lambda_m^{\text{soft}}(\lambda_n a) = \frac{J_m(\lambda_n a)}{H_m^{(1)}(\lambda_n a)}, \quad (2a)$$

$$\Lambda_m^{\text{hard}}(\lambda_n a) = \frac{J'_m(\lambda_n a)}{H_m^{(1)'}(\lambda_n a)}, \quad (2b)$$

where the prime denotes differentiation with respect to the argument. [See Eqs. (14)–(16) in Ref. 17 and Eqs. (10)–(12) in Ref. 18. Note that the factor $1/\rho(z_s)$ in Eq. (1) corresponds to a depth-dependent density profile.]

Following a similar analysis for the case of a penetrable cylindrical obstacle, it is found that the total pressure field outside the obstacle, p_w , can be represented as the sum of an axisymmetric incident field p_i [the same as in the previous case, see Eq. (1)], and a scattered field, p_w^{sc} , given by a double series, as follows:

$$p_w(r, z, \theta) = p_i + p_w^{\text{sc}} = \frac{i}{4\rho(z_s)} \sum_{n=1}^{\infty} g_n(z_s) H_0^{(1)}(\lambda_n R) g_n(z) + \sum_{n=1}^{\infty} \sum_{m=0}^{\infty} e_m A_{nm}^{\text{sc}} H_m^{(1)}(\lambda_n r_s) g_n(z_s) \times \tilde{H}_m^{(1)}(\lambda_n r) g_n(z) \cos m\theta, \quad (3)$$

where

$$\tilde{H}_m^{(1)}(\lambda_n r) = \frac{H_m^{(1)}(\lambda_n r)}{H_m^{(1)}(\lambda_n a)}, \quad (4)$$

and A_{nm}^{sc} are appropriate constants to be determined. A normalization of the form (4) is often used to avoid overflow problems in numerical computations (see, e.g., Ref. 10 and Ref. 1, § 5.9.1).

The pressure field inside the cylindrical obstacle, p_b , is represented by a double series of the form (see, for example, Chap. XXI of Ref. 19)

$$p_b(r, z, \theta) = \sum_{n=1}^{\infty} \sum_{m=0}^{\infty} e_m B_{nm} \tilde{J}_m(\mu_n r) \psi_n(z) \cos m\theta, \quad (5)$$

where μ_n and $\psi_n(z)$ are, respectively, the eigenvalues and the normalized eigenfunctions of the vertical problem corresponding to the stratification of the obstacle,

$$\tilde{J}_m(\mu_n r) = \frac{J_m(\mu_n r)}{J_m(\mu_n a)} \quad (6)$$

is a normalization analogous to (4), and B_{nm} are appropriate constants to be determined.

The representations (3) and (5) are valid independently from the type of matching (or boundary) condition to be imposed on the cylindrical surface. Applying the matching conditions at the cylindrical interface, linear systems of equations are obtained, permitting the calculation of the coefficients A_{nm}^{sc} and B_{nm} . Before proceeding to the construction of these systems, we recall here the orthonormality conditions for the vertical eigenfunctions in the water and in the obstacle

$$\int_0^H \frac{1}{\rho_w(z)} g_n(z) g_\nu(z) dz = \int_0^H \frac{1}{\rho_c(z)} \psi_n(z) \psi_\nu(z) dz = \begin{cases} 1, & n = \nu \\ 0, & n \neq \nu \end{cases}, \quad (7)$$

and the Graf's addition theorem for the Hankel functions²⁰

$$H_0^{(1)}(\lambda_n R) = \sum_{m=0}^{\infty} e_m H_m^{(1)}(\lambda_n r_s) J_m(\lambda_n r) \cos m\theta. \quad (8)$$

Then, exploiting the pressure continuity condition at $r = a$, and using (7), (8), and the orthonormality of the azimuthal eigenfunctions $\cos m\theta$, we obtain

$$A_{vm}^{\text{sc}} = \sum_{n=1}^{\infty} \frac{O_{vn}^w}{g_\nu(z_s) H_m^{(1)}(\lambda_\nu r_s)} B_{nm} - \frac{i}{4\rho(z_s)} J_m(\lambda_\nu a), \quad (9)$$

where

$$O_{vn}^w = \int_0^H \frac{1}{\rho_w(z)} \psi_n(z) g_\nu(z) dz. \quad (10)$$

Working similarly with the normal-velocity continuity condition at $r = a$, we obtain

$$A_{\nu m}^{\text{sc}} = \sum_{n=1}^{\infty} \frac{O_{\nu n}^b}{g_{\nu}(z_s) H_m^{(1)}(\lambda_{\nu} r_s)} \frac{\mu_n J'_m(\mu_n a)}{\lambda_{\nu} J_m(\mu_n a)} \frac{H_m^{(1)}(\lambda_{\nu} a)}{H_m^{(1)'}(\lambda_{\nu} a)} \times B_{nm} - \frac{i}{4\rho(z_s)} \frac{H_m^{(1)}(\lambda_{\nu} a)}{H_m^{(1)'}(\lambda_{\nu} a)} J'_m(\lambda_{\nu} a), \quad (11)$$

where

$$O_{\nu n}^b = \int_0^H \frac{1}{\rho_b(z)} \psi_n(z) g_{\nu}(z) dz. \quad (12)$$

The coupling coefficients $O_{\nu n}^{w,b}$ are weighted inner products between the water and the obstacle vertical eigenfunctions, the weights being the water or the obstacle densities, respectively.

Eliminating coefficients $A_{\nu m}^{\text{sc}}$ from the system of Eqs. (9) and (11), the following system for the coefficients B_{nm} is obtained:

$$\sum_{n=1}^{\infty} P_{\nu n, m} B_{nm} = X_{\nu m}, \quad \nu = 1, 2, \dots, \quad m = 0, 1, \dots, \quad (13)$$

where

$$P_{\nu n, m} = O_{\nu n}^w - O_{\nu n}^b \frac{\mu_n J'_m(\mu_n a)}{\lambda_{\nu} J_m(\mu_n a)} \frac{H_m^{(1)}(\lambda_{\nu} a)}{H_m^{(1)'}(\lambda_{\nu} a)}, \quad (14)$$

and

$$X_{\nu m} = -\frac{g_{\nu}(z_s)}{\rho(z_s)} \frac{H_m^{(1)}(\lambda_{\nu} r_s)}{2\pi\lambda_{\nu} a H_m^{(1)'}(\lambda_{\nu} a)}. \quad (15)$$

Observing Eq. (13), one can note that the azimuthal index m serves as an independent parameter, i.e., system (13) can be solved for each m independently. (This is due to the fact that the functions $\cos m\theta$ are the azimuthal eigenfunctions both outside and inside the cylindrical obstacle.) In the course of numerical calculations the infinite system (13) is truncated, for each m , to a finite one on the basis of the number of propagating modes. The issue is a little complicated in this case, since the number of propagating modes in the water is, in general, different than the number of propagating modes in the obstacle. See Sec. III for a detailed discussion.

Asymptotic expressions of the coefficients $P_{\nu n, m}$ and $X_{\nu m}$ for large values of m , obtained by using well-known recurrence relations for Bessel functions and their asymptotic formulas for large m , are given by

$$P_{\nu n, m} \approx O_{\nu n}^w + O_{\nu n}^b, \quad X_{nm} \approx \frac{g_{\nu}(z_s)}{2\pi\rho(z_s)} \frac{1}{m} \left(\frac{a}{r_s}\right)^m, \quad m \rightarrow \infty. \quad (16)$$

The second equation shows that $|X_{nm}| < (C_n/m) \beta^m$ ($\beta = a/r_s < 1$) uniformly with respect to n .

In the general case of arbitrary ssdps of the water and cylinder layers, the eigenvalues and eigenfunctions of the corresponding vertical problems are calculated numerically, e.g., using the algorithms presented in Refs. 21–23. Accordingly, the integrals $O_{\nu n}^w$ and $O_{\nu n}^b$ in Eq. (14) are evaluated by means of numerical integration, and the system (13) is solved for B_{nm} by means of the method of singular value

decomposition (SVD). (Information about the SVD method can be found, for example, in Refs. 24 and 25.) However, in the isovelocity case (i.e., when c_w , c_b , ρ_w , ρ_b are z -independent) the system (13) has an analytical solution, which is given in the following subsection.

B. Analytical solution to the constant-parameter case

In the case where the ssdps in water and cylinder are constant, the eigenvalues λ_n , μ_n , and the eigenfunctions $g_n(z)$, $\psi_n(z)$ are expressed analytically as

$$\lambda_n^2 = 1 - \left(n - \frac{1}{2}\right)^2 \left(\frac{\pi}{k_w H}\right)^2, \quad g_n(z) = \sqrt{\frac{2\rho_w}{H}} \sin\left\{\left(n - \frac{1}{2}\right) \frac{\pi}{H} z\right\}, \quad n = 1, 2, \dots, \quad (17a)$$

$$\mu_n^2 = 1 - \left(n - \frac{1}{2}\right)^2 \left(\frac{\pi}{k_b H}\right)^2, \quad \psi_n(z) = \sqrt{\frac{2\rho_b}{H}} \sin\left\{\left(n - \frac{1}{2}\right) \frac{\pi}{H} z\right\}, \quad n = 1, 2, \dots. \quad (17b)$$

Consequently, the integrals $O_{\nu n}^w$ and $O_{\nu n}^b$, given by Eqs. (10) and (12), respectively, take the values

$$O_{\nu n}^w = \sqrt{\frac{\rho_b}{\rho_w}} \delta_{\nu n}, \quad (18a)$$

$$O_{\nu n}^b = \sqrt{\frac{\rho_w}{\rho_b}} \delta_{\nu n}, \quad (18b)$$

where $\delta_{\nu n}$ is the Kronecker's delta. That is, $O_{\nu n}^{w,b}$ are different from zero only for $n = \nu$ and $O_{\nu n}^w O_{\nu n}^b = 1$. Then, the matrix $P_{\nu n, m}$ becomes diagonal and the infinite system (13) is solved at once, giving

$$B_{nm} = \sqrt{\gamma} X_{nm} / P_{nm}, \quad (19)$$

where

$$P_{nm} \equiv P_{nn, m} = 1 - \gamma \frac{\mu_n J'_m(\mu_n a)}{\lambda_n J_m(\mu_n a)} \frac{H_m^{(1)}(\lambda_n a)}{H_m^{(1)'}(\lambda_n a)}, \quad (20)$$

$$\gamma = \rho_w / \rho_b.$$

Substituting expressions (18a) and (19) into (9), the coefficients A_{nm}^{sc} are found to be

$$A_{nm}^{\text{sc}} = -\frac{1}{\rho(z_s)} \left[\frac{J_m(\mu_n a)}{2\pi D_{nm}} + \frac{i}{4} J_m(\lambda_n a) \right], \quad (21)$$

where

$$D_{nm} = \lambda_n a J_m(\mu_n a) H_m^{(1)'}(\lambda_n a) - \gamma \mu_n a J'_m(\mu_n a) H_m^{(1)}(\lambda_n a). \quad (22)$$

Introducing the expressions (19) for B_{nm} and (21) for A_{nm}^{sc} into Eqs. (5) and (3), respectively, we finally obtain the pressure field p_b , inside the cylindrical obstacle, and p_w , in the surrounding water, as follows:

$$p_b = - \sum_{n=1}^{\infty} \sum_{m=0}^{\infty} e_m \left[\frac{\sqrt{\gamma}}{2\pi\rho(z_s)D_{nm}} \right] \times H_m^{(1)}(\lambda_n r_s) g_n(z_s) J_m(\mu_n r) \psi_n(z) \cos m\theta, \quad (23)$$

$$p_w = p_i - \sum_{n=1}^{\infty} \sum_{m=0}^{\infty} e_m \left[\frac{i}{4\rho(z_s)} \frac{J_m(\lambda_n a)}{H_m^{(1)}(\lambda_n a)} \right] \times H_m^{(1)}(\lambda_n r_s) g_n(z_s) H_m^{(1)}(\lambda_n r) g_n(z) \cos m\theta - \sum_{n=1}^{\infty} \sum_{m=0}^{\infty} e_m \left[\frac{J_m(\mu_n a)}{2\pi\rho(z_s)H_m^{(1)}(\lambda_n a)D_{nm}} \right] \times H_m^{(1)}(\lambda_n r_s) g_n(z_s) H_m^{(1)}(\lambda_n r) g_n(z) \cos m\theta, \quad (24)$$

where p_i is the axisymmetric field of a point source in the unobstructed waveguide [see Eq. (3)].

A very interesting feature resulting from Eq. (24) is that the second term on the right-hand side represents the scattered field corresponding to an ideally soft cylindrical obstacle, p_{sc}^{soft} [see Eqs. (1)–(2a)], studied in Refs. 17 and 18. That is, the total pressure field p_w , outside a penetrable cylindrical obstacle, equals the total pressure field corresponding to an ideally soft cylindrical obstacle, $p_i + p_{sc}^{\text{soft}}$, plus a pressure field, p_{sc}^{obs} , which involves coefficients expressing the modal penetrability of the obstacle. Also, from Eqs. (23) and (24), it becomes clear that the series expansions of the fields p_b , p_{sc}^{soft} , and p_{sc}^{obs} have the same structure. (This is also valid for the scattered field of an ideally hard cylinder, p_{sc}^{hard} .) The coefficients included in brackets in Eqs. (23) and (24) have the general form $C_{nm} = C_{nm}(\gamma; \lambda_n a; \mu_n a)$.

Existing analytic solutions to simpler propagation/scattering problems can be easily obtained as special cases of solution (23) and (24):

- (i) *The special case of a ‘‘water obstacle’’ (no obstacle).* Setting, in Eqs. (23) and (24), $c_b = c_w$ and $\rho_b = \rho_w$, using Graf’s addition theorem [Eq. (8)] and taking into account the relation²⁰

$$W\{J_m(x), H_m^{(1)}(x)\} = J_m(x)H_m^{(1)'}(x) - J_m'(x)H_m^{(1)}(x) = \frac{2i}{\pi x}, \quad (25)$$

we obtain that both pressure fields p_w and p_b degenerate to the axisymmetric field in the unobstructed waveguide, p_i , as expected.

- (ii) *The limiting case of an ideally hard obstacle.* Assuming that $\rho_b \rightarrow \infty$ (i.e., $\gamma \rightarrow 0$) and using Eq. (25), the third term, p_{sc}^{obs} , in Eq. (24) can be expressed as the difference $p_{sc}^{\text{hard}} - p_{sc}^{\text{soft}}$. This means that p_w tends to the pressure field corresponding to an ideally hard cylindrical obstacle, $p_i + p_{sc}^{\text{hard}}$ [see Eqs. (1)–(2b)]. Also, in this case, p_b tends to zero. Both results are physically expected.

C. Scattering cross section

Recall that the time-average (t -average for short) acoustic power radiated by a source enclosed by a surface S is given by

$$\mathcal{P}_{av} = \int \int_S \mathbf{I}_{av} \cdot \mathbf{n} dS, \quad (26)$$

where

$$\mathbf{I}_{av} = \frac{1}{2} \Re\{p \mathbf{v}^*\} \quad (27)$$

is the t -average acoustic energy flux or acoustic intensity, and \mathbf{n} is the unit normal pointing out of the volume containing the source.²⁶ In Eq. (27), p and \mathbf{v} are the complex amplitudes of the pressure and the fluid velocity, respectively, and the asterisk denotes a complex conjugate.

According to Eq. (26), the t -average power \mathcal{P}_{sc} of a scattered field due to the presence of a cylindrical object in a waveguide is expressed as

$$\mathcal{P}_{sc} = \int \int_S I_{sc} dS, \quad (28)$$

where I_{sc} is the t -average intensity of the scattered wave and S is now taken to be a cylindrical surface of arbitrary radius r enclosing the source. Assuming that the sound source is far enough from the object so that the incident sound wave is approximately plane, and the receiver is at large distance from the scattering object, the scattering cross section σ_{sc} is given by²⁷

$$\sigma_{sc} = \frac{\mathcal{P}_{sc}}{\overline{I_p}} = \frac{\iint_S I_{sc} dS}{\overline{I_p}} = \frac{\int_0^{2\pi} \int_0^H |p_{sc}|^2 r d\theta dz}{|\overline{p_p}|^2}, \quad (29)$$

where p_{sc} is the pressure corresponding to the scattered field, p_p is the pressure of the incident, effectively plane, wave at the scatterer’s position, I_p is the t -average intensity of the incident field at the scatterer’s position, and the bar denotes averaging over depth, i.e., it stands for the operator $H^{-1} \int_0^H (\cdot) dz$. Note that the depth-averaging appearing in Eq. (29) is due to the fact that the incident, effectively plane, wave is actually depth dependent because of the waveguide.

The goal in this subsection is to derive an analytic formula for the scattering cross section of an impenetrable or penetrable cylindrical obstacle in a plane waveguide, in terms of the expansion coefficients of the pressure field.

The pressure field p_p is the asymptotic approximation to p_i [see Eq. (3)] for large horizontal distance at the scatterer’s position, and it is given by¹

$$p_p = p_p(z) = \frac{i}{\rho(z_s) \sqrt{8\pi r_s}} e^{-i\pi/4} \sum_{n=1}^{\infty} g_n(z_s) g_n(z) \frac{e^{i\lambda_n r_s}}{\sqrt{\lambda_n}}. \quad (30)$$

Also, substituting the large-argument asymptotics for Hankel functions $H_m^{(1)}(x)$ (Ref. 20) in the formulas for the scattered field of an impenetrable or penetrable cylindrical obstacle [Eqs. (1) and (3), respectively], we obtain

$$p_{sc}^{\text{imp}} = - \frac{1}{2\pi\rho(z_s) \sqrt{r r_s}} \sum_{n=1}^{\infty} \sum_{m=0}^{\infty} e_m (-1)^m \Lambda_m(\lambda_n a) \times \frac{e^{i\lambda_n(r+r_s)}}{\lambda_n} g_n(z_s) g_n(z) \cos m\theta, \quad (31)$$

and

$$p_{sc}^{pen} = \frac{2i}{\pi \sqrt{r r_s}} \sum_{n=1}^{\infty} \sum_{m=0}^{\infty} e_m (-1)^m \frac{A_{nm}^{sc}}{H_m^{(1)}(\lambda_n a)} \times \frac{e^{i\lambda_n(r+r_s)}}{\lambda_n} g_n(z_s) g_n(z) \cos m\theta, \quad (32)$$

where the coefficients $\Lambda_m(\lambda_n a)$ are given by Eq. (2).

To calculate the scattering cross section using Eqs. (29)–(32), integrals of the form

$$I_1 = \int_0^H \left| \sum_n b_n g_n(z) \right|^2 dz, \quad (33a)$$

$$I_2 = \int_0^H \int_0^{2\pi} \left| \sum_{nm} d_{nm} g_n(z) \cos m\theta \right|^2 d\theta dz, \quad (33b)$$

have to be calculated, where b_n and d_{nm} are complex constants. Exploiting the orthogonality of the vertical and azimuthal eigenfunctions, $g_n(z)$ and $\cos m\theta$, respectively, we obtain

$$\frac{1}{\rho(z)} I_1 = \sum_{n=1}^{\infty} |b_n|^2, \quad (34a)$$

$$\frac{1}{\rho(z)} I_2 = \pi \sum_{n=1}^{\infty} \sum_{m=0}^{\infty} \epsilon_m |d_{nm}|^2, \quad (34b)$$

where $\epsilon_0 = 2$, $\epsilon_m = 1$, $m > 0$.

Combining properly Eqs. (29)–(34) now, it is found that the scattering cross sections for the cases of an impenetrable and a penetrable cylindrical obstacle in a plane waveguide are given by

$$\sigma_{sc}^{imp} = 2H \frac{\sum_{n=1}^{\infty} \sum_{m=0}^{\infty} \epsilon_m e_m^2 g_n^2(z_s) |\Lambda_m(\lambda_n a)|^2 / \lambda_n^2}{\sum_{n=1}^{\infty} g_n^2(z_s) / \lambda_n} \quad (35)$$

and

$$\sigma_{sc}^{pen} = 2H \frac{\sum_{n=1}^{\infty} \sum_{m=0}^{\infty} \epsilon_m e_m^2 g_n^2(z_s) |4\tilde{A}_{nm}^{sc}|^2 / \lambda_n^2}{\sum_{n=1}^{\infty} g_n^2(z_s) / \lambda_n}, \quad (36)$$

respectively, where

$$|\Lambda_m(\lambda_n a)|^2 = \frac{1}{1 + \tau_m^2(\lambda_n a)}, \quad (37)$$

$$\tau_m^{soft}(\lambda_n a) = \frac{Y_m(\lambda_n a)}{J_m(\lambda_n a)}, \quad (38a)$$

$$\tau_m^{hard}(\lambda_n a) = \frac{Y'_m(\lambda_n a)}{J'_m(\lambda_n a)}, \quad (38b)$$

and

$$\tilde{A}_{nm}^{sc} = \frac{\rho(z_s)}{H_m^{(1)}(\lambda_n a)} A_{nm}^{sc}. \quad (39)$$

For the case of an isovelocity ocean waveguide, the coefficients A_{nm}^{sc} are given by Eq. (21). Results of a systematic

numerical investigation of σ_{sc}^{imp} and σ_{sc}^{pen} versus frequency for various characteristic values of environmental parameters are presented in Sec. III B 2.

III. NUMERICAL RESULTS

Except for the well-established and widely used notions of *propagating modes* and *evanescent modes*, the following terminology is used throughout this section:

- (i) *azimuthal-evanescent modes*: modes decaying exponentially with respect to the azimuthal index m . (For a detailed discussion about azimuthal-evanescent and azimuthal-propagating modes, see Ref. 17, where these notions have been introduced.);
- (ii) *w-p modes*: propagating modes in the water column;
- (iii) *w-e modes*: evanescent modes in the water column;
- (iv) *o-p modes*: propagating modes in the obstacle column;
- (v) *o-e modes*: evanescent modes in the obstacle column;
- (vi) n_{max}^w, n_{max}^b : the number of vertical terms required to achieve numerical convergence of p_w and p_b , respectively;
- (vii) n_{cr}^w, n_{cr}^b : the number of *w-p* and *o-p* modes, respectively;
- (viii) m_{max}^w, m_{max}^b : the number of azimuthal terms required to achieve numerical convergence of p_w^{sc} and p_b , respectively;
- (ix) m_{cr}^w, m_{cr}^b : the critical value separating azimuthal-propagating from azimuthal-evanescent modes in water and obstacle, respectively.

A. General aspects of the numerical investigation

Whenever a calculation of infinite series is taking place, a basic question is the proper truncation of the series. As concerns the vertical terms, since the field is calculated sufficiently far from the source point, the contribution from the *w-e* modes outside the obstacle is expected to be negligible; see also Ref. 5. In fact, preliminary numerical results have confirmed the assumption that an accurate calculation of p_w can be achieved by retaining only the *w-p* modes.

Taking into account that $c_w < c_b$ and $\rho_w < \rho_b$, it follows that $n_{cr}^w > n_{cr}^b$. For consistency reasons, the number of vertical terms retained in the series (3) and (5), representing the pressure field in water and obstacle, respectively, is taken to be $n_{max}^w = n_{max}^b = n_{cr}^w > n_{cr}^b$. This means that a number of $n_{cr}^w - n_{cr}^b$ *o-e/w-p* modes, i.e., evanescent with respect to the obstacle column but propagating with respect to the water column, have to be included in the calculations of both fields. The essential argument behind the choice $n_{max}^w = n_{max}^b$ is the requirement that the (finite) dimension of the linear subspace in which the pressure field is projected (by truncation) should be the same throughout the physical domain $D = D^w \cup D^b$. Otherwise, a vertical pattern lying in an n_{max}^w -dimensional subspace should be matched with another vertical pattern lying in a different-dimension subspace. This is conceptually awkward and, in most cases, numerically inefficient.

In accordance with the above remarks, the pressure fields p_w and p_b can be written in the form

$$p_w = \underbrace{p_w(n \leq n_{cr}^b)}_{I_a} + \underbrace{p_w(n_{cr}^b < n \leq n_{cr}^w)}_{I_b} + \underbrace{p_w(n > n_{cr}^w)}_{I_c}, \quad (40a)$$

$$p_b = \underbrace{p_b(n \leq n_{cr}^b)}_{II_a} + \underbrace{p_b(n_{cr}^b < n \leq n_{cr}^w)}_{II_b} + \underbrace{p_b(n > n_{cr}^w)}_{II_c}, \quad (40b)$$

where, for example, $I_b = p_w(n_{cr}^b < n \leq n_{cr}^w)$ denotes the part of the pressure field obtained by retaining the $n_{cr}^w - n_{cr}^b$ terms in n -summation.

The numerical investigation of the terms I_a, \dots, II_c , for a significant number of different sets of parameters, leads to the following conclusions:

- (i) In the water, the sum I_a of the n_{cr}^b w - p/o - p modes has the dominant contribution to the acoustic field. The sum I_b of the $n_{cr}^w - n_{cr}^b$ w - p/o - e has an essential contribution in the insonified region (i.e., between the source and the obstacle) and a much smaller one behind the obstacle. The sum I_c of the w - e/o - e modes is negligible, as already mentioned.
- (ii) Within the obstacle, the sum II_a of the w - p/o - p modes contributes almost 100% to the acoustic field. The sum II_b of the $n_{cr}^w - n_{cr}^b$ w - p/o - e modes is non-negligible only in a narrow zone around the front boundary of the obstacle. The remaining evanescent modes (sum II_c) are of no importance to the acoustic field.

Quantitative information about the terms I_b and II_b will be given in the following subsection.

As concerns the rate of azimuthal convergence of the double series expressing the scattered field outside the obstacle, p_w^{sc} , and the total field inside the obstacle, p_b , the following conclusions have been drawn:

- (i) In the water, numerical investigation gives results quite similar with those already established in Ref. 17, in relation to the study of the diffraction series for an impenetrable cylindrical obstacle. According to that, $m_{cr}^w = \lambda_n a$. Consequently, an estimation of a lower bound of m_{max}^w is derived by imposing $m_{max}^w > m_{cr}^w (n = 1) = \lambda_1 a$. [This result is due to the fact that the rate of azimuthal convergence of the normal-mode series for the diffraction field in both cases—penetrable or impenetrable obstacle—is dominated by the rate of decay of the Bessel functions $J_m(\lambda_n a)$, considered as functions of the index m . See Ref. 17 for a detailed analysis.]
- (ii) By a similar argumentation, it can also be proved that, inside the obstacle, $m_{cr}^b = \lambda_n r$. Consequently, an estimation of a lower bound of m_{max}^b is derived by imposing $m_{max}^b > m_{cr}^b (n = 1) = \lambda_1 r$. This relation directly implies that the more a receiver approaches the center of the obstacle the less azimuthal terms are required for the convergence of the series (5).

The accuracy and the efficiency of the numerical calculation of the pressure fields p_w and p_b also depend on the

algorithms used for the numerical calculation of the eigenvalues and eigenvectors, Bessel and Hankel functions, as well as on the algorithm performing singular value decomposition for solving the linear systems (13) in the general case of depth-dependent ssdps. Real eigenvalues and the corresponding eigenvectors have been calculated by KRAKEN,^{21,22} while imaginary eigenvalues and the associated eigenvectors have been calculated using MATLAB. The other calculations have been based on subroutines from *Numerical Recipes*.²⁵ All codes used are robust and have been tested in numerous cases.

B. Results

1. Transmission loss

Because of the multiparametric character of the studied problem, it is inevitable that some parameters should be kept constant. In the present subsection the unaltered parameters are the depth of the waveguide $H = 250$ m, the radius of the cylindrical obstacle $a = 200$ m, the distance between source and cylinder axis $r_s = 1000$ m, and the water density $\rho_w = 1$ g/cm³. The frequency of the source ranges from 20 to 500 Hz, corresponding to nondimensional wave numbers $17 < k_w a < 420$ and $13 < k_b a < 322$.

The contribution of propagating and evanescent modes to the acoustic field outside and inside a penetrable cylindrical obstacle is illustrated by means of a set of transmission loss (TL) graphs [Fig. 2(a)–(d)] at four different frequencies for an environment with parameters (in addition to the ones given above) $z_s = 100$ m, $c_b = 1950$ m/s, $\rho_b = 2.1$ g/cm³, $c_w = 1500$ m/s. Three distinct curves are presented in each figure: (i) TL versus range along the horizontal line passing through the source point and the center of the obstacle, as obtained by summing up to n_{cr}^w modes, i.e., by retaining the n_{cr}^b w - p/o - p modes plus $n_{cr}^w - n_{cr}^b$ w - p/o - e modes (solid line); (ii) same as (i) but excluding from the calculations the $n_{cr}^w - n_{cr}^b$ w - p/o - e modes (dashed line); and (iii) the absolute difference between (i) and (ii) (dotted line starting from nearly 0 dB). In all figures the source is considered at the origin of the horizontal axis, and the position of the cylindrical obstacle is determined by the long-dashed vertical lines. It is clearly shown that the contribution of the $n_{cr}^w - n_{cr}^b$ modes to the acoustic field is significant in the insonified region, which means that the backscattered field is essentially affected by the sum I_b in Eq. (40a). On the other hand, I_b has a minimal contribution to the field in the shadow zone, except in the case of very low frequencies [see Fig. 2(a)]. As concerns the acoustic field inside the obstacle, the sum of the $n_{cr}^w - n_{cr}^b$ modes [II_b in Eq. (40b)] has an effect only in a narrow zone around the front boundary (i.e., the insonified side) of the obstacle.

In Fig. 3(a)–(d) transmission losses are calculated and compared on a horizontal plane passing through the source for four different cases, corresponding to cylindrical obstacles with different acoustic properties: (i) ideally soft cylinder [homogeneous Dirichlet condition on the boundary, Fig. 3(a)]; (ii) acoustic material with $c_b = 1950$ m/s, $\rho_b = 2.1$ g/cm³ [moraine, Fig. 3(b)]; (iii) acoustic material with $c_b = 5000$ m/s, $\rho_b = 5$ g/cm³ [Fig. 3(c)]; and (iv) ideally hard

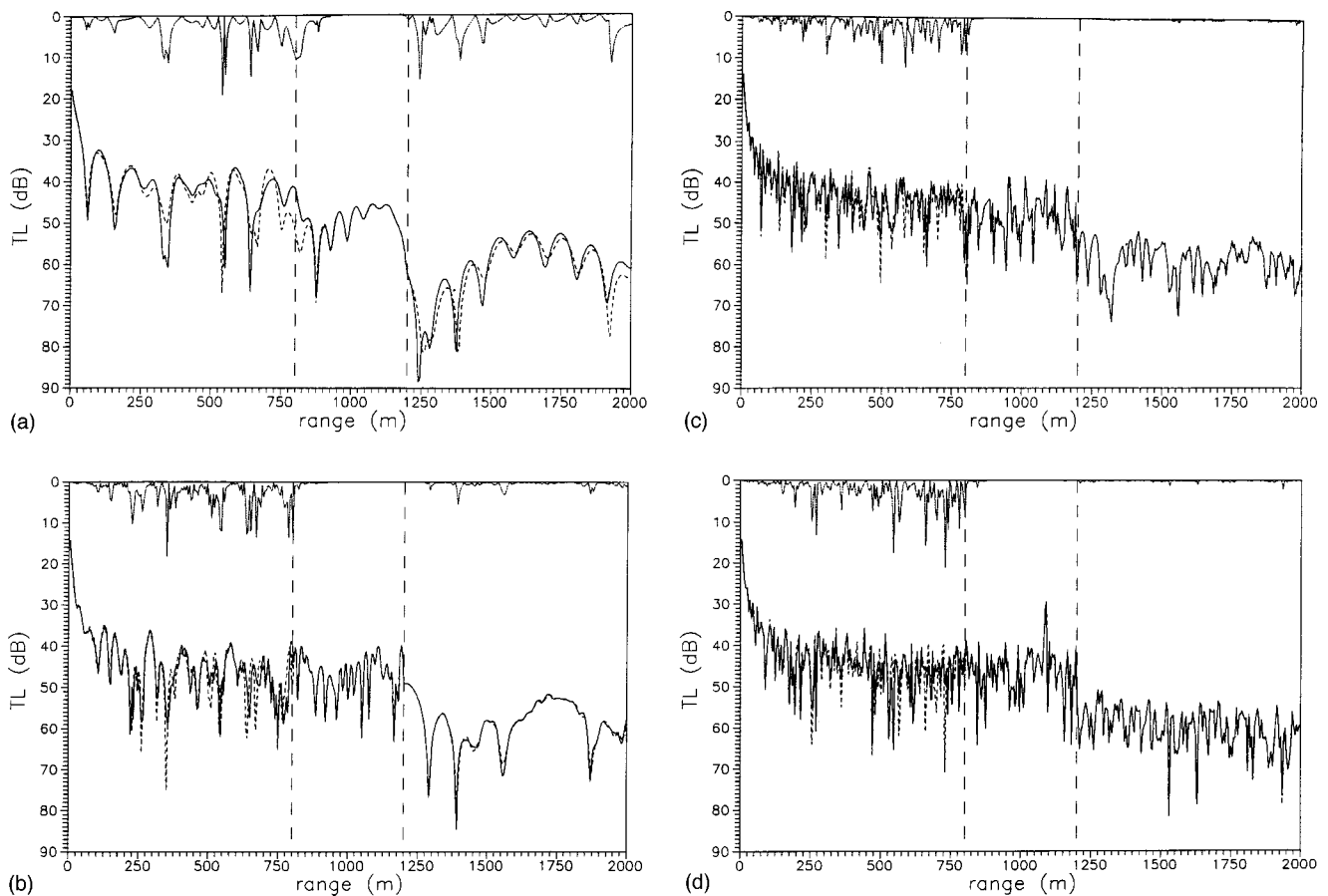


FIG. 2. (a) TL versus range at $f=20$ Hz along the horizontal line passing through the source point and the center of a penetrable cylindrical obstacle with $c_b=1950$ m/s, $\rho_b=2.1$ g/cm³. The other parameters of the environment are $H=250$ m, $a=200$ m, $z_s=100$ m, $r_s=1000$ m, $c_w=1500$ m/s, $\rho_w=1$ g/cm³. The acoustic field is calculated taking into account (i) n_{cr}^w modes (solid curve), and (ii) n_{cr}^b modes (dashed curve). The dotted curve starting from nearly 0 dB represents the absolute difference between (i) and (ii). (b) The same as (a), except that $f=60$ Hz. (c) The same as (a), except that $f=200$ Hz. (d) The same as (a), except that $f=500$ Hz.

cylinder [homogeneous Neumann condition on the boundary, Fig. 3(d)]. The results presented correspond to a source frequency $f=500$ Hz, source depth $z_s=H/3$, and a constant sound speed in water $c_w=1500$ m/s. It is observed that the acoustic field in the insonified region is very similar in all cases, but a quite different distribution of losses appears in the shadow zone. Furthermore, by comparing the four cases, it can be seen that the strongest scatterer seems to be the ideally soft cylinder.

Finally, in Fig. 4(a)–(c) transmission losses are calculated on a vertical slice passing through the source and the axis of layered cylindrical obstacles, for the depth-dependent sound-speed profile in water depicted in Fig. 5. The source frequency and depth are $f=60$ Hz and $z_s=100$ m, respectively. In these figures, the layered obstacles are a clay-water cylinder [i.e., “floating obstacle,” $c_b=1600$ m/s, $\rho_b=1.6$ g/cm³, Fig. 4(a)], a water-moraine cylinder [i.e., “seamount,” Fig. 4(b)], and a clay-moraine cylinder obtained as a combination of the two previous cases [Fig. 4(c)]. In all cases the horizontal interfaces are at $h_b=H/2$. Well-formed energy paths (rays) and shadow areas can be seen in all figures. It can be observed that those paths of concentrated energy, resulting from multiple scattering between the horizontal interfaces of the obstacle and the waveguide bound-

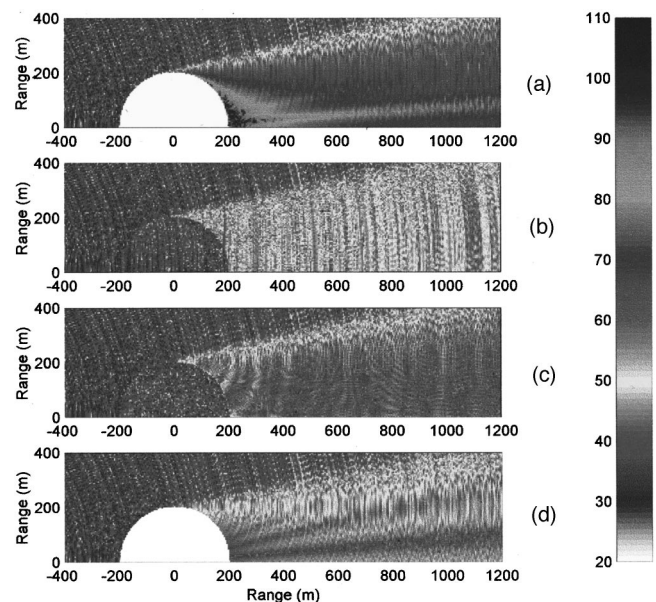


FIG. 3. TL at $f=500$ Hz calculated on a horizontal plane passing through the source point. The cylindrical obstacle is impenetrable in (a) (ideally soft) and (d) (ideally hard), and penetrable in (b) ($c_b=1950$ m/s, $\rho_b=2.1$ g/cm³) and (c) ($c_b=5000$ m/s, $\rho_b=5$ g/cm³). The source depth is $z_s=H/3$, and the other parameters are as in Fig. 2(a).

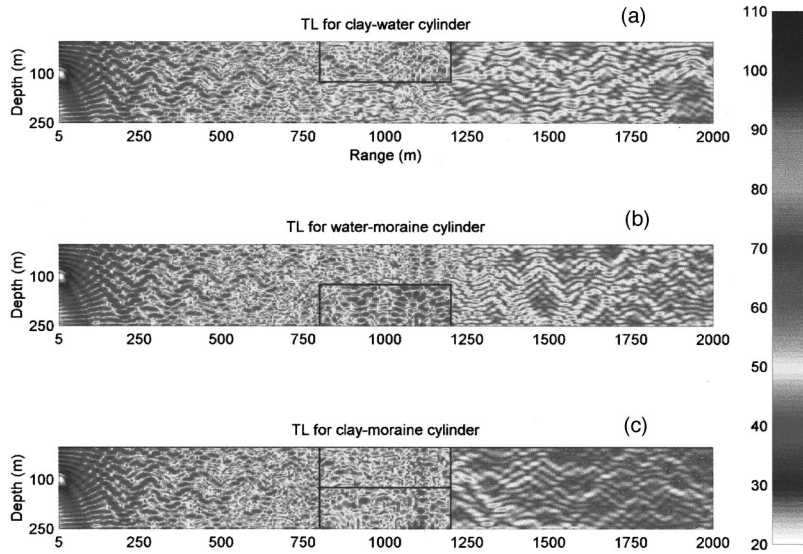


FIG. 4. TL at $f=60$ Hz calculated on a vertical slice passing through the source and the axis of the obstacle for the depth-dependent ssdp of Fig. 5. (a) Floating obstacle: $c_b=1600$ m/s, $\rho_b=1.6$ g/cm³ (clay), (b) seamount: $c_b=1950$ m/s, $\rho_b=2.1$ g/cm³ (moraine), (c) clay-moraine obstacle. The horizontal interface in (a)–(c) is at $h_b=H/2$, and the other parameters are as in Fig. 2(a).

ary, are more intensive in Fig. 4(b) than in Fig. 4(a). Note that in Fig. 4(b) the waveguide reflector is an ideally soft boundary (the free surface) and the values of the acoustic properties of the obstacle are higher than in Fig. 4(a). The following general remarks arise from Fig. 4(a)–(c): (i) in the insonified region the total pressure field is similar in all cases, (ii) increasing the values of the obstacle acoustic parameters implies that more energy is trapped inside the obstacle, (iii) the concentration of the trapped energy is higher in the area between the axis of the obstacle and the back interface, due to the inner propagation and backscattering from that interface. (Colored versions of figures 3 and 4, better revealing the above described features, are available upon request.)

2. Scattering cross section

Systematic numerical investigation of the nominator and the denominator of Eqs. (35) and (36) as frequency f increases has resulted in the conjectures

$$\lim_{f \rightarrow \infty} \sum_{n=1}^{\infty} g_n^2(z_s)/\lambda_n = \frac{H}{4}, \quad (41)$$

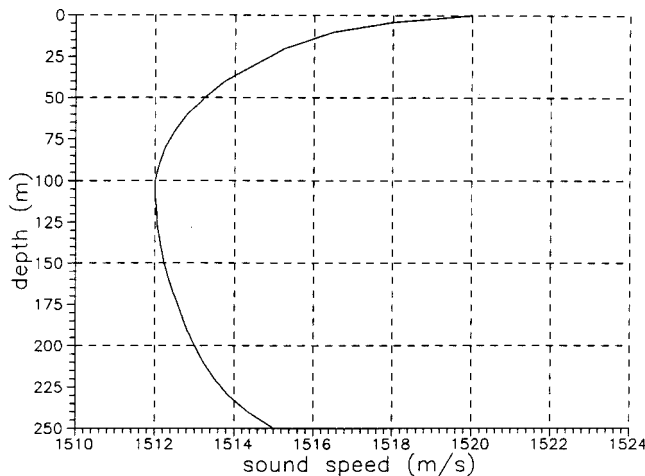


FIG. 5. The depth-dependent sound-speed profile used for the TL calculation in Fig. 4(a)–(c).

$$\lim_{f \rightarrow \infty} \sum_{n=1}^{\infty} \sum_{m=0}^{\infty} \epsilon_m e_m^2 g_n^2(z_s) |\Lambda_m(\lambda_n a)|^2 / \lambda_n^2 = \frac{Ha}{2}, \quad (42)$$

$$\lim_{f \rightarrow \infty} \sum_{n=1}^{\infty} \sum_{m=0}^{\infty} \epsilon_m e_m^2 g_n^2(z_s) |4\tilde{A}_{nm}^{sc}|^2 / \lambda_n^2 = \frac{Ha}{2}, \quad (43)$$

which imply that

$$\lim_{f \rightarrow \infty} \sigma_{sc}^{imp} = \lim_{f \rightarrow \infty} \sigma_{sc}^{pen} = 4Ha. \quad (44)$$

It is worth noticing that Eq. (44) is expected on the basis of geometrical-acoustics considerations, which require

$$\sigma_{sc}^g \rightarrow 2A_{proj}, \quad (45)$$

where A_{proj} is the projected area of the scattering obstacle.²⁶ An attempt for the analytic proof of Eqs. (41)–(43) has not been undertaken up to now.

The scattering cross sections obtained by Eqs. (35) and (36) are presented versus frequency in Fig. 6(a)–(d) for various combinations of the parameters of an isovelocity environment. In all figures $c_w=1500$ m/s, $\rho_w=1$ g/cm³, $z_s=100$ m, and $Ha=5 \times 10^4$ m².

Figure 6(a) shows the scattering cross section of an ideally soft and an ideally hard cylindrical obstacle, σ_{sc}^{soft} and σ_{sc}^{hard} , respectively, versus frequency for three different combinations of the waveguide depth H and the obstacle radius a . It is clearly seen that: (i) for all combinations of H and a , and for all frequencies, $\sigma_{sc}^{hard} < \sigma_{sc}^{soft}$; (ii) σ_{sc}^{hard} and σ_{sc}^{soft} are symmetric with respect to the axis $\sigma = \sigma_{sc}^g = 4Ha$; and (iii) σ_{sc}^{hard} is always less than the limiting value σ_{sc}^g , while σ_{sc}^{soft} is always greater than the limiting value σ_{sc}^g . Additionally, for $Ha = \text{const.}$, the scattering cross section converges faster to σ_{sc}^g as a increases.

In Fig. 6(b) and (c) the scattering cross section of cylindrical obstacles of the same geometry ($H=500$ m, $a=100$ m) versus frequency is presented for different values of their

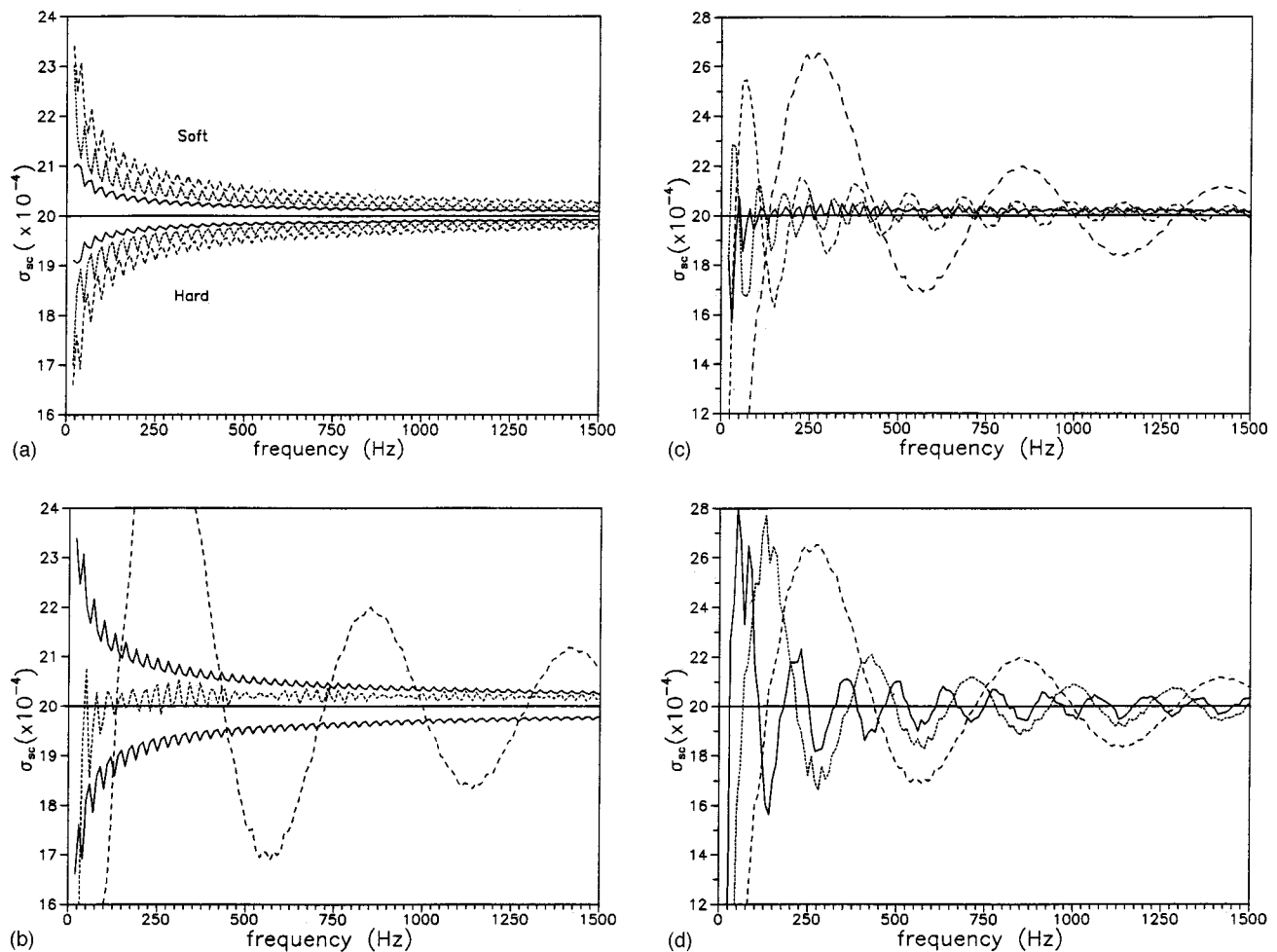


FIG. 6. (a) Scattering cross section versus frequency for an ideally soft (curves above $\sigma_{sc}^g = 4Ha = 20 \times 10^4 \text{ m}^2$) and an ideally hard (curves below σ_{sc}^g) cylindrical obstacle. Medium-dash line: $H = 500 \text{ m}$, $a = 100 \text{ m}$; short-dash line: $H = 250 \text{ m}$, $a = 200 \text{ m}$; solid line: $H = 125 \text{ m}$, $a = 400 \text{ m}$. The other parameters of the environment are $c_w = 1500 \text{ m/s}$, $\rho_w = 1 \text{ g/cm}^3$, and $z_s = 100 \text{ m}$. (b) Scattering cross section versus frequency for cylindrical obstacles with $H = 500 \text{ m}$ and $a = 100 \text{ m}$. Medium-dash line: $c_b = 1520 \text{ m/s}$, $\rho_b = 1.4 \text{ g/cm}^3$; short-dash line: $c_b = 1950 \text{ m/s}$, $\rho_b = 2.1 \text{ g/cm}^3$; solid line: ideally soft (above σ_{sc}^g) and ideally hard (below σ_{sc}^g) obstacle. The other parameters of the environment are as in (a). (c) Scattering cross section versus frequency for penetrable cylindrical obstacles with $H = 500 \text{ m}$ and $a = 100 \text{ m}$. Long-dash line: $c_b = 1520 \text{ m/s}$, $\rho_b = 1.4 \text{ g/cm}^3$; medium-dash line: $c_b = 1578 \text{ m/s}$, $\rho_b = 1.5 \text{ g/cm}^3$; short-dash line: $c_b = 1677 \text{ m/s}$, $\rho_b = 1.83 \text{ g/cm}^3$; solid line: $c_b = 1950 \text{ m/s}$, $\rho_b = 2.1 \text{ g/cm}^3$. The other parameters of the environment are as in (a). (d) Scattering cross section versus frequency for penetrable cylindrical obstacles with $c_b = 1520 \text{ m/s}$, $\rho_b = 1.4 \text{ g/cm}^3$. Medium-dash line: $H = 500 \text{ m}$, $a = 100 \text{ m}$; short-dash line: $H = 250 \text{ m}$, $a = 200 \text{ m}$; solid line: $H = 125 \text{ m}$, $a = 400 \text{ m}$. The other parameters of the environment are as in (a).

acoustic properties. It is shown that the cross section of a penetrable obstacle exhibits pronounced oscillations which decay while tending to the high-frequency limit σ_{sc}^g . The lower are the values of the obstacle acoustic parameters (c_b , ρ_b); the larger are the amplitude and the cycles of the oscillations, which implies slower convergence to σ_{sc}^g . Furthermore, when c_b , ρ_b become high enough, the curve $\sigma_{sc}(f)$ for the penetrable obstacle lies between the corresponding curves for the ideally soft and hard obstacle [see Fig. 6(b), curve for moraine cylindrical obstacle].

Finally, in Fig. 6(d) the scattering cross section of clay-silt cylindrical obstacles versus frequency is presented for three different combinations of H and a . It is observed that, for $Ha = \text{const.}$, the cycles of the oscillations around σ_{sc}^g become larger and $\sigma_{sc}(f)$ converges more slowly to its limiting value as a decreases.

IV. CONCLUSIONS AND DISCUSSION

In the present work the problem of the three-dimensional (3D) acoustic scattering from horizontally mul-

tilayered penetrable cylindrical obstacles in shallow-water environments with arbitrary horizontal stratification and depth-dependent sound-speed and density profiles is studied. The field outside the obstacle is decomposed into a known incident field (the axisymmetric, unobstructed contribution of the source) and a scattered field, which is represented as a double-series, normal-mode expansion using Hankel functions. The pressure field inside the obstacle is represented as a double-series, normal-mode expansion using Bessel functions. The application of the matching conditions at the water-obstacle interface introduces mode coupling and results in an infinite system of equations which, after an appropriate truncation, is solved numerically for the unknown coefficients of the series expansion. In the special case of constant sound-speed and density profiles, an exact analytical solution for the pressure fields is provided.

Specific attention has been paid to the fact that there is a number of coupled modes, propagating with respect to the water column but evanescent with respect to the obstacle

column. Extensive numerical investigation has shown that this part of the solution should be included in the calculations in order to obtain physically meaningful and accurate results for the pressure fields, both inside and outside the obstacle. More specifically, results for the transmission loss (TL) for an isovelocity cylindrical obstacle in an isovelocity ocean waveguide reveal that the contribution of water-propagating/obstacle-evanescent ($w-p/o-e$) modes constitutes a significant part of the backscattered field for low and low-to-moderate frequencies. Under this light, the methodology of Ref. 12, in which the $w-p/o-e$ modes are ignored, is incomplete, in the sense that it is not able to provide an accurate description of the field if the number of propagating modes is different outside and inside the cylinder. Note that, in the above reference, numerical results are given only for source frequency $f=10$ Hz, in which case the number of propagating modes is the same throughout the domain.

TL results are also given for double-layered acoustic cylindrical structures (a floating obstacle, a seamount, and a composite "island") in a typical depth-dependent ocean environment, and representative comparisons are presented for the transmission losses corresponding to impenetrable and penetrable cylindrical obstacles. Furthermore, analytical expressions for the scattering cross section of cylindrical obstacles are derived in terms of the expansion coefficients of the pressure field, and numerical results are presented versus frequency in the range $20 \text{ Hz} \leq f \leq 1500 \text{ Hz}$. The calculated cross section of both impenetrable and penetrable obstacles tends to the asymptotic value obtained by means of geometrical acoustics.

The systematic numerical investigation of the problem treated in this paper renders the present solution a useful benchmark for 3D propagation/scattering codes working at low to moderate frequencies. The emphasis is on shallow-water scattering predictions, since no fully 3D, fully coupled scattering models exist, as the authors are aware of. The authors are willing to investigate different environments and/or calculate different quantities, upon request by other researchers, for intercomparison and validation purposes.

The simplistic assumption of the rigid (perfectly reflecting) bottom is a limitation of the presented benchmark solution with respect to a category of numerical models that can give very good predictions over more realistic (penetrable) ocean bottoms, but do not behave well if the bottom is rigid. (For example, split-step parabolic equation models based on split-step Fourier transform.) However, it should be noted here that the benchmark solutions given in this work can be easily extended to more realistic (penetrable) ocean bottoms: no modifications of the pressure-field series expansions (inside and outside the cylinder) and the methodology of the solution are required; the only difference is that the eigenvalues and eigenfunctions of the new vertical eigenvalue problems (which will include the layers of the bottom) have to be calculated by means of appropriate, currently available, codes. (See, e.g., Refs. 22, 23, 28.) On the other hand, 3D parabolic equation propagation models based on finite-difference and finite-element techniques can handle a rigid bottom. Thus, if the latter are further developed to include

backscattering, they could benefit from the benchmark solution as exactly given in this paper.

The presented method can be generalized to obstacles of more complex structure, e.g., radially layered penetrable scatterers, a problem which is currently under consideration.²⁹ Other important subjects for further study arising from this work are a systematic high-frequency asymptotic treatment along the lines of Ref. 17, and the extension of the method for handling cylindrical obstacles made of elastic material.

ACKNOWLEDGMENTS

The helpful discussions of both authors with Dr. K.A. Belibassakis are deeply appreciated. The authors also thank the reviewers for their constructive comments. The present work was partially supported by the Greek Navy. The second author was also supported by the Greek State Scholarships Foundation.

- ¹F. B. Jensen, W. A. Kuperman, M. B. Porter, and H. Schmidt, *Computational Ocean Acoustics* (American Institute of Physics, New York, 1994).
- ²*Oceanography and Acoustics: Prediction and Propagation Models*, edited by A. R. Robinson and D. Lee (American Institute of Physics, New York, 1994).
- ³D. Lee and M. H. Schultz, *Numerical Ocean Acoustic Propagation in Three Dimensions* (World Scientific, Singapore, 1995).
- ⁴A. Tolstoy, "3D propagation issues and models," *J. Comput. Acoust.* **4**, 243–271 (1996).
- ⁵G. A. Athanassoulis and A. M. Prospathopoulos, "Three-dimensional acoustic scattering of a source-generated field from a cylindrical island," *J. Acoust. Soc. Am.* **100**, 206–218 (1996).
- ⁶K. Chandra, S. K. Isler, and C. Thompson, "Acoustic scattering in three-dimensional fluid media," *J. Acoust. Soc. Am.* **98**, 3462–3468 (1995).
- ⁷Z. Ye, "A novel approach to sound scattering by cylinders of finite length," *J. Acoust. Soc. Am.* **102**, 877–884 (1997).
- ⁸D. K. Dacol, "A coupled-amplitude approach to solving the Helmholtz equation," *J. Acoust. Soc. Am.* **101**, 2566–2570 (1997).
- ⁹R. B. Evans, "Three dimensional acoustic scattering from a cylindrical inclusion in a waveguide," in *Computational Acoustics*, edited by D. Lee, A. Cakmak, and R. Vichnevetsky (North-Holland, Amsterdam, 1990), Vol. 2, pp. 123–132.
- ¹⁰R. B. Evans, "A coupled mode solution for acoustic propagation in a waveguide with stepwise depth variations of a penetrable bottom," *J. Acoust. Soc. Am.* **74**, 188–195 (1983).
- ¹¹G. A. Athanassoulis and A. M. Prospathopoulos, "3D acoustic scattering of source-generated acoustic field by an axisymmetric island," in *Proceedings of the XIth Symposium on Hydroacoustics*, Gdynia-Jurata, Poland, 24–27 May 1994, edited by Z. Jagodzinski and R. Salamon (The Academy of Polish Navy, Gdynia-Jurata, Poland), pp. 75–82.
- ¹²M. I. Taroudakis, "A coupled-mode formulation for the solution of the Helmholtz equation in water in the presence of a conical sea-mount," *J. Comput. Acoust.* **4**, 101–121 (1996).
- ¹³R. B. Evans, personal communication (November 1997).
- ¹⁴J. A. Fawcett and T. W. Dawson, "Fourier synthesis of three-dimensional scattering in a two-dimensional oceanic waveguide using boundary integral equation methods," *J. Acoust. Soc. Am.* **104**, 3296–3304 (1998).
- ¹⁵J. A. Fawcett, "An efficient three-dimensional boundary integral equation method for solving azimuthally symmetric scattering problems in the oceanic waveguide," *J. Acoust. Soc. Am.* **94**, 2307–2314 (1993).
- ¹⁶J. A. Fawcett, "Coupled-mode modeling of acoustic scattering from three-dimensional, axisymmetric objects," *J. Acoust. Soc. Am.* **102**, 3387–3393 (1997).
- ¹⁷G. A. Athanassoulis and K. A. Belibassakis, "All-frequency normal-mode solution of the three-dimensional acoustic scattering field from a cylindrical obstacle in a waveguide," *J. Acoust. Soc. Am.* **101**, 3371–3384 (1997).

- ¹⁸G. A. Athanassoulis, A. M. Prospathopoulos, and K. A. Belibassakis, "A normal-mode solution for 3D acoustic scattering from a cylindrical island," in *Proceedings of the 3rd European Conference on Underwater Acoustics*, Iraklion, Crete, Hellas, 24–28 June 1996, edited by J. S. Papadakis (Crete University Press, Heraklion, Greece, 1996), pp. 273–278.
- ¹⁹E. Skudrzyk, *The Foundations of Acoustics, Basic Mathematics and Basic Acoustics* (Springer, New York, 1971).
- ²⁰M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions*, 9th ed. (Dover, New York, 1972).
- ²¹M. B. Porter and E. L. Reiss, "A numerical method for ocean acoustic normal modes," *J. Acoust. Soc. Am.* **76**, 244–252 (1984).
- ²²M. B. Porter, "The KRAKEN Normal Mode Program," Saclant Undersea Research Centre (1994).
- ²³S. J. Levinson, E. K. Westwood, R. A. Koch, S. K. Mitchell, and C. V. Sheppard, "An efficient and robust method for underwater acoustic normal-mode computations," *J. Acoust. Soc. Am.* **97**, 1576–1585 (1995).
- ²⁴G. H. Golub and C. F. Van Loan, *Matrix Computations*, 2nd ed. (Johns Hopkins University Press, Baltimore, 1989).
- ²⁵W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes in Fortran, The Art of Scientific Computing*, 2nd ed. (Cambridge University Press, Cambridge, 1992).
- ²⁶A. D. Pierce, *Acoustics: An Introduction to Its Physical Principles and Applications* (McGraw-Hill, New York, 1981).
- ²⁷C. S. Clay and H. Medwin, *Acoustical Oceanography: Principles and Applications* (Wiley, New York, 1977).
- ²⁸E. K. Westwood and R. A. Koch, "Elimination of branch cuts from the normal-mode solution using gradient half spaces," *J. Acoust. Soc. Am.* **106**, 2513–2523 (1999).
- ²⁹A. M. Prospathopoulos, "Three-dimensional scattering from axisymmetric obstacles in ocean waveguides," Ph.D thesis, Dept. of Naval Architecture and Marine Engineering, National Technical University of Athens (2000).

Time reversal in a waveguide: Study of the temporal and spatial focusing

Philippe Roux^{a)} and Mathias Fink

Laboratoire Ondes et Acoustique, ESPCI, Université Paris VII, CNRS UMR 7587, 10 rue Vauquelin, 75005 Paris, France

(Received 24 February 1999; accepted for publication 17 February 2000)

Temporal and spatial focusing properties of time-reversal mirrors (TRMs) are studied in a waveguide. The experiments are done using an ultrasonic TRM in an idealized waveguide. The width of the focal spot, and the spatial and temporal sidelobe levels are experimentally and numerically analyzed with respect to the characteristics of the waveguide-TRM system. An algorithm is developed to compute directly in the time domain the time-reversed field. This algorithm is based on the application of the mirror theorem to both the source and the TRM placed in the waveguide. Because time reversal is a stable and robust process, some of the ultrasonic results can be extended to ocean acoustics. Applications to underwater acoustic transmissions as well as ultrasonic medical imaging are discussed. © 2000 Acoustical Society of America.

[S0001-4966(00)04205-3]

PACS numbers: 43.30.Pc, 43.20.Gp, 43.35.Cg [DLB]

INTRODUCTION

Time-reversal invariance in acoustics means that for every burst of sound emitted from a source—and possibly reflected, refracted, or scattered by any propagation medium—there exists a set of waves that precisely retraces all these complex paths and converges at the original source, as if time were going backwards. Based on this concept, time-reversal cavities^{1,2} and time-reversal mirrors (TRMs)^{3–6} have been developed to focus ultrasonic pulses through strongly heterogeneous media such as multiscattering media,⁷ chaotic cavities,⁸ or waveguides.⁹ The advantage of time reversal, compared to any other technique, is that it provides a robust set of waves without any signal analysis, which compensates for the transfer function of the medium. Thus, TRMs lead to a spatial focusing and a temporal compression. Spatial focusing means that the time-reversed field focuses back exactly at the source. Temporal compression means that the time-reversed signal at the source is similar to the signal previously emitted by the source. Recently, we have shown that a TRM takes advantage of the heterogeneity of the medium to improve the focusing quality at the source.⁷ Following these results, the aim of this paper is twofold: on the one hand, we carefully study the spatial and temporal focusing properties of an ultrasonic TRM in a waveguide; on the other hand, we show that some of these results may be extended in ocean acoustics with some applications to acoustic transmission in an underwater acoustic channel. However, time reversal in underwater acoustics is not a new field. In the sixties, Parvulescu *et al.*^{10,11} performed experiments in shallow water at sea with one transducer working in a time-reversed mode. They observed temporal compression and they explained the major results by the method of images we will use in Sec. I. Nevertheless, their experiments did not embody the spatial focusing property of TRMs. More recently, Jackson

et al.^{12,13} and Dowling^{14,15} have developed the theory of phase conjugation (the Fourier conjugate of time reversal) in ocean acoustics. At the same time, Feuillade *et al.*¹⁶ computed numerical time-reversal experiments in shallow water where spatial focusing was studied. Finally, Kuperman *et al.*¹⁷ and Hodgkiss *et al.*¹⁸ demonstrated time-reversal experiments in an underwater acoustic channel. At a 445-Hz frequency, they used a 24-elements TRM to perform focusing and multipaths compensation up to 30 km. One of their important results was to show that time reversal is a very robust and stable process which is not very sensitive to noise or ocean variability. Thus, when performing a time-reversal experiment, the focusing in a “real” environment (the ocean) and an idealized system (an ultrasonic waveguide) are quite comparable. Working with ultrasound may then be a good way, particularly cheap and easy to set up, to study time-reversal properties in a waveguide very generally.

The paper is structured as follows: In Sec. I, we describe the experimental setup and the basic results obtained after time reversal in a waveguide. In Sec. II, we represent the TRM-waveguide system with its analog in free space and we deduce from this representation a simple way to compute, directly in the time domain, the acoustic focusing after time reversal. In Sec. III, we experimentally and numerically study the focal spot and the sidelobe levels as a function of the characteristics of the system: the length (L) and the height (H) of the guide, the wavelength (λ), and the number of transducers (N) of the TRM. In Sec. IV, we study the temporal compression after time reversal as a function of the same parameters. These results lead to the application of time reversal to acoustic transmission in an underwater acoustic channel as well as to medical applications in ultrasonics. Finally, we present our investigations of the influence of a nonstatic interface (such as waves at the surface of the ocean) on the quality of the time-reversal focusing.

^{a)}Electronic mail: Philippe.Roux@espci.fr

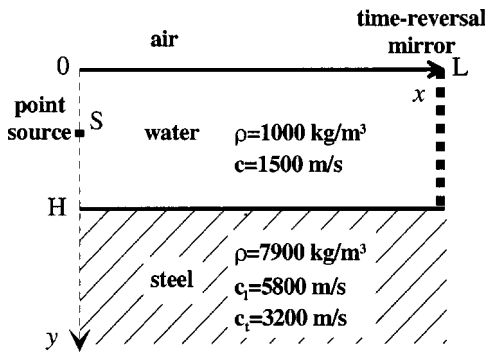


FIG. 1. Schematic of the acoustic waveguide: the guide length (L) ranges from 40 to 80 cm and the water depth (H) from 1 to 5 cm. The central acoustic wavelength (λ) is 0.5 mm. The array element spacing is 0.42 mm. The TRM is always centered at the middle of the water depth.

I. EXPERIMENTAL SETUP AND FIRST RESULTS

In this section, we describe first the basics of a time-reversal experiment and we show that the time-reversal focusing properties are greatly improved when the TRM is placed in a confined medium.

The setup for a time-reversal experiment in a waveguide is presented in Fig. 1. The water channel is bounded by two interfaces. A point-like ultrasonic source is located on one side of the waveguide and can be used either as a source or as a receiver. A time-reversal mirror is located at a distance L from the source. A complete time-reversal experiment is divided in three stages: first, the transducer source located at S transmits a pulse. In the second stage, the incident field is recorded, time-reversed, and transmitted back into the waveguide. In the last stage, the time-reversed field is measured in the source plane by the source transducer, which is used as a receiver. This mobile receiver can be translated around the initial position S along the y axis to measure step by step the time-reversed pressure field $p_{tr}(y, t)$. In this case we are interested both in the temporal compression and in the spatial focusing of the time-reversed field in the plane of the initial source. The quality of the temporal compression at the source can be deduced from the curve $p_{tr}(0, t)$ and the directivity pattern of the time-reversed beam is obtained from the curve $d(y) = \max_t \{p_{tr}(y, t)\}$.

The waveguide is made of two plane-parallel interfaces (one steel/water and one air/water interface). The length of the guide ranges from $L = 400$ mm to $L = 800$ mm, and the water depth from $H = 10$ mm to $H = 50$ mm. The TRM is a 128-transducer linear array identical to the source transducer. The width of an individual element is 0.39 mm, and the spacing between two elements is 0.42 mm. Thus, the maximal aperture of the array is 50 mm. The number of elements used during a time-reversal experiment is then adjusted to the water depth (for example, only 96 transducers are used for a 4-cm-depth waveguide). The mirror works at a 3-MHz central frequency with a 50% relative bandwidth at -6 dB (central wavelength = 0.5 mm). Each transducer element has its own amplifier, an 8-bit A/D converter, a storage memory, and an 8-bit D/A converter working at a 20-MHz sampling rate, which permit us to retransmit a time-reversed version of the recorded signals.

Several comments are to be made about the experimental setup:

- (i) The length of the source and the length of the array elements are $d = 12$ mm, which is significantly larger than the wavelength and the width of each element. However, these transducers cannot be considered as line sources because the propagation distance L is always much larger than the far-field distance R_{farfield} of each of them ($R_{\text{farfield}} \approx d^2/4\lambda = 72$ mm). Thus, the source and the TRM lie in each other's far field and the acoustic radiation is inversely proportional to the range, as for a point-like source.
- (ii) The steel density is 7.9 and the longitudinal and transversal wave speeds are $c_l = 5800$ m/s and $c_t = 3200$ m/s, respectively. With these parameters, no energy penetrates into the bottom for grazing angles smaller than $\theta_c = 65^\circ$ (grazing angles are measured with respect to a horizontal axis). This critical angle is much larger than the angular directivity of our ultrasonic transducers, which means that no energy is dissipated into the bottom.
- (iii) The TRM is always centered at the middle of the water depth, whatever the number of transducers used. However, the source transducer will not always be placed at the center of the waveguide.
- (iv) The sidewalls of the tank are at more than 30 cm from the xy propagation plane. Given the waveguide length, this means that the first echoes from the tank boundaries, if any, would arrive much later than the last recorded echoes reflected by the waveguide interfaces [see Fig. 2(a)].

Figure 2(a) shows the transmitted field recorded by the array after propagation through the waveguide. After arrival of a first wavefront corresponding to the direct path, we notice the arrival of a set of multipath signals corresponding to the multiple reflections of the incident wave on the interfaces. Figure 2(b) represents the signal recorded on one transducer of the TRM. After time reversal, Fig. 3(a) shows a remarkable temporal and spatial focusing of the time-reversed field at the initial source. Two major points are to be underlined: from the temporal point of view, multipath effects are compensated. The signal measured at S in the waveguide is nearly identical to the one you would find after a time-reversal experiment in free space [Fig. 3(b)]. This means that the transfer function of the waveguide has been completely compensated by the time-reversal process. Second, from the spatial point of view, the time-reversal focusing is much smaller than the one observed after a time-reversal experiment with the same TRM in free space (Fig. 4). Thus, time reversal takes advantage of the presence of interfaces in the medium to improve spatial focusing.

Following these first results, we study now the temporal and spatial properties of time-reversal focusing in a waveguide and some of its applications in ultrasonics and ocean acoustics.

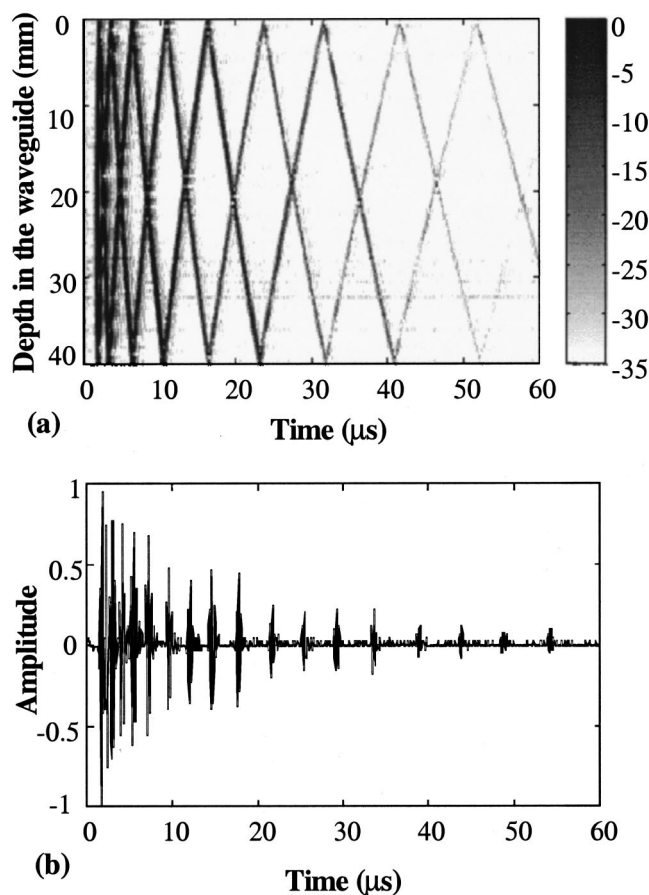


FIG. 2. (a) Spatial-temporal representation of the incident acoustic field received by the TRM after propagation through the guide; $H=40$ mm; $L=74$ cm; source depth=20 mm; (b) Temporal signal measured on one transducer of the array; receiver depth=10 mm.

II. FREE-SPACE REPRESENTATION OF TIME REVERSAL IN A WAVEGUIDE

The goal of this section is to give a free-space representation of the TRM-waveguide system and to deduce from this representation an algorithm to compute directly in the time domain the time-reversed field at the source.

We restrict our work to the study of a Pekeris waveguide.¹⁹ In the experimental configuration, λ is always much smaller than the characteristic lengths of the waveguide L and H . Ray theory can thus be applied to describe acoustic propagation. In this case, Snell's laws imply that each waveguide interface acts as a mirror for acoustic rays. The application of the method of images²⁰ allows us to transpose the forward and backward propagation in a waveguide between a point source in A and a TRM in B into a free-space configuration. To understand this transposition, we describe time reversal as a two-stage process: the forward and backward propagation.

In a first stage, the signal is emitted from A and is received on the TRM in B after propagation through the waveguide. The rays then appear to be emitted from a set of virtual sources in free space. These sources are the images of the actual source with respect to the guide interfaces (Fig. 5). The signal received on each individual transducer of the TRM is the superposition of the signal coming from each

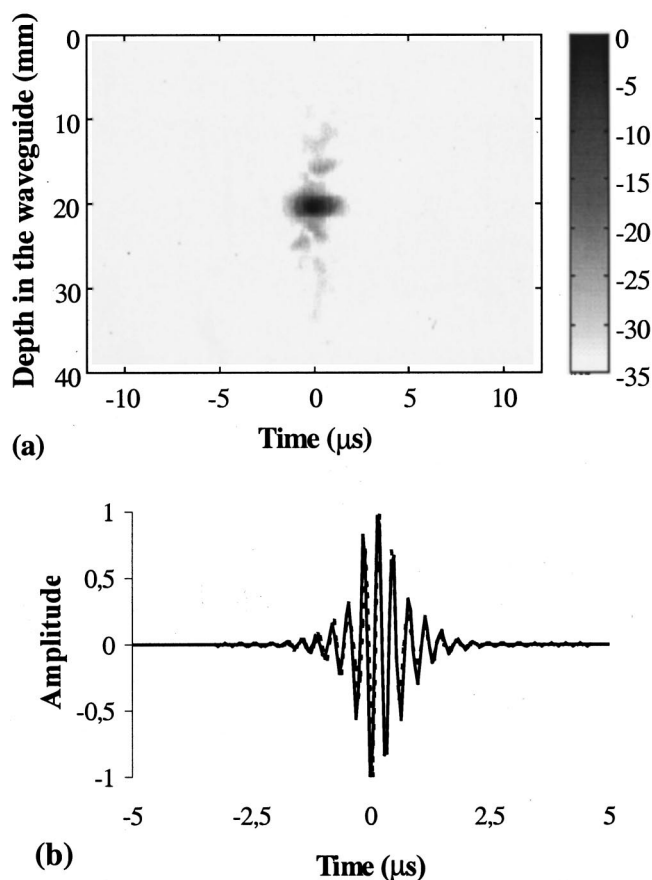


FIG. 3. (a) Spatial-temporal representation of the time-reversed field measured in the plane of the point source; $L=74$ cm; $H=40$ mm; source depth=20 mm; a 96-transducer TRM; (b) bold line: zoom of the time-reversed signal measured at the point source between -5 and $+5$ μ s (receiver depth=20 mm); dashed line: time-reversed signal measured at the source in free space.

image of the point source A . We experimentally confirm this result by recording the field on the TRM after propagation through the waveguide and by time-reversing this field in free space, meaning we remove the waveguide and transmit the time-reversed signal in a medium without interface. The time-reversed field then focuses at the source but also on the virtual images of the source with respect to the interfaces (Fig. 6). The directivity pattern is asymmetric because the

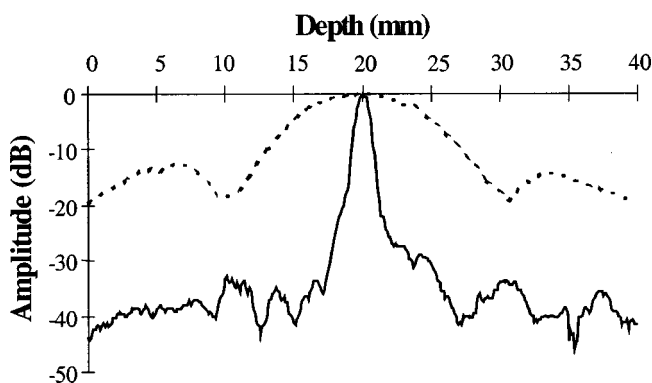


FIG. 4. Directivity pattern of the time-reversed field in the plane of source: the dotted line corresponds to free space, the full line to the waveguide; $H=40$ mm; $L=74$ cm; a 96-transducer TRM; source depth=20 mm.

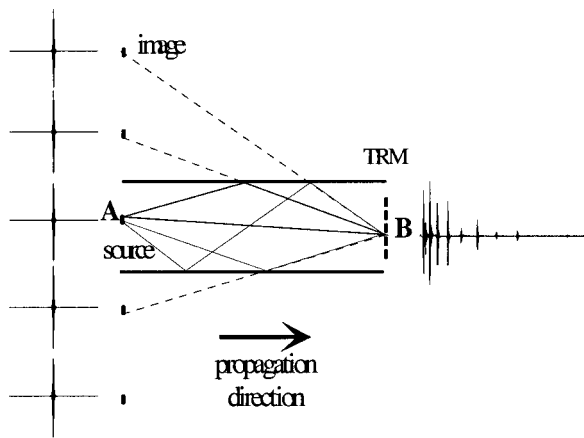


FIG. 5. Application of the method of images to the point source in the waveguide.

source is not centered at the middle of the water depth: for a source at depth y_0 , the images are located at $\pm y_0 \pm 2kH$, where k is an integer.

In a second stage, the time-reversal field is sent back from the TRM in B and we measure in A the time-reversed field after back propagation through the guide. From A , the TRM is now equivalent in free space to a set of virtual TRMs, which are the images of the actual TRM with respect to the wave guide interfaces (Fig. 7). This result has already been shown by the fact that the focal spot after time reversal in the waveguide is much smaller than the one obtained with the same TRM in free space (see Fig. 4). In other words, the smaller focal spot is due to an increase of the effective TRM aperture due to the presence of virtual TRM images in the waveguide.

Thus, the application of the method of images is two-fold; it increases the effective number of sources in the forward propagation stage and also the number of TRMs in the backward propagation stage. This allows us to describe the complicated TRM-waveguide system with a simpler free-space system created by a set of sources and mirrors. These sources are the images of the actual source with respect to the interfaces and are placed in front of a set of TRMs, which

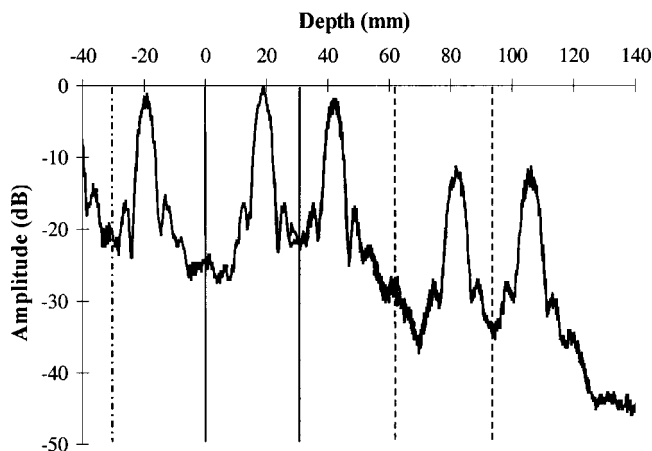


FIG. 6. Directivity pattern of the time-reversed field after back propagation in free space. The bold lines correspond to the waveguide actual interfaces and the dashed lines to the images of the interfaces; $L=33$ cm, $H=32$ mm, a 66-transducer TRM; source depth=19 mm.

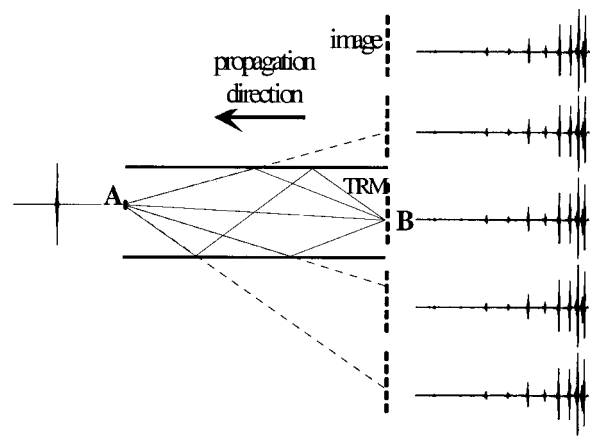


FIG. 7. Application of the method of images to the TRM in the waveguide.

are the images of the actual TRM with respect to the same interfaces (Fig. 8).

Theoretically the number of images is infinite. However, this number is experimentally limited by two different effects in ocean acoustics and in ultrasonics:

- (i) In ocean acoustics, the number of echoes observed in a waveguide is limited by the reflection coefficient at the bottom interface. In first approximation, this leads to an attenuation as a function of range equal to

$$R(\theta_{\text{grazing}}) = R_1^n(\theta_{\text{grazing}}) * (-1)^m, \quad (1)$$

where $R_1(\theta_{\text{grazing}})$ is the reflection coefficient measured at the water–bottom interface for a plane wave and a grazing angle θ_{grazing} , n corresponds to the number of reflections on this interface, and m the number of reflections on the water–air interface. For a given grazing angle θ_{grazing} , the larger the number of reflections (n), the longer the propagation range and the

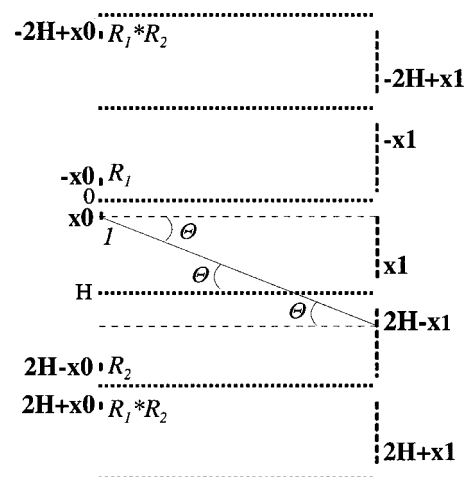


FIG. 8. Representation of the time-reversal–free-space system analog to the time-reversal–waveguide system: on the left are located the images of the actual source x_0 at $\pm x_0 \pm 2kH$; on the right are located the images of an actual TRM transducer x_1 at $\pm x_1 \pm 2pH$. The coefficients $R_1^n * R_2^m$ correspond to the attenuation of every source relative to its position. In the case of a water–air interface, we have $R_1 = -1$. A ray is plotted between the source x_0 and the receiver $2H - x_1$. The grazing angle at the interface Θ is equal to the launch and the reception angles.

stronger the attenuation. The reflection coefficient R may be complex for a reflection below the critical angle θ_c or because of absorption in the bottom.²⁰ Basically, this reflection coefficient implies that long-distance propagation in ocean waveguides occurs within a cone whose half-angle is equal to the critical angle θ_c .

- (ii) In ultrasonics, attenuation due to reflection is negligible in the case of a steel–water interface. However, the size of each transducer is no longer small compared to the acoustic wavelength. This leads to an angular directivity $\text{ang}(\theta_{\text{dir}})$ which affects the acoustic field both in transmission and in reception. For a launch or a reception angle θ_{dir} (with respect to the horizontal axis), the angular directivity is approximated by an attenuation coefficient equal in the far field to

$$\text{ang}(\theta_{\text{dir}}) \propto \frac{\sin\left(\frac{ka \sin(\theta_{\text{dir}})}{2}\right)}{\frac{ka \sin(\theta_{\text{dir}})}{2}}, \quad (2)$$

where k is the acoustic wave number at the central frequency of the transducer and the size of the transducer. Once again, the angular directivity limits long-distance propagation within a cone whose half-angle depends on the transducer characteristics. Note that in the case of a Pekeris waveguide, we have $\theta_{\text{dir}} = \theta_{\text{grazing}}$ (Fig. 8).

It is interesting to see that the propagation cone in our ultrasonic time-reversal experiments ($\sim 30^\circ$ at -10 dB) is similar to the one measured in underwater acoustics with classical bottom characteristics (typically $15\text{--}35^\circ$ for continental-shelf sediments). This is the reason why time-reversal provides similar results in ultrasonics and in underwater acoustics.

In the free-space visualization, each image of the source and the TRM is weighted according to its attenuation coefficient and the angular directivity (see Fig. 8). Thus, the images further away from the waveguide have a lower contribution to the field.

According to this representation, a time-reversal experiment consists first of transmitting a pulse by all of the sources and second, of time-reversing this field back by the combined TRMs. This leads to a simple algorithm, in the time domain, to compute the acoustic focusing after time reversal:

- (1) According to the waveguide characteristics, construct the free-space analog system with the images of the source and the TRM. For a source at coordinates $r_0 = (0, y_0)$, the images are at coordinates $r_k = (0, \pm y_0 + 2kH)$, where k is an integer. For a TRM made of N transducers at $r_i = (L, y_i)$, with $i \in [1, N]$, the TRM images are at $r_{p,i} = (L, \pm y_i + 2pH)$, where p is an integer (see Fig. 8).
- (2) Compute the Green's functions $G(r_{p,i}, t | r_k, 0)$ between the sources at r_k and time 0, and the TRMs at $r_{p,i}$ and time t for the forward propagation. Similarly, compute the Green's functions $G(r, t | r_{p,i}, 0)$ between the TRMs at

$r_{p,i}$ and each point $r = (0, y)$, with $y \in [0, H]$ in the plane of the source for the backward propagation. Do not forget the attenuation coefficient of each emitter and the directivity angle of each emitter and each receiver according to their positions.

- (3) Compute the time-reversed field at point r using the formula

$$S(r, t) = \sum_k \sum_p \sum_i G(r, t | r_{p,i}, 0) \otimes G(r_{p,i}, T - t | r_k, 0) \otimes f(T - t), \quad (3)$$

where $f(t)$ is the transmitted signal and T is such as $G \equiv 0$ for $t > T$. The first sum refers to the source images and the second to the TRM images.

In the case of a uniform sound-speed profile as in the Pekeris waveguide, the acoustic propagation between the sources and the receivers is described by the 3D free-space Green's function, which leads to the simple formula

$$G(r_{p,i}, t | r_k, 0) \approx R(\theta) \text{ang}^2(\theta) \frac{\delta(t - |r_{p,i} - r_k|/c)}{|r_{p,i} - r_k|} \quad (4)$$

with

$$\theta_{\text{dir}} = \theta_{\text{grazing}} = \theta = \cos^{-1} \left(\frac{|r_{p,i} - r_k|}{L} \right).$$

Note that the angular directivity is squared in Eq. (4) because the field is transmitted and received by transducers with the same characteristics.

If the sound speed is not uniform in the waveguide, then along with each source and TRM there exists a sound-speed profile symmetric with respect to the interfaces. In this case, a ray code must be used to compute the free-space Green's function taking into account acoustic refraction.²¹ Of course, time reversal and Eq. (3) remains valid but the new Green's function $G(r_{p,i}, t | r_k, 0)$ depends dramatically on the sound-speed profile and must be calculated numerically.

The biggest advantage of this algorithm is to work directly in the time domain and to use symmetry properties due to the waveguide geometry to compute the acoustic focusing after time reversal. Thus, in the case of a wideband source, the computational time is greatly reduced compared to the more classical monochromatic-modal formulation of acoustic propagation in a waveguide.

Furthermore, two points are to be underlined:

- (i) We have restricted our application of ray theory to a Pekeris waveguide. We have not investigated the classical ray theory problems encountered in a depth-dependent sound-speed profile, such as caustics, for example. Ray theory is a big topic, well-known and available in many references. The aim of this article is not to give a tutorial on this subject but to show that ray tracing is a convenient tool to study time reversal directly in the time domain. Furthermore, even if caustics may dramatically affect the forward propagation between the sources and the TRM by changing the energy distribution in the waveguide, their pres-

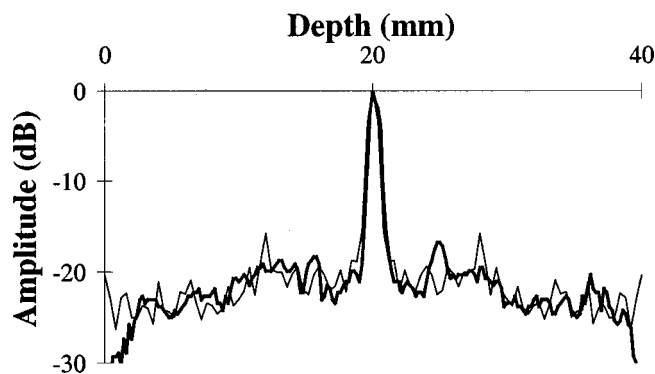


FIG. 9. Directivity pattern of the time-reversed field obtained experimentally (bold line) and numerically (thin line) for a ten-transducer TRM. The TRM spans the whole water depth; $H = 40$ mm; source depth = 20 mm; $L = 74$ cm.

ence does not break time-reversal symmetry and thus, they do not influence the way the time-reversed field focuses back at the source.

- (ii) We have described reflection at the water–bottom interface by the reflection coefficient calculated in the case of a monochromatic and plane incident wave. We have not considered the pulse distortion after reflection on this interface which occurs when the reflection coefficient is complex. One way to take care of the pulse distortion is to modify the impulse response (and so the Green’s function) as described by Cron *et al.*²² With a complex reflection coefficient

$$R(\theta) = R_{\text{real}}(\theta) + iR_{\text{imag}}(\theta), \quad (5)$$

the impulse response is no longer defined by $h(t) = R(\theta)\delta(t)$, but by the equation

$$h(t) = R_{\text{real}}(\theta)\delta(t) - R_{\text{imag}}(\theta)\frac{1}{\pi t}. \quad (6)$$

This induces a distortion of the wave, which leads to a spreading of the signal and to a decrease of its amplitude. In order to take care of this wave distortion in our free-space configuration, we use the modified 3D Green’s function

$$G(r_{p,i}, t | r_k, 0) \approx \frac{\text{ang}^2(\theta)}{|r_{p,i} - r_k|} \left(R_{\text{real}}(\theta) \delta\left(t - \frac{|r_{p,i} - r_k|}{c}\right) - R_{\text{imag}}(\theta) \frac{1}{\pi\left(t - \frac{|r_{p,i} - r_k|}{c}\right)} \right). \quad (7)$$

Figure 9 shows a comparison between an experimental and a numerical directivity pattern obtained after time reversal in an ultrasonic waveguide. The length of the guide is 74 cm, the water depth is 4 cm, and the source is placed at the center of the guide. The TRM is made of ten equally spaced transducers which spread over the whole water depth. The numerical result has been obtained using the algorithm described above and the 3D Green’s function as written in Eq. (7). Numerical and experimental results are in good agreement, confirming that a simple model is enough to describe

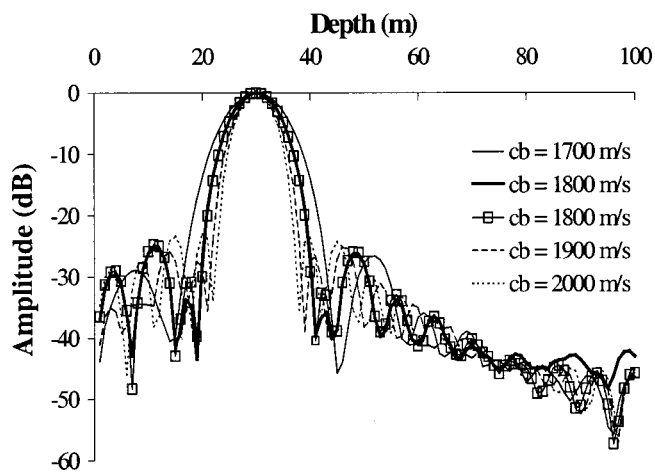


FIG. 10. Directivity pattern of the time-reversed field numerically computed in an underwater waveguide with different bottom velocities cb ; central frequency = 200 Hz, -6 -dB frequency bandwidth = 50 Hz, $L = 10$ km, $H = 100$ m, $\rho_b = 2000$ kg/m³, bottom attenuation = 0.2 dB/ λ , source depth = 30 m, a 20-transducer TRM. The two $c = 1800$ m/s curves correspond to two TRMs whose central position is shifted in depth by 1 wavelength. Since the frequency bandwidth is relatively small, the simulation is based here on normal-mode propagation.

the quality of focusing after time reversal. More particularly, it is obvious that the positions of the source and the TRM in the simulation are not exactly the same as in the experiment. Thus, the good agreement between experimental and numerical results shows that time-reversal focusing is not very sensitive to small errors in the transducer locations.

Indeed, time reversal is known to be a robust process in general and a stable process in the particular case of an acoustic waveguide. Time reversal is robust because it acts as a correlator¹¹ and is therefore not very sensitive to random noise. The stability of time reversal in a waveguide can be understood through a simple example: suppose the impulse response of the guide is made of ten echoes due to ten reflections on the interfaces. After propagation through the waveguide, the incident field depends on the interference between these ten echoes. It is well-known that the transmitted field is strongly affected by any mismatch of the waveguide characteristics or the receiver location.²³ Suppose now we time-reverse the field by a TRM made of 20 transducers. We are interested now in the time-reversed field in the plane of the source. Because of reciprocity, the impulse response between the TRM and the source is still made of ten echoes. This means that each TRM transducer time-reverses ten echoes and that each of these ten echoes generates ten echoes. For a 20-transducer TRM, the total number of echoes received after time reversal at each point in the plane of the source is then equal to $20 \times 10 \times 10 = 2000$. Once again, all these echoes interfere together and give rise to the spatial and temporal focusing observed in Fig. 2. But, because of the big number of echoes interfering together after time reversal, the time-reversed field appears not to be very sensitive to the phase and to the amplitude of each of them. This means that an approximated knowledge of the transducers location (Fig. 9) or of the bottom characteristics (Fig. 10) does not change time-reversal focusing, whereas it may dramatically affect the incident field. As a consequence, the advantage of the

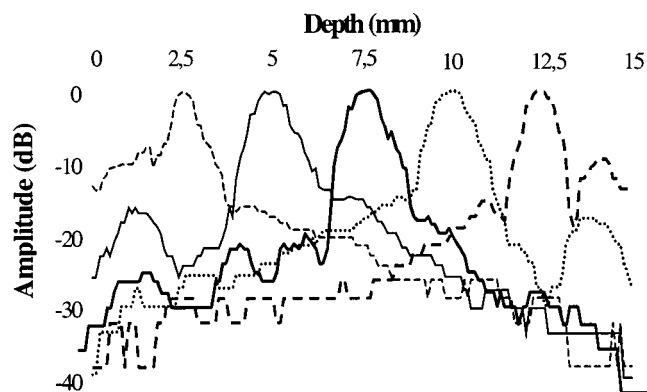


FIG. 11. Experimental directivity pattern of the time-reversed field for five positions of the source y_0 in the waveguide: $y_0 = 2.5$ mm, $y_0 = 5$ mm, $y_0 = 7.5$ mm, $y_0 = 10$ mm, $y_0 = 12.5$ mm; $L = 74$ cm; $H = 1.5$ cm; a 37-transducer TRM.

algorithm described above is that it combines a fast way to compute the time-reversed field and an acceptable accuracy even if the actual environment is roughly simplified.

III. SPATIAL FOCUSING AFTER TIME REVERSAL IN A WAVEGUIDE

The quality of the time-reversal focusing is determined by the characteristic parameters of the TRM-waveguide system. Recall these parameters are the length L , the height (H) of the guide, the wavelength (λ), and the number of transducers (N) of the TRM. In the following, we only work with $L/H \gg 1$ and $H/\lambda \gg 1$.

A. Position of the source

First, we remark that, besides these parameters, the position of the source greatly affects the sidelobe structure. Figure 11 represents five directivity patterns obtained in the same waveguide configuration, which show that the sidelobe levels are dramatically dependent on the depth of the initial source position.

In particular, sidelobes are more important when the source is close to the interfaces. This can be understood easily by the free-space visualization of time reversal in the waveguide. Indeed, the effective source can be considered in free space as the actual source *and* its images. Similarly, time reversal is performed now by the actual TRM *and* its images. This means that the time-reversed field will focus back on each source as seen in Fig. 6. Consequently, if the actual source is close to an interface, the sidelobes of the closest image spill over into the “real” waveguide, which explains the results of Fig. 11. For ease of study, we will now always place the source in the middle of the waveguide.

B. Focal spot

The directivity pattern obtained after time reversal in different waveguides is presented in Fig. 12. This figure shows that the focal spot does not depend on the waveguide parameters L or H . According to diffraction laws, the width (Δ) of the focal spot in free space depends on the length of the focusing array (D)

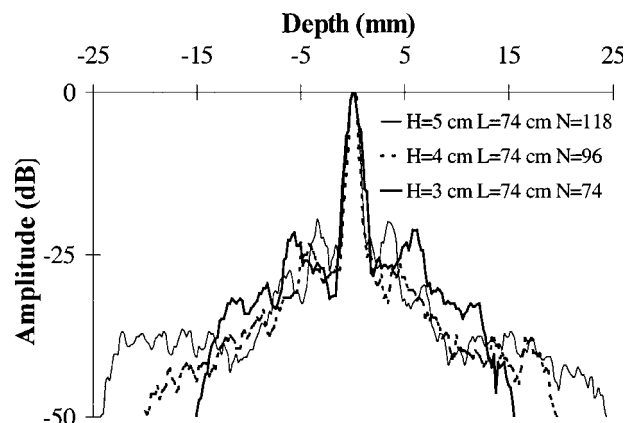


FIG. 12. Experimental directivity pattern of the time-reversed field in three different waveguides. The receiver depth is measured with respect to the source depth y_0 : $y_0 = 15$ mm for $H = 30$ mm; $y_0 = 20$ mm for $H = 40$ mm; $y_0 = 25$ mm for $H = 50$ mm. The TRM spans the whole water depth. The width of the focal spot does not change as a function of the waveguide depth.

$$\Delta \approx \frac{\lambda L}{D}. \quad (8)$$

In the waveguide, the focal spot is smaller because the reflections of the incident field on the interfaces virtually increase the size of the TRM. The distance D corresponds to the effective aperture of the TRM defined by its images. We have seen in Sec. II that the number of images is limited by two different effects in ultrasonics or in ocean acoustics: in ultrasonics, the field transmitted or received by each transducer is weighted by its angular directivity [Eq. (2)]. Neglecting the reflection coefficient at the steel–water interface, the effective aperture of the TRM is then determined by the product of the angular directivity of the source and the receiver. In our configuration, all the transducers are the same and we call θ_d the half-width directivity angle. Thus, the width of the focal spot is independent of the waveguide and is roughly equal to

$$\Delta \approx \frac{\lambda}{2 \tan(\theta_d)}. \quad (9)$$

In an underwater acoustic channel, propagation is also limited inside a cone defined by the critical angle θ_c . However, attenuation in the bottom decreases the contribution of the highest propagating mode as a function of range L . Thus, the width of the focal spot is slowly dependent of L as shown experimentally by Hodgkiss *et al.*¹⁸ This paper roughly defines the width of the focus as the ratio of the water depth to the number of contributing modes if the sound speed is not strongly dependent on depth. In analogy with Eq. (9), this is equivalent to

$$\Delta \approx \frac{\lambda}{2 \tan(\theta_m)}, \quad (10)$$

where θ_m is the grazing angle of the highest contributing mode in a waveguide of length L .

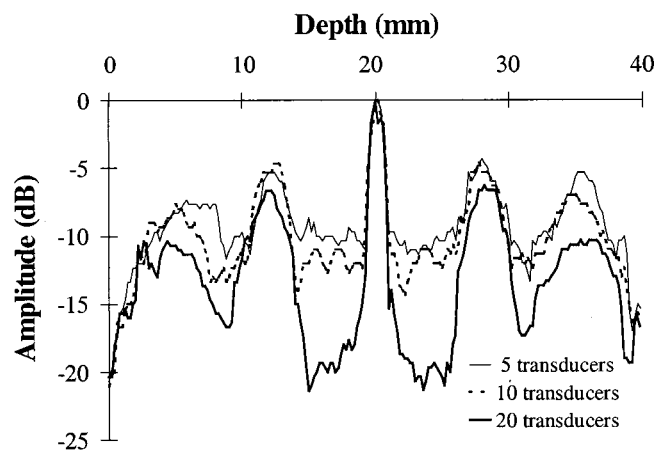


FIG. 13. Directivity pattern of the time-reversed field that puts into evidence sidelobes for small TRMs centered in the waveguide: 5 transducers correspond to a 2.5-mm-large TRM, 10 transducers to a 5-mm-large TRM, and 20 transducers to a 10-mm-large TRM; $H=40$ mm; $L=74$ cm; source depth=20 mm.

C. Sidelobes

In free space, sidelobes due to spatial aliasing are avoided when the spatial sampling of a horizontal array is greater than or equal to $\lambda/2$. The TRM used in this experiment is a vertical array with λ spacing. This is enough to neglect the sidelobes due to a poor spatial sampling because vertical wavelengths are usually much longer for acoustic waves propagating in a waveguide.¹⁸ However, Fig. 13 represents three directivity patterns obtained after time reversal in a waveguide which clearly show the presence of sidelobes due to spatial aliasing. These sidelobes appear when the TRM becomes small compared to the water depth H because of the new periodicity induced by the guide interfaces. We can visualize then a superarray made up of our small TRMs with spacing H (see Fig. 8). According to diffraction laws, this periodicity generates high sidelobes on each side of the main lobe at distance $k(\lambda L)/H$, where k is an integer. Experimentally, we have $(\lambda L)/H=8$ mm, which corresponds to the position of the sidelobes observed in Fig. 13. In all future measurements, the TRM will adequately span the water column.

We now try to quantitatively understand the dependence of the sidelobes on L , H , and N . The sidelobe amplitude is slowly varying versus water depth, which allows us to define the sidelobe level as its average amplitude around the focal spot. More precisely, we introduced in Sec. I the directivity pattern along the y axis by $d(y)=\max_t\{p_{tr}(y,t)\}$, where $p_{tr}(y,t)$ is the time-reversed beam in the source plane. We now define the sidelobe level by $\langle 20 \log_{10}[d(y)] \rangle$ for y varying over five wavelengths on each side of the focal spot. In other words, the sidelobe level is the average envelope expressed in dB around the focal spot. A heuristic study based on the free-space configuration has shown that one of the relevant parameters of the sidelobe study is the number NL/H .²⁴ We confirm this hypothesis by experimentally and numerically studying the sidelobe level as a function of the three parameters N , L , and H .

Figure 14 represents directivity patterns measured in a waveguide as a function of N . This shows that the sidelobe

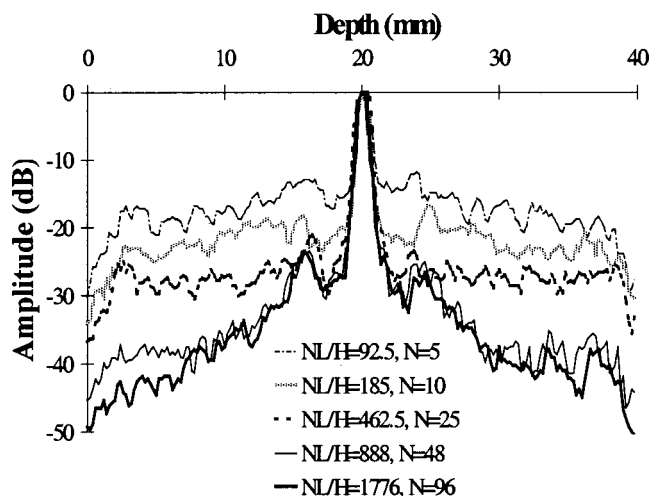


FIG. 14. Experimental directivity pattern of the time-reversed field versus NL/H ; the N transducers of the TRM span the whole water depth; $L=74$ cm; $H=40$ mm; source depth=20 mm.

level decreases as N increases down to a limit reached here when $N \sim 50$, or $NL/H \sim 1000$. Figure 15 represents the sidelobe level measured experimentally and computed numerically as a function of NL/H in different configurations. These results confirm the sidelobe behavior observed experimentally in Fig. 14: the sidelobe level roughly follows an inverse linear decrease as a function of NL/H before reaching a limit for $NL/H \sim 1000$. It is clear that the limit value of the parameter NL/H is not the same for ultrasonic experiments (where it probably depends on the transducer size) and in ocean acoustics (where it surely depends on the bottom property). Thus, the sidelobe level is a very important but very complicated issue. Here are two points that should be theoretically and experimentally investigated in the future:

- (i) In our experiments, a saturation of the sidelobe level is clearly observed for $NL/H \sim 1000$ when L/H or N are increased. Increasing N means more transducers

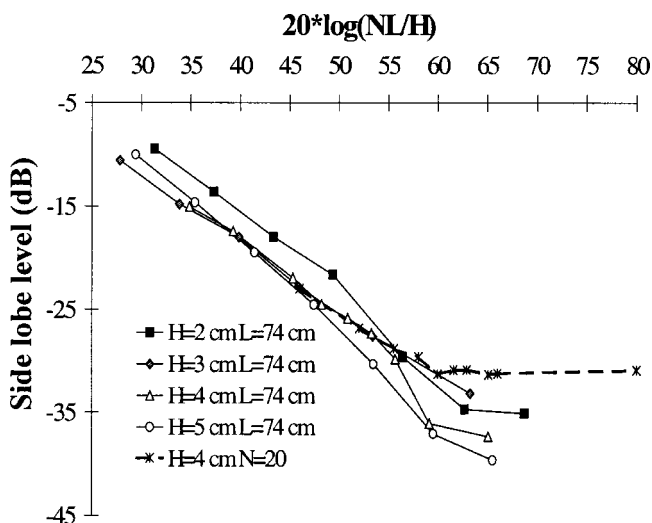


FIG. 15. Logarithmic decrease of the sidelobe level versus NL/H in different waveguides: the first four curves are experimental results as a function of the number of transducers N : the average slope coefficient is equal to -0.92 ; the dashed line corresponds to numerical results as a function of the waveguide length L and confirms the saturation of the sidelobe level.

and so more spatial information at a given range. Increasing L/H means more reflections on the waveguide interfaces and so more temporal information on each transducer. What is the coupling between spatial and temporal information on the sidelobe level after time reversal?

- (ii) What is the effect of the bandwidth of the emitted pulse on the sidelobe level? At first sight, a shorter pulse means a larger bandwidth and leads to lower sidelobes after time reversal.

IV. TEMPORAL COMPRESSION AFTER TIME REVERSAL IN A WAVEGUIDE

We have seen in the Introduction how multipaths were compensated after time reversal in a waveguide [Fig. 3(b)]. Indeed, the time-reversal process enables the realization of an “optimal” spatial and temporal filter matched to the waveguide transfer function. This is related to the reciprocity theorem, which states that the positions of a source and receiver can be interchanged without modifying the resulting field. Thus, if the waveguide transfer function is characterized by its impulse response, or equivalently Green’s function defined at point r_i and time t by $\tilde{G}(r_i, t | r_s, 0)$ from a source at point r_s and time 0, reciprocity means that

$$\tilde{G}(r_i, t | r_s, 0) = \tilde{G}(r_s, t | r_i, 0). \quad (11)$$

Under this assumption and assuming ideal transducers, the time-reversed field observed at the source location becomes

$$P_{tr}(r_s=0, t) = \sum_{i=1}^N \tilde{G}(r_i, t | r_s=0, 0) \otimes \tilde{G}(r_i, T-t | r_s=0, 0). \quad (12)$$

Note that the waveguide Green’s function \tilde{G} takes into account the reflection on the interfaces and is therefore more complicated than the free-space Green’s function G defined in Sec. I.

Each individual contribution i in Eq. (12) is the autocorrelation function of the impulse response between the point r_s and the transducer i at point r_i : it is a symmetrical signal with a maximum at the same time T for each transducer. However, for a given transducer i , the sidelobes of

$$\tilde{G}(r_i, t | r_s=0, 0) \otimes \tilde{G}(r_i, T-t | r_s=0, 0) \quad (13)$$

are located at different times that depend on the various multipaths between the source and the transducer. When we add a sufficient number of transducers, the sidelobes interfere destructively and all the maxima add constructively at $t = T$. Thus, if the TRM samples the whole water depth, the time-reversed field can be approximated by a Dirac δ -function (Fig. 16). The duration of the time-reversed signal at the initial source is then related only to the transfer function of the transducers, which is actually nonideal.

A. Acoustic transmission using time reversal

The temporal compression observed after time reversal in a waveguide has direct applications to acoustic transmission in underwater acoustics. The issue in underwater acous-

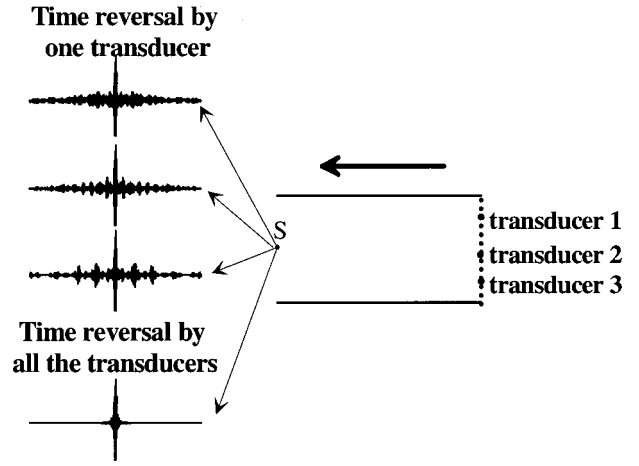


FIG. 16. Temporal compression after time reversal in a waveguide obtained separately for three different transducers and for the whole TRM.

tic transmission is to send a given signal through a shallow-water environment. If the emitted signal is long, the presence of interfaces leads to the appearance of multiple arrivals, which interfere and damage the quality of the acoustic transmission.

How can a TRM improve acoustic transmissions? Suppose that the goal of a TRM at B is to send a signal $f(t)$ to A through a waveguide (Fig. 1). First, A transmits a short pulse supposed to be a Dirac impulse $\delta(t)$. The TRM at B records the pressure field $h_i(t)$ with $t \in [0, T]$, which is the impulse response of the waveguide between the source and each transducer i of the TRM. After time reversal, we have shown that the time-reversed signal received in A is similar to the emitted Dirac impulse, which means

$$\sum_i h_i(t) \otimes h_i(T-t) = \delta(t). \quad (14)$$

As a matter of fact, if each transducer i of the TRM now transmits the signal $h_i(T-t) \otimes f(t)$, then the signal received at the source A is

$$\sum_i h_i(t) \otimes h_i(T-t) \otimes f(t) = f(t). \quad (15)$$

Following this procedure, Fig. 17 shows the example of the transmission through an ultrasonic waveguide of a train of five pulses. The pulses were separated in time by a time greater than the inverse of the transfer function bandwidth. The quality of the transmission depends dramatically on the number of transducers used on the TRM: the transmission is good for 48 transducers and the signal is unreadable if the time reversal is done with only one transducer. Moreover, we remark that point A is the only place in the waveguide to receive the transmitted information because the time-reversal field is only adapted spatially and temporally at this point. Acoustic transmission using time reversal is then not only accurate but also confidential because the transmitted information will remain obscure to anyone not close to point A .

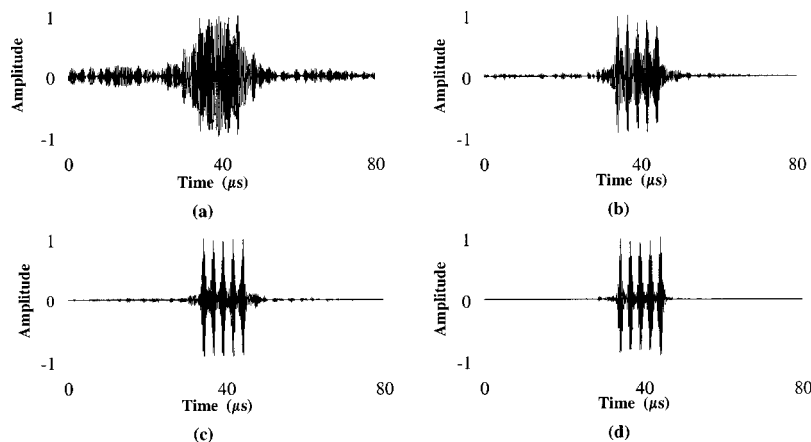


FIG. 17. Acoustic transmission of a train of five pulses using time reversal in a waveguide for a TRM made of (a) 1 transducer, (b) 8 transducers, (c) 16 transducers, and (d) 48 transducers; $L = 74$ cm; $H = 40$ mm; source depth = 20 mm; the TRM spans the whole water depth.

B. Amplification of the signal after time reversal in a waveguide

Another issue of temporal compression is the comparison between the amplitude of the time-reversed signal in the waveguide and in free-space configuration. The presence of interfaces in the guide leads to the generation of multiple echoes, which means that each transducer receives and then retransmits a much bigger energy than in a free-space medium where only one wavefront is expected. After time reversal, all the energy transmitted by the TRM is concentrated on the focal spot. The amplitude of the time-reversed field at the source is then much higher in a waveguide than the amplitude obtained with the same TRM in free space (Fig. 18).

Two points are to be highlighted: both in free space and a waveguide, the amplitude of the time-reversed field increases linearly with the number of transducers on the TRM, because each transducer receives and then retransmits approximately the same energy whatever its position in the waveguide. Generally, this is no longer true in ocean acoustics and, for example, in a strongly down-refracting sound-speed profile where incident energy is refracted to the bottom.^{18,25} Next, we see that the shallower the waveguide, the more echoes are received on the TRM and the higher the amplitude of the time-reversed signal at the initial source.

As a matter of fact, the temporal compression may be used to obtain a high-value pressure peak even with a low electrical power input applied to the transducers. Our group

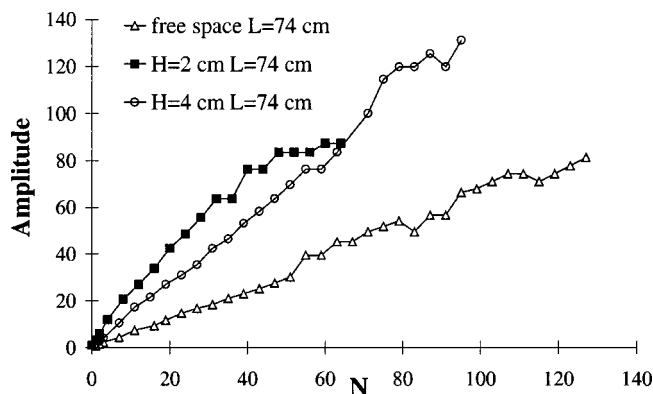


FIG. 18. Comparison of the maximum of the time-reversed signal measured at the source versus N in free space and in two different waveguides.

is now working on shock wave generators using this principle with a view to medical applications.

C. Time reversal in a nonstatic waveguide

We now study the influence of a nonstatic waveguide on the temporal compression observed after time reversal. The waveguide is 74 cm long and 4 cm deep. The 96-transducer TRM spans the whole water depth. Waves at the air–water interface of the guide are generated like waves at the sea surface to create a dynamic interface. The waves propagate from the source toward the TRM and their frequency of excitation is around 15 Hz, which corresponds to a 15-mm wavelength. The root-mean-square height of the waves is on the order of the acoustic wavelength (~ 0.5 mm).

First, we remark that the direct wavefront and the wave that reflects only once on the bottom are not affected by this disturbance. Because these echoes are generally the most energetic ones, waves at the sea surface have little effect on time reversal as predicted theoretically by Kuperman *et al.*¹⁷ On the other hand, as soon as the whole time-reversed field is affected by the disturbance, it may have a consequent influence on the time-reversal compression observed at the source. To explore the dynamic effects on time reversal, we select in our time-reversal window only the echoes which have been reflected at least once by the surface of the guide. Then, we perform time reversal in two configurations: an adaptive and a nonadaptive mode. In the adaptive mode, the time-reversed field is transmitted back into the waveguide as fast as allowed by the electronics of the TRM. In the nonadaptive mode, the time-reversed field is delayed before being retransmitted. The time delay is on the order of a few seconds, which is long compared to the dynamic time scale of the surface waves. Figure 19 represents the time-reversed field at the source for 100 shots in the adaptive and the nonadaptive mode.

In the nonadaptive mode, the medium has changed between the reception of the incident field and the transmission of the time-reversed field. The temporal compression at the initial source is then randomly affected by the surface. A similar nonadaptive time-reversal experiment has already been reported in Ref. 18 through an underwater acoustic channel where the disturbances were probably due to internal waves evolving on a few minutes' time scale.

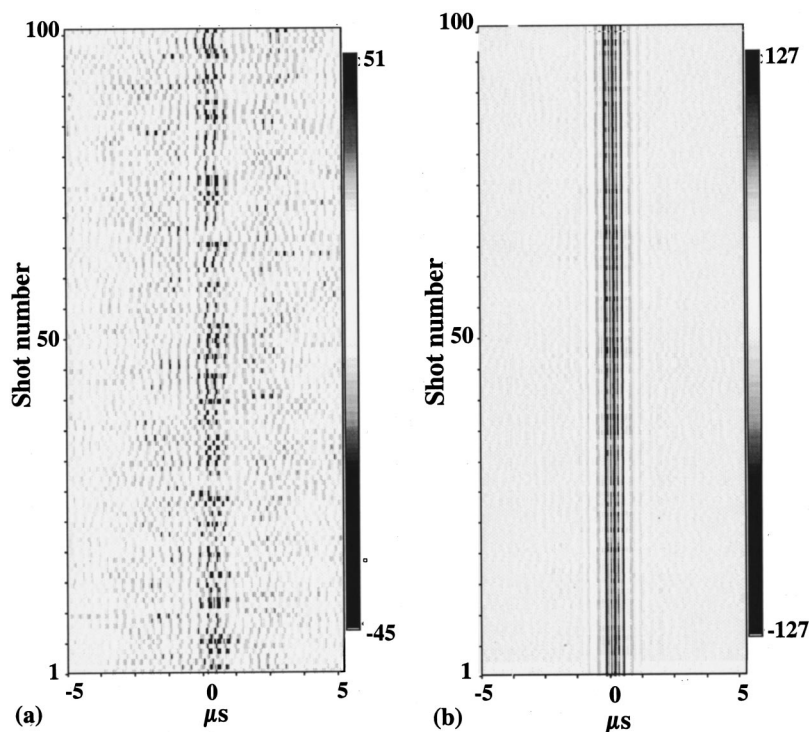


FIG. 19. Representation of 100 successive time-reversed signals measured at the source in an unstationary waveguide: (a) in the nonadaptive mode, (b) in the adaptive mode. The x axis corresponds to the time and the y axis to the shot number. The gray scale corresponds to the amplitude of the time-reversed signal. $L = 74$ cm; $H = 40$ mm; a 96-transducer TRM.

In the adaptive mode, the time-reversal process and the acoustic propagation is short enough (~ 2 ms) to consider the nonstatic surface as frozen during this time interval. In this case, the surface waves act as a rough static interface and the temporal compression after time reversal is similar to the one observed in a plane waveguide. In the ocean, acoustic propagation in shallow water lasts mostly between 1 and 20 s. Compared to the internal dynamic time scale of the ocean (around a few minutes), it should be possible to perform a good and stable time-reversal compression with a TRM working in the adaptive mode.

Finally, we quantify the effect of the nonstatic surface by a statistical study of the time-reversed amplitude at the source in the adaptive and nonadaptive mode (Fig. 20). As expected, we observe that the average amplitude in the nonadaptive mode is four times weaker than in the adaptive mode, which is itself close to the time-reversed amplitude obtained in a plane and static waveguide.

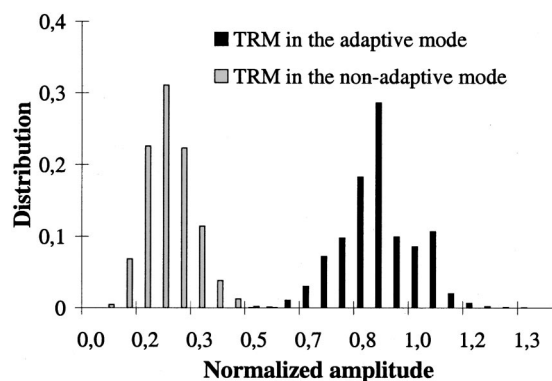


FIG. 20. Distribution of the amplitude maxima of 5000 time-reversed signals measured at the source in an unstationary waveguide. Amplitude 1 corresponds to the maximum of the time-reversed signal at the source in a static waveguide; $L = 74$ cm; $H = 40$ mm; a 96-transducer TRM.

V. CONCLUSION

We have presented in this paper experimental and numerical results on time-reversal properties in an ultrasonic waveguide. As long as the basic differences between ultrasonics and ocean acoustics are understood, some of these results are applicable to an underwater acoustic channel. We have shown in particular that the spatial sidelobe level around the focal spot depends on NL/H , where L and H are the length and the height of the guide and N the number of transducers of the TRM. On the other hand, the temporal compression observed after time reversal may lead to applications to acoustic transmission in ocean acoustics as well as high-intensity focusing in the medical field. Following some recent works in a wedge waveguide,²⁶ we plan to experimentally and numerically study time-reversal focusing in range-dependent waveguides in the future.

ACKNOWLEDGMENTS

The authors would like to thank Julien De Rosny for his useful contribution to the experiments and James H. Rose, Geoffrey Edelmann, and the referees for their wise comments on the manuscript.

¹D. Cassereau and M. Fink, "Time-reversal of ultrasonic fields. III. Theory of the closed time-reversal cavity," *IEEE Trans. Ultrason. Ferroelectr. Freq. Control* **39**, 579–592 (1992).

²D. Cassereau and M. Fink, "Focusing with plane time-reversal mirrors: An efficient alternative to closed cavities," *J. Acoust. Soc. Am.* **94**, 2373–2386 (1993).

³M. Fink, C. Prada, F. Wu, and D. Cassereau, "Self-focusing in inhomogeneous media with time-reversal acoustic mirrors," *Proc. IEEE Ultrason. Symp.* (2), 681–686 (1989).

⁴M. Fink, "Time-reversal of ultrasonics fields. I. Basic principles," *IEEE Trans. Ultrason. Ferroelectr. Freq. Control* **39**, 555–566 (1992).

⁵M. Fink, "Time-reversal in acoustics," *Contemp. Phys.* **37**(2), 95–109 (1996).

- ⁶M. Fink, "Time-reversed acoustics," *Phys. Today* **50**(3), 34–40 (1997).
- ⁷A. Derode, P. Roux, and M. Fink, "Robust acoustic time-reversal with high-order multiple scattering," *Phys. Rev. Lett.* **75**, 4206–4209 (1995).
- ⁸C. Draeger and M. Fink, "One-channel time-reversal of elastic waves in a chaotic 2D silicon cavity," *Phys. Rev. Lett.* **79**, 407–410 (1997).
- ⁹P. Roux, B. Roman, and M. Fink, "Time-reversal in an ultrasonic waveguide," *Appl. Phys. Lett.* **70**(14), 1811–1813 (1997).
- ¹⁰A. Parvulescu and C. S. Clay, "Reproducibility of signal transmissions in the ocean," *Radio Electron. Eng.* **29**, 223–229 (1965).
- ¹¹A. Parvulescu, "Matched-signal (Mess) processing by the ocean," *J. Acoust. Soc. Am.* **98**, 943–960 (1995).
- ¹²D. R. Jackson and D. R. Dowling, "Phase-conjugation in underwater acoustics," *J. Acoust. Soc. Am.* **89**, 171–181 (1991).
- ¹³D. R. Dowling and D. R. Jackson, "Narrow-band performance of phase-conjugate arrays in dynamic random media," *J. Acoust. Soc. Am.* **91**, 3257–3277 (1992).
- ¹⁴D. R. Dowling, "Phase-conjugate array focusing in a moving medium," *J. Acoust. Soc. Am.* **94**, 1716–1718 (1993).
- ¹⁵D. R. Dowling, "Acoustic pulse compression using passive phase-conjugate processing," *J. Acoust. Soc. Am.* **95**, 1450–1458 (1994).
- ¹⁶C. Feuillade and C. S. Clay, "Source imaging and sidelobe suppression using time-domain techniques in a shallow-water waveguide," *J. Acoust. Soc. Am.* **92**, 2165–2172 (1992).
- ¹⁷W. 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).
- ¹⁸W. S. Hodgkiss, H. C. Song, W. A. Kuperman, T. Akal, C. Ferla, and D. R. Jackson, "A long-range and variable focus phase conjugation experiment in shallow water," *J. Acoust. Soc. Am.* **105**, 1597–1604 (1998).
- ¹⁹C. L. Pekeris, "Theory of propagation of explosive sound in shallow water," *Mem.-Geol. Soc. Am.* **27** (1948).
- ²⁰L. M. Brekhovskikh and Y. P. Lysanov, *Fundamental of Ocean Acoustics*, 2nd ed. (Springer, New York, 1990).
- ²¹F. B. Jensen, W. A. Kuperman, M. B. Porter, and H. Schmidt, *Computational Ocean Acoustics* (Springer, New York, 1994).
- ²²B. F. Cron and A. H. Nuttall, "Phase distortion of a pulsed caused by bottom reflection," *J. Acoust. Soc. Am.* **37**, 486–492 (1965).
- ²³A. B. Baggeroer, W. A. Kuperman, and H. Schmidt, "Matched field processing: Source localization in correlated noise as an optimum parameter estimation problem," *J. Acoust. Soc. Am.* **83**, 571–587 (1988).
- ²⁴P. Roux, "The acoustic time reversal mirrors: Application to the focalization in a waveguide and to the characterization of hydrodynamic flows," Thesis (in French), Paris VI University, 1997.
- ²⁵H. 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).
- ²⁶C. Feuillade and C. S. Clay, "Broadband source imaging in a shallow water wedge by an array of receivers," *J. Acoust. Soc. Am.* **96**, 501–504 (1994).

Analysis of resolution for an amplitude-steered array

Catherine H. Frazier^{a)}

*Bioacoustics Research Laboratory, Department of Electrical and Computer Engineering,
University of Illinois at Urbana-Champaign, Urbana, Illinois 61801*

W. Jack Hughes^{b)}

*Applied Research Laboratory, The Pennsylvania State University, P.O. Box 30, State College,
Pennsylvania 16801*

William D. O'Brien, Jr.^{c)}

*Bioacoustics Research Laboratory, Department of Electrical and Computer Engineering,
University of Illinois at Urbana-Champaign, Urbana, Illinois 61801*

(Received 10 June 1999; accepted for publication 16 February 2000)

In 1976, Hughes and Thompson introduced the idea of steering the maximum response of a linear array by amplitude weighting the output signals of the elements, thus eliminating the need for time delays or phase-shift networks. Currently that amplitude-steered array concept is being extended to a broadband two-dimensional array that can be used for real-time three-dimensional imaging. In shifting the use of the amplitude-steered array from underwater acoustic communications to imaging, we must consider different issues of the array's performance such as lateral and axial resolution. For the linear amplitude-steered array, we show that both lateral and axial resolution are limited by the length of the array. The dependence of axial resolution on the length of the array is a unique feature of the amplitude-steered array, leading to an interesting tradeoff between lateral and axial resolution. A theoretical basis for the dependence is developed and simulation results are given. © 2000 Acoustical Society of America. [S0001-4966(00)03705-X]

PACS numbers: 43.30.Yj, 43.30.Wi [DLB]

INTRODUCTION

The concept of steering an array using amplitude weighting was introduced by Hughes and Thompson in 1976.¹ Although amplitude weighting is essentially an apodization, in this discussion the desired effect is to steer the main lobe, not to lower sidelobes. The intent of amplitude steering is to tilt the maximum response of the beam pattern without using multiple delay lines or phase-shift networks, which are bulky. Amplitude weighting is implemented passively, using fixed weights, in order to achieve this advantage in terms of cost and complexity.

In the original formulation of the amplitude-steered array, the beam was steered to a particular direction at a single frequency, and the fact that the steering direction changed with frequency was considered a drawback of the design. In this study, we use the change in steering direction with frequency to design a fast imaging system. By exciting the array with an impulsive or broadband linear frequency modulated (FM) chirp signal, the maximum array response is swept over a range of angles. A sector is scanned using a single broadband transmit pulse, leading to fast two-dimensional imaging of the sector, compared to conventional imaging that uses one transmit pulse for each steering direction. Even more exciting than two-dimensional imaging is the potential expansion to real-time three-dimensional imaging, which is currently being developed.

Figure 1 is a diagram of the imaging geometry. The steering direction of the main beam is defined by the angle θ measured from the array broadside direction. The axial and lateral directions are defined with reference to the steering direction. The axial direction is along the steering direction, and the lateral direction is perpendicular to it. A linear FM chirp with defined start and stop frequencies would sweep the main lobe over the sector defined in the figure by angled solid black lines.

Axial resolution is defined as the spatial length of the received pulse. The amplitude-steered array operated in broadband mode possesses a unique quality that axial resolution is dependent on the length of the array. In traditional imaging with an array of equiamplitude elements, where the transducer is shock-excited, axial resolution is independent of the length of the array, but dependent on the wavelength at the resonance frequency and on the Q of the transducer, where Q is defined as 2π times the energy stored at resonance divided by the energy dissipated per cycle.² For the amplitude-steered array, resolution in the lateral direction can be traded for resolution in the axial direction by changing the array length. It is the unique relationship between axial and lateral resolution of the amplitude-steered array that is explored in this paper.

Section I gives a brief overview of amplitude steering and image formation using this array. We recommend that readers refer to Ref. 1 for a more complete discussion of amplitude steering. Section II defines axial and lateral resolution for the purpose of this paper. Section III describes the tradeoff between axial and lateral resolution. In Sec. IV,

^{a)}Electronic mail: hillsley@uiuc.edu

^{b)}Electronic mail: wjh2@psu.edu

^{c)}Electronic mail: wdo@uiuc.edu

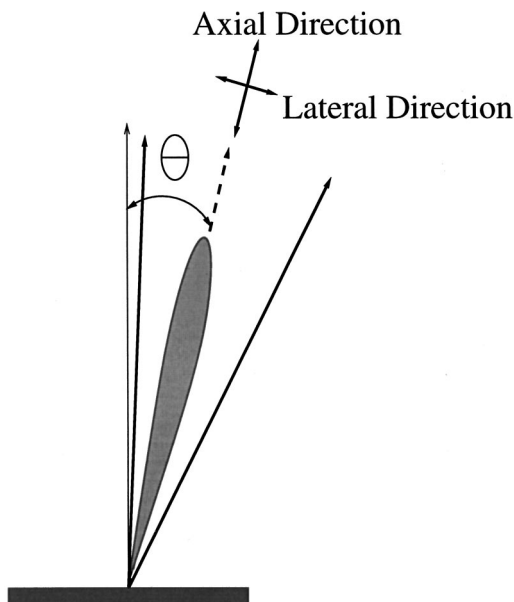


FIG. 1. Description of the imaging scenario. The steering direction of the main beam is given by the angle θ . Axial and lateral directions are defined relative to the steering direction.

simulation results are presented. And in Sec. V, conclusions are given.

I. AMPLITUDE STEERING AND IMAGE FORMATION

We will use the amplitude-steered array in pulse-echo mode, so each of the elements is both a transmitter and a receiver. In this discussion, we first consider the one-way far-field pressure of a linear array of equally spaced, equiamplitude point sources. To steer the main lobe to θ_0 , the signal from each element can be phase-shifted by multiplying by the factor $e^{-j(nk_0d \sin \theta_0)}$, where n is an index used to enumerate the elements, k_0 is the wave number at a particular frequency, and d is the distance between the elements. This phase shift will only steer the main lobe to θ_0 at the frequency used to calculate k_0 ; therefore, a fixed phase shift is generally used for narrow-band operation. The pressure field can be expressed as

$$P(r, \theta) = \frac{e^{-jkr}}{r} \sum_n e^{jnd(k \sin \theta - k_0 \sin \theta_0)} = \frac{Ne^{-jkr}}{r} H(\theta). \quad (1)$$

Now, we consider only the array pattern, $H(\theta)$. We assume that we have an array with an even number of elements, N , and we measure the phase for each element relative to the center of the array. To simplify the expression, we replace $(kd/2) \sin \theta$ with u and $(k_0d/2) \sin \theta_0$ with ϕ .

$$H(\theta) = \frac{1}{N} \sum_{n=-N/2+1}^{N/2} e^{j(2n-1)(u-\phi)}. \quad (2)$$

Combining terms of the sum in pairs, we have

$$H(\theta) = \frac{2}{N} [\cos(u-\phi) + \cos(3(u-\phi)) + \dots + \cos((N-1)(u-\phi))]. \quad (3)$$

Using the trigonometric identity for the cosine of a sum, Eq. (3) can be rewritten as

$$H(\theta) = \frac{2}{N} [\cos \phi \cos u + \cos 3\phi \cos 3u + \dots + \cos((N-1)\phi) \cos((N-1)u) + \sin \phi \sin u + \sin 3\phi \sin 3u + \dots + \sin((N-1)\phi) \sin((N-1)u)]. \quad (4)$$

A corresponding equation for an array with an odd number of elements is

$$H(\theta) = \frac{2}{N} [0.5 + \cos 2\phi \cos 2u + \cos 4\phi \cos 4u + \dots + \cos((N-1)\phi) \cos((N-1)u) + \sin 2\phi \sin 2u + \sin 4\phi \sin 4u + \dots + \sin((N-1)\phi) \sin((N-1)u)]. \quad (5)$$

Using Eq. (4), we shift our interpretation of how the beam steering is achieved. At the beginning of this discussion, we steered the beam by phase-shifting a linear array of equiamplitude elements. Now we steer the beam by weighting the elements. In Eq. (4), $\cos((2n-1)\phi)$ and $\sin((2n-1)\phi)$ terms are constants which we interpret as amplitude weights on the elements. The $\cos((2n-1)u)$ and $\sin((2n-1)u)$ terms represent combinations of pairs of elements on opposite sides of the center of the array in phase or 180° out of phase. Therefore, Eq. (4) can be interpreted as the sum of the outputs of two arrays. The first array [sum of $\cos((2n-1)u)$ terms] combines elements on opposite sides of the center with equal amplitudes and in phase. We call this the phase-symmetric array. The second array [sum of $\sin((2n-1)u)$ terms] combines elements on opposite sides of the array with equal amplitudes and 180° out of phase. We call this the phase-antisymmetric array. According to the equation, the outputs of the two arrays should be added in phase; however, the output of a phase-antisymmetric array is inherently 90° out of phase with the output of a phase-symmetric array. The outputs of the two arrays, therefore, must be added with an additional $\pm 90^\circ$ shift.³ The sign of the shift determines whether the beam is tilted toward positive θ_0 or toward negative θ_0 . Conceptually, we have used two arrays or one array with two sets of weights to achieve beamsteering. In practice, the array layout is designed such that the phase-symmetric and phase-antisymmetric arrays share the same space. The $\cos(2n-1)\phi$ and $\sin(2n-1)\phi$ factors in Eq. (4) are the amplitude weights which determine the steering direction of the main beam, but they do not have any effect on sidelobe levels. Further apodization can be applied to achieve reduced sidelobe levels.

If the array described above is excited by a frequency different from the design frequency used to calculate ϕ , the maximum response will occur at an angle different from the designed steering direction, θ_0 .

$$\phi = \frac{k_0d}{2} \sin \theta_0 = \frac{k_fd}{2} \sin \theta_f. \quad (6)$$

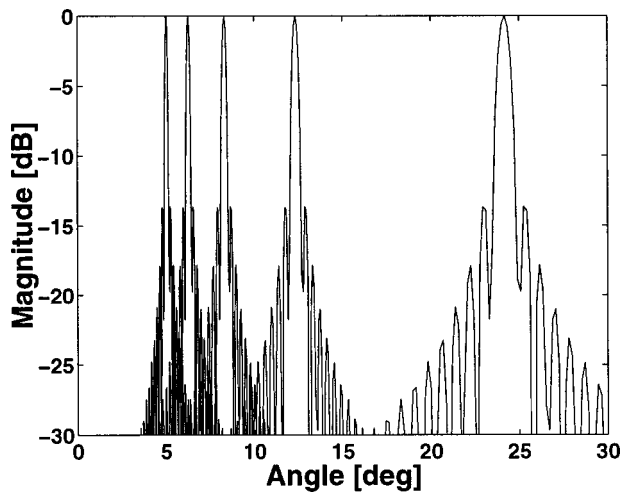


FIG. 2. Beams from a 9.76-cm-length aperture steered to 5° at 5.6 MHz. Beams are shown for 5.6 MHz (5°), 4.5 MHz (6.23°), 3.4 MHz (8.25°), 2.3 MHz (12.25°), and 1.2 MHz (24°).

The subscript, f , was added to k and θ to emphasize that the new steering direction is calculated for a specific frequency, f . Rearranging to solve for the new steering direction, θ_f , gives

$$\theta_f = \sin^{-1} \left(\frac{k_0}{k_f} \sin \theta_0 \right), \quad (7)$$

which can be restated as

$$\theta_f = \sin^{-1} \left(\frac{f_0}{f} \sin \theta_0 \right). \quad (8)$$

As the frequency increases, the angle that the beam is steered away from broadside decreases. An example of several beams from one array is shown in Fig. 2. The 9.76-cm-length array is designed to steer to 5° at 5.6 MHz. Beams are shown for 5.6 MHz (5°), 4.5 MHz (6.23°), 3.4 MHz (8.25°), 2.3 MHz (12.25°), and 1.2 MHz (24°). Note that as the frequency decreases, the beamwidths increase and the spacing between the beams also increases.

In Fig. 2, we have shown a discrete set of beams. By transmitting a linear FM chirp, the beam is swept continuously over a sector as the frequency changes. The array, operated in pulse-echo mode, transmits a pulse and receives the echo. The output of the array is a single radio frequency (rf) signal from which the image is made. To form a two-dimensional image, the received signal is first filtered with a matched filter for pulse-compression. Then the short-time Fourier transform (STFT) is calculated. The position of the fast Fourier transform (FFT) window gives range information, and the frequencies contained in the window give lateral position information. The image is the spectrogram, which is the magnitude squared of the STFT.

II. DEFINITION OF RESOLUTION

We will use one-point definitions for lateral and axial resolution, which correspond to measuring the axial and lateral extent of the point spread function. This definition is used in contrast to a two-point definition, which would specify the minimum separation between two points that al-

lows the points to be distinguished. For the amplitude-steered array of point elements operated at a particular frequency f , the array pattern can be written as

$$H(\theta) = \frac{1}{N} \frac{\sin((N/2)kd(\sin \theta - \sin \theta_f))}{\sin((1/2)kd(\sin \theta - \sin \theta_f))}. \quad (9)$$

We define lateral resolution as the -3 -dB width of the one-way array pattern. The -3 -dB points on either side of the maximum are the points where the argument of the periodic sinc is ± 1.3894 . We use the expression

$$\frac{N}{2} kd(\sin \theta_{-3 \text{ dB}} - \sin \theta_f) = \pm 1.3894. \quad (10)$$

Solving for $\theta_{-3 \text{ dB}}$ on each side of the maximum array response,

$$\begin{aligned} \theta_{-3 \text{ dB}}^+ &= \sin^{-1} \left(\sin \theta_f + \frac{0.4423\lambda}{Nd} \right), \\ \theta_{-3 \text{ dB}}^- &= \sin^{-1} \left(\sin \theta_f - \frac{0.4423\lambda}{Nd} \right), \\ \alpha &= \theta_{-3 \text{ dB}}^+ - \theta_{-3 \text{ dB}}^-, \end{aligned} \quad (11)$$

where α is the lateral resolution in degrees.

Axial resolution is defined as the spatial length of the pulse

$$AR = \frac{ct}{2}, \quad (12)$$

where c is the speed of sound in the medium and t is the temporal length of the pulse. For conventional imaging, where the transducer is shock-excited, this description of axial resolution can be related to an expression for axial resolution which depends on the wavelength at the resonance frequency and on the Q of the transducer, where Q is defined as 2π times the energy stored at resonance divided by the energy lost per cycle.² That expression is given by

$$AR = \frac{Q\lambda}{4}. \quad (13)$$

An alternate definition of Q , $Q = f_r / \Delta f$, where f_r is the resonance frequency and Δf is measured at half power points, is consistent with the previous definition of Q .⁴

We seek a definition of axial resolution that is similar in form to Eq. (13). However, our definition of resolution will not depend on the Q of the transducer, but rather on the inverse of the relative bandwidth of the received signal, $f/\Delta f$ which we call Q_{sig} . We transmit a chirp and then apply matched filtering for pulse compression. The filter operates on the time signal received by the array. The output of the pulse compression operation is approximately a sinc if the time-bandwidth product is large enough,⁵

$$\rho = \frac{\sin(2\pi\Delta f x/c)}{2\pi\Delta f x/c}, \quad (14)$$

where ρ is the correlation between the received signal and the impulse response of the filter, Δf is the bandwidth of the received chirp, c is the speed of sound, and x is the range which is calculated as $ct/2$ when the array is operated in

pulse-echo mode. We use the -3 -dB points of ρ to find the axial resolution. The function falls to -3 dB, relative to the maximum, when the argument is equal to ± 1.3894 . We write

$$\frac{2\pi f}{c} \frac{\Delta f}{f} x_{-3 \text{ dB}} = 1.3894, \quad (15)$$

and solve for $2x_{-3 \text{ dB}}$, where the factor of 2 accounts for the -3 -dB points on each side of the maximum. Frequency, f , has been multiplied in the numerator and denominator of Eq. (15) so that we may find an expression for axial resolution in terms of wavelength. The expression for axial resolution is

$$AR = 2x_{-3 \text{ dB}} = 0.4423 Q_{\text{sig}} \lambda, \quad (16)$$

where Q_{sig} is the inverse of the relative bandwidth of the received signal. Although this expression does not explicitly depend on the length of the array, we will see that Q_{sig} depends on the length through the beamwidth in the next section, where we discuss the tradeoff between axial and lateral resolution. We could have reduced Eq. (15) to an expression for $x_{-3 \text{ dB}}$ that was a function of Δf ; however, we prefer to have an expression in terms of the wavelength so we can see how resolution will change with center frequency, rather than with bandwidth. Note that for good resolution, we would like both $\theta_{-3 \text{ dB}}$ and $x_{-3 \text{ dB}}$ to be small.

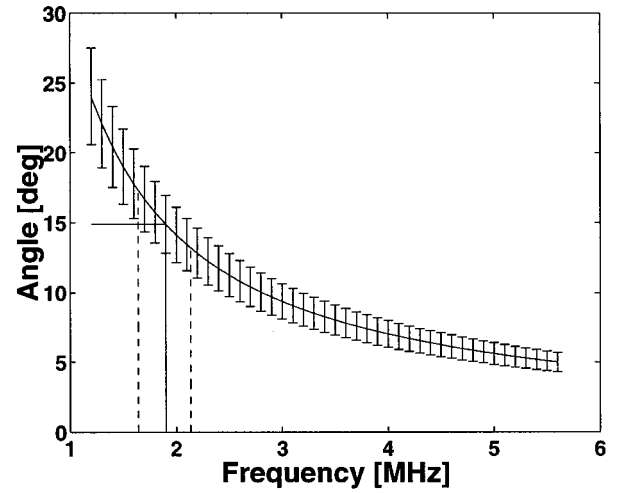
Finite element size will change the beamwidth, which affects Q_{sig} and therefore axial resolution; however, the effect is generally small for practical imaging arrays. For finite elements, the overall array pattern is multiplied by the beam pattern of the individual element, which can narrow the main beam, thus improving the lateral resolution and degrading the axial resolution. The array pattern for the array of point sources is given by Eq. (9). For an array of rectangular elements, the unsteered beam pattern is given by

$$H(\theta) = \frac{1}{N} \frac{\sin\left(\frac{Nkd}{2} \sin \theta\right)}{\sin\left(\frac{kd}{2} \sin \theta\right)} \frac{\sin\left(\frac{ka}{2} \sin \theta\right)}{\frac{ka}{2} \sin \theta}, \quad (17)$$

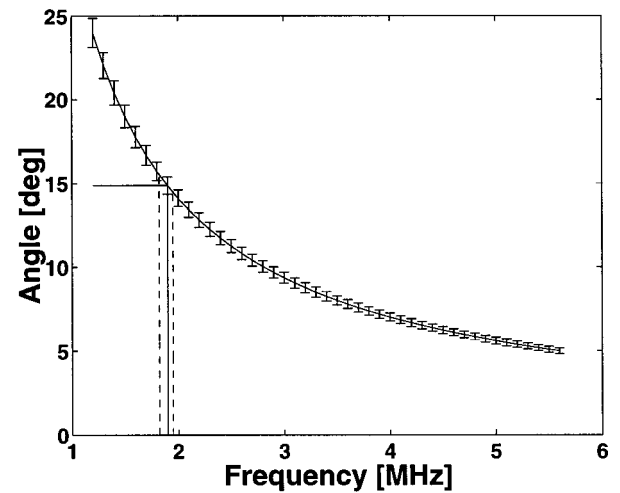
where a is the width of the element. The largest possible width of the element is the center-to-center spacing of the elements ($a = d$), which would mean that the aperture is continuous. If N is large, the expression for the width of the main beam will be dominated by the periodic sinc term, the array pattern. Likewise, if $a \ll d$, the expression for the width of the main beam will be dominated by the array pattern. In the other extreme, if $a = d$ and $N = 2$, the effect of the element size is to double the length of the array compared to the array of two-point elements. In that case and in similar cases where N is small and $a \approx d$, the finite element size will have a larger effect improving lateral resolution and degrading axial resolution.

III. AXIAL AND LATERAL RESOLUTION TRADEOFF

The amplitude-steered array spatially separates frequencies by virtue of the fixed phase shift used to calculate amplitude weights. If the frequencies could be completely separated, i.e., if the beams were infinitely narrow, a point target



(a)



(b)

FIG. 3. Lateral and axial resolution for (a) 1-cm-length and (b) 4-cm-length apertures. The line plotted is the steering direction versus frequency. The error bars indicate the -3 -dB beamwidth at each frequency.

within the insonified region would produce a single frequency return, implying that two-point targets separated only in range could not be distinguished. In reality, it is possible to distinguish multiple targets in the same direction at different ranges because beams are not infinitely thin, but overlap due to the finite length of the array.

Figure 3 conceptually shows the tradeoff between axial and lateral resolution. In parts (a) and (b), the curve plotted shows the steering direction versus frequency. The error bars indicate the -3 -dB beamwidth at each frequency. We can see from the error bars that lateral resolution improves with increasing frequency, as expected. Axial resolution can also be determined from these plots. At a particular steering direction, by observing the range of frequencies that overlap, we can determine the bandwidth at that particular direction. In both parts (a) and (b) of Fig. 3, the frequencies that overlap at a steering direction of 15° are located between the dashed lines. The extent of the dashed lines along the frequency axis tells us Δf . We know the “resonance fre-

quency,” the frequency with the greatest amplitude at 15°, from the curve. Therefore, we can determine the wavelength and Q_{sig} . Axial resolution, defined by Eq. (16), can be calculated. Q_{sig} is approximately constant over the range of frequencies shown; therefore, axial resolution also improves with increasing frequency.

Figure 3(a) shows field characteristics for a 1-cm-length array. If the length of the aperture is increased from 1 to 4 cm, the lateral resolution improves, illustrated by the decrease in size of the error bars between Fig. 3(a) and (b). However, that decrease in beamwidth implies a decrease in the range of frequencies that overlap in a particular direction, illustrated by the more narrow range between the dashed lines in Fig. 3(a) than in (b). Therefore, Q_{sig} increases with increased array length, and the axial resolution is worsened.

IV. SIMULATION AND RESULTS

We analyze the tradeoff between axial and lateral resolution for a linear array by simulating the received signal from point targets when the array is used in pulse-echo mode. Three arrays with different lengths are used. The first array has 452 elements with center-to-center spacing of 0.216 mm (9.76-cm-length aperture). For comparison, we also simulate arrays with 694 elements (15-cm-length aperture) and 347 elements (7.5-cm-length aperture), but otherwise similar designs. The amplitude weighting is determined so that the main beam is steered to 5° at 5.6 MHz. The transmitted signal is a linear FM chirp with frequency swept from 1.2 to 5.6 MHz. The targets are placed at 20 m, well beyond the intended maximum range, so that they are in the far field for all steering directions and all array lengths. The angular positions of the targets range from 6 to 24°. The speed of sound is assumed to be 1500 m/s for all simulations. Attenuation is not included. The transducer we simulate has a broadband, low- Q impulse response. The main effect of the transducer’s transfer function is to reduce the amplitude of targets away from the resonance frequency of the transducer, not to affect resolution. Therefore, in the simulations we replace the transducer’s impulse response with an impulse.

The operation of the linear amplitude steered array has been simulated using the Field II program, developed by J. A. Jensen.^{6,7} Figure 4 shows an image of six-point targets imaged using the 452-element array. We can immediately see that both axial and lateral resolution improve with decreasing steering direction, corresponding to increasing frequency.

In order to quantify the tradeoff between axial and lateral resolution, resolutions were measured from images of point targets. In addition to the resolution tradeoff that arises due to the length of the array, there is an additional tradeoff between axial and lateral resolution due to the processing. A STFT is used to form the images. The length of the sliding FFT window affects both lateral and axial resolution of targets. A very short FFT window implies poor frequency resolution and therefore poor lateral resolution, but it also implies good time localization and therefore good axial resolution. The FFT window can be increased to improve lateral resolution until the fundamental limit on lateral reso-

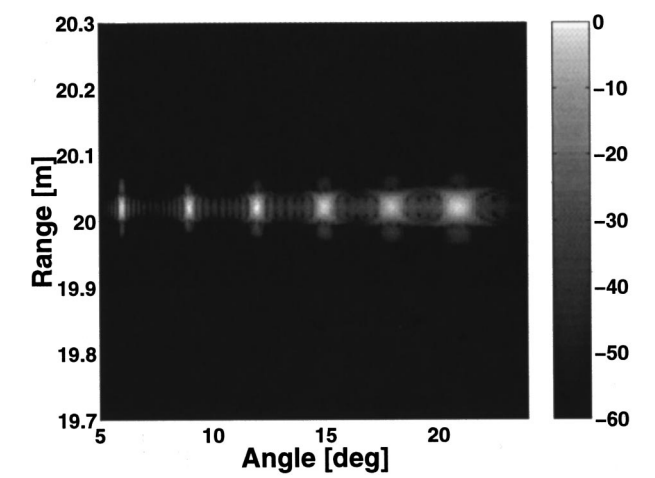


FIG. 4. Image of six simulated targets using 9.76-cm array steered to 5° at 5.6 MHz. The STFT calculation uses a 50.8-μs Hanning window.

lution due to the length of the array is reached, but a long window means poor axial resolution.

Lateral and axial resolutions were measured from images formed using different length FFT windows in order to reduce the effect of processing on the measurements, so that we could measure the tradeoff in resolution only due to the length of the array. Axial resolution was measured by setting the window length to be short, i.e., 8.333 μs corresponding to ten cycles of 1.2 MHz. This window length was considered long enough to define a signal at the lowest frequency of the chirp. Lateral resolution was measured by setting the window length to be long, i.e., 0.147 ms. This window length was chosen by increasing the length of the FFT window until the improvement in lateral resolution at 6° was less than 0.5%. The image in Fig. 4 is shown using a window length of 50.8 μs, which is a compromise between the two extremes.

Measurements were made of the −3-dB beamwidths in the lateral direction and the −3-dB signal length in the axial direction (Table I). For all steering directions, lateral resolution improves with increasing array length, and axial resolution degrades with increasing array length. Calculated and measured beamwidths and axial resolutions are shown in Figs. 5 and 6, respectively. Calculated beamwidths were

TABLE I. Axial and lateral resolution measurement results.

Angle (deg)	6	9	12	15	18	21
f (MHz)	4.669	3.120	2.347	1.886	1.579	1.362
λ (mm)	0.321	0.481	0.639	0.795	0.950	1.10
BW (deg)						
7.5 cm	0.211	0.331	0.404	0.573	0.722	0.808
10 cm	0.174	0.254	0.316	0.440	0.539	0.613
15 cm	0.119	0.174	0.214	0.289	0.342	0.421
Δf (kHz)						
7.5 cm	165	110	75	70	60	50
10 cm	130	85	60	55	45	40
15 cm	95	55	40	35	30	20
AR (mm)						
7.5 cm	3.845	7.690	8.972	10.25	11.54	15.38
10 cm	6.409	8.972	11.54	12.82	15.38	17.94
15 cm	8.972	12.82	16.67	19.23	24.35	26.92

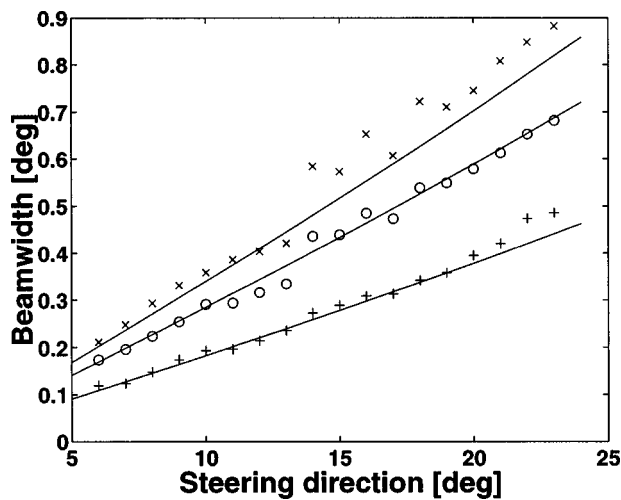


FIG. 5. Comparison of calculated and measured beamwidths. Measured values are given for 7.5-cm length array (\times), 9.76-cm length array (\circ), and 15-cm length array ($+$). For values measured from simulated images, the STFT calculation uses a 0.147-ms Hanning window. Calculated values are indicated by the solid lines.

found using Eq. (11). Figure 5 shows that beamwidths measured from simulated data agreed with calculated beamwidths to a steering direction of 14° . At higher steering angles, corresponding to lower frequencies, the measured beamwidths were larger than calculated beamwidths. This agreement may be improved by using an even longer FFT window. Axial resolution was calculated using Eq. (16), where Q_{sig} was calculated using the predicted beamwidths for each frequency and assuming that the array had a flat frequency response. If the steering direction was within the steered -3-dB beamwidths of a frequency, f , then f was included within Δf used to calculate Q_{sig} . Figure 6 shows that the calculated and measured axial resolutions agreed well. The average Q_{sig} predicted for the 7.5-, 9.76-, and 15-cm apertures are 27.60 ± 0.094 , 35.90 ± 0.021 , and 55.18 ± 0.054 . The average Q_{sig} measured for the 7.5-, 9.76-, and

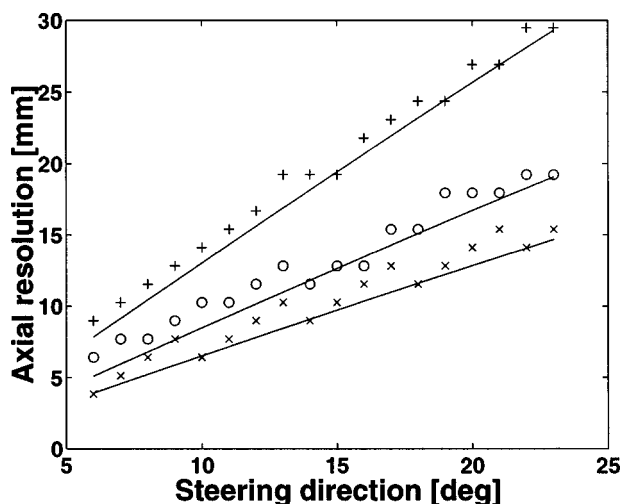


FIG. 6. Comparison of calculated and measured axial resolutions. Measured values are given for 7.5-cm length array (\times), 9.76-cm length array (\circ), and 15-cm length array ($+$). For values measured from simulated images, the STFT calculation uses an 8.333- μs Hanning window. Calculated values are indicated by the solid lines.

15-cm apertures are 28.5 ± 1.85 , 36.1 ± 2.01 , and 54.1 ± 5.45 , respectively.

V. DISCUSSION AND CONCLUSION

The amplitude-steered array is an array developed to decrease the complexity of electronics used to steer a beam. Amplitude weighting is implemented passively, using fixed weights, so that advantages in terms of cost and complexity over the electronically steered array can be achieved. The array can be used to do fast two-dimensional imaging by exciting it with a broad band of frequencies. When operated with a broadband pulse, the axial resolution is dependent on the length of the array relative to the wavelength, which is different from the operation of conventional transducers.

Fast two-dimensional imaging refers to collecting data for a two-dimensional image with one transmit pulse. The speed of data collection is compared to a conventional, electronically steered array, which transmits at least one pulse per steering direction or line in the image. For each pulse transmitted, the system waits for the pulse to be received from the farthest range of interest before the next pulse is transmitted. In order for the conventional array to collect all the image data with a single transmit pulse, the transmit beam must be broad, so that the entire region of interest is insonified. Then in receive mode, the signals received by the individual elements are stored and used to calculate beams in each of the directions of interest. This imaging method is as fast as ours; however it suffers from a poor signal-to-noise ratio and some loss of resolution because focusing is accomplished only on receive, not on transmit. Other researchers are working on different methods to form orthogonal beams, so that the transmit pattern has multiple main lobes and the targets can be separated with processing.⁸

With the amplitude-steered array, the entire region of interest is insonified by transmitting a broadband pulse. The beams in each direction can be focused by using a focusing lens in front of the array. The signal-to-noise ratio can be improved by lowering the chirp rate, that is, by transmitting a longer signal. Because of the pulse compression, axial resolution is not dependent on the time duration of the pulse, but rather on the bandwidth of the received pulse. The limiting factor in the length of the pulse transmitted is the depth of the "dead zone," the shallow region that cannot be imaged because the transducer is still transmitting; it cannot receive signals.

ACKNOWLEDGMENTS

This work was supported by DARPA under Contract No. BAA 97-33. The authors would like to thank the anonymous reviewers for their careful review and helpful comments.

¹W. J. Hughes and W. Thompson, Jr., "Tilted directional response patterns formed by amplitude weighting and a single 90° phase shift," *J. Acoust. Soc. Am.* **59**, 1040–1045 (1976).

²D. A. Christensen, *Ultrasonic Bioinstrumentation* (Wiley, New York, 1988).

³E. J. Skudrzyk, *The Foundations of Acoustics* (Springer-Verlag, New York, 1971).

- ⁴A. D. Pierce, *Acoustics: An Introduction to its Physical Principles and Applications* (Acoustical Society of America, Woodbury, New York, 1991).
- ⁵D. C. Munson, Jr. and R. L. Visentin, "A signal processing view of strip-mapping synthetic aperture radar," *IEEE Trans. Acoust., Speech, Signal Process.* **37**(12), 2131–2147 (1989).
- ⁶Program is available at <http://www.it.dtu.dk/~jaj/field/field.html>.
- ⁷J. A. Jensen, "FIELD: A program for simulating ultrasound systems," *Med. Biol. Eng. Comput.* **34**(1), Suppl. 1, 351–353 (1996).
- ⁸J. Shen and E. S. Ebbini, "A new coded-excitation ultrasound imaging system—Part I: Basic principles," *IEEE Trans. Ultrason. Ferroelectr. Freq. Control* **43**(1), 131–140 (1996).

A model for longitudinal and shear wave propagation in viscoelastic media

Thomas L. Szabo^{a)}

Agilent Technologies, 3000 Minuteman Road, Andover, Massachusetts 01810

Junru Wu^{b)}

Department of Physics, University of Vermont, Burlington, Vermont 05405

(Received 7 December 1998; accepted for publication 16 February 2000)

Relaxation models fail to predict and explain loss characteristics of many viscoelastic materials which follow a frequency power law. A model based on a time-domain statement of causality is presented that describes observed power-law behavior of many viscoelastic materials. A Hooke's law is derived from power-law loss characteristics; it reduces to the Hooke's law for the Voigt model for the specific case of quadratic frequency loss. Broadband loss and velocity data for both longitudinal and shear elastic types of waves agree well with predictions. These acoustic loss models are compared to theories for loss mechanisms in dielectrics based on isolated polar molecules and cooperative interactions. © 2000 Acoustical Society of America. [S0001-4966(00)04705-6]

PACS numbers: 43.35.Bf, 43.35.Cg, 43.35.Mr [HEB]

BACKGROUND

In many instances, loss data taken from infrasound to ultrasound frequencies (Bamber, 1986; White, 1965; Collins and Lee, 1956; Szabo, 1995) follow a simple frequency, f , power law with constants α_0 and α_1 ,

$$\alpha(f) = \alpha_0 + \alpha_1 |f|^y, \quad (1)$$

where y is a positive number usually less than 2 and most often, α_0 is set to zero. Absorption data for $\alpha(f)$ are shown in Fig. 1 for a wide range of viscoelastic materials and frequency ranges. The longitudinal wave loss of bovine liver (Polhammer, 1973), shown with an exponent $y=1.3$ for 1–100 MHz, is typical of materials that are in the commonly occurring range, $1 \leq y \leq 2$. Human myocardium (Kudo *et al.*, 1997), once thought to follow a Voigt model (Ahuja, 1979), also falls in this range for y at ultrasound frequencies, as do many other tissues (Goss *et al.*, 1979; Bamber, 1986; Duck, 1990). At lower frequencies, many plastics, including data given later in this paper, aromatic polyurethanes (Guess and Cambell, 1995), and noncrosslinked polymers (Bagley and Torvik, 1983), fall into this range of exponent y . At very high ultrasound frequencies, single crystalline materials such as yttrium indium garnet (YIG), follow a quadratic power law with shear wave loss less than that of the longitudinal wave. Perhaps the most striking examples of the absorption following a frequency power law are those in seismology literature. White (1965) summarizes absorption data that as linear with frequency for six decades of frequency. Data extracted from his summary for two types of granite in Fig. 1 indicate this linear dependence from 140 Hz to 2.2 MHz. Note here shear wave loss is greater than that of a longitudinal wave, a more typical result.

Effects of absorption are usually described by a material transfer function in the frequency domain of the form, $\exp\{-\alpha(f) + i\beta(f)z\}$, where z is the distance of propagation from a source. Usually β is taken to mean $\beta = 2\pi f/V_0$, where V_0 is a constant speed of sound in the medium. In order to completely describe the effects of absorption in a causal way, velocity dispersion, $V(f)$, must be included; this means that $\beta(f)$ is a more complicated function of frequency. We present methods for determining $\beta(f)$ from data to complete the description for absorption of the power law type in Sec. I.

The loss characteristics of Fig. 1 are not predicted by a classic relaxation theory such as the Voigt dashpot model (Herzfeld and Litovotz, 1959; Kinsler *et al.*, 1982; Bamber, 1986; Auld, 1990) which has the well-known attenuation and velocity characteristics illustrated in Fig. 2. The basis of this model is a modified Hooke's law which has an added viscous term relating stress T (a second rank tensor) to the time derivative of strain S (also a second rank tensor),

$$T = c:S + \eta: \frac{\partial S}{\partial t}, \quad (2)$$

in dyadic notation (see the Appendix), where c is an elastic stiffness constant fourth rank tensor and η is a viscosity constant fourth-rank tensor.

Many attempts have been made to make relaxation models fit the types of data illustrated by Fig. 1. When no resonance is visible in absorption versus frequency data, an assumption is often made that a resonance occurs at a very high frequency beyond the measurement range, so that a low frequency approximation can be applied. From Fig. 2(a), at low frequencies, the absorption α has a frequency-squared dependence for absorption. For these frequencies, the speed of sound is nearly constant at low frequencies [in Fig. 2(b), the propagation factor divided frequency, β/ω is inverse velocity]. These characteristics do not match the type of data in Fig. 1.

^{a)}Electronic mail: tom_szabo@agilent.com

^{b)}Electronic mail: jwu@zoo.uvm.edu

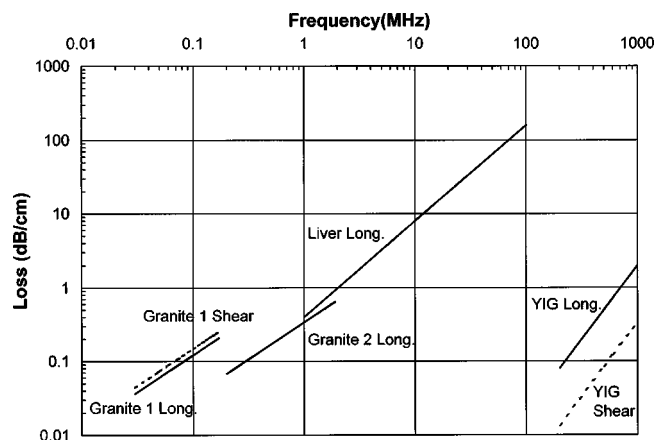


FIG. 1. Data for shear and longitudinal wave loss which show power-law dependence over four decades of frequency.

Others have suggested that, over a specified range of frequencies, a combination of relaxation functions can be used to approximate a power law. White (1965) discusses the attempts made to match data. Postma (1958) indicated such approximations were possible, and Horton (1959) achieved a reasonable fit to the linear loss of Pierre shale over a range of frequencies by using four relaxation constants. In principle, data can be fitted by a sum of relaxation constants over a frequency range as described by Tschoegel (1989). The application of multiple relaxation constants to match a desired power-law variation over a restricted frequency range will result in a time response consisting of a superposition of exponentially decaying relaxation constants, quite unlike the time response obtained for the power-law case in Secs. II and III. Models have also been proposed involving either a distribution or continuum of relaxation constants (Bamber,

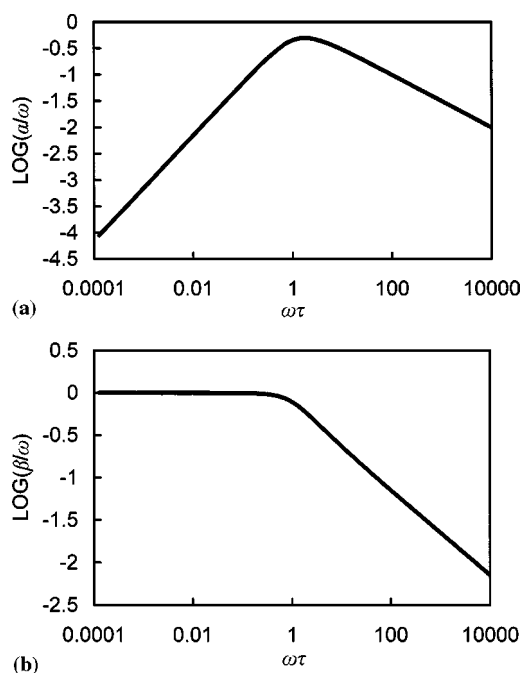


FIG. 2. (a) Absorption α divided by angular frequency ω vs angular frequency ω times τ for the Voigt dashpot model. Loss peak at $\omega_p = 1/\tau$, where τ is relaxation time. (b) Excess dispersion β divided by angular frequency ω vs angular frequency ω times τ .

1986; Aki and Richards, 1980; Nachman *et al.*, 1990; Szabo, 1993).

Ferry (1980, p. 59) wrote that “although an arbitrary set of parameters would suffice to predict macroscopic behavior, it would not be unique and would be of little value for theoretical interpretation. This difficulty can be avoided, however, by substituting continuous spectra.” White (1965) remarked that it is doubtful that these approaches would match linear frequency dependence of loss data observed over six decades of frequency. These attempts to fit data mask the possibility that these losses could depend on fundamental natural mechanisms other than relaxation phenomena. This topic is discussed in Sec. IV.

INTRODUCTION

Is there another paradigm better matched to observed losses in viscoelastic media than relaxation phenomena? This study extends previous work for fluids (Szabo, 1993, 1994) to describe losses in viscoelastic media based on a frequency power law, Eq. (1), such as all the cases depicted in Fig. 1. Elastic wave equations describe lossless propagation. New wave equations which contain an extra loss term appropriate for viscoelastic media are derived in the Appendix.

The model developed here is based on the assumption that the loss per wavelength is small, about $\alpha/\lambda \leq 0.6$, as explained in Sec. I and the end of the Appendix. A consequence of the smallness approximation is that the model becomes slightly less accurate at extremely high frequencies. For example, for loss of longitudinal YIG of Fig. 1, a conservative estimate of a high-frequency limit is 3.8×10^{12} Hz, which is well beyond the practical measurement range for this material. As the wavelength approaches the scale of material structure at these extremely high frequencies, other loss mechanisms (not in this theory) are expected to dominate.

Power-law data, as indicated by Fig. 1 and by data for liquids (Szabo, 1995) tend to preserve their unique frequency dependence over an extremely wide range. This model holds for any reasonable positive value of the exponent y and has been shown to agree well with data with values from $y=0.5$ to 2 (Szabo, 1995; He, 1998a, 1998b, 1999).

A consequence of this model is a more comprehensive Hooke's law in which the relation between stress and strain includes a convolution operator $r(S,t)$ which replaces the derivative term in Eq. (2)

$$T = c:S + \eta:r(S,t). \quad (3)$$

A direct relationship between this law and absorption and dispersion is part of the derivation in the Appendix. The Hooke's laws in Eqs. (2) and (3) are illustrated symbolically in Fig. 3. Specific forms of these laws are in Sec. II. There, we shall show that Hooke's law from the Voigt model is a special instance of Eq. (3).

As mentioned earlier, phase velocity dispersion is a causal by-product of absorption. General relations between absorption and velocity dispersion for viscoelastic media (Sec. I) apply even when the underlying physical causes of loss are unknown and the loss characteristic and/or velocity dispersion are acquired through measurement. These rela-

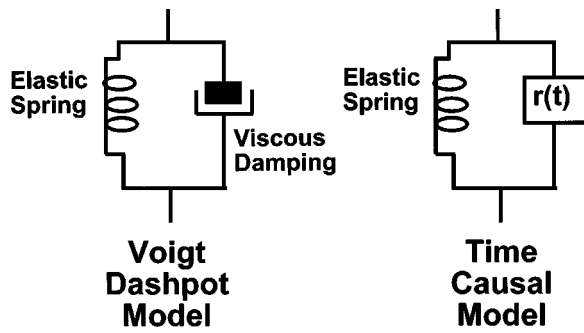


FIG. 3. Left: the schematic representation of Hooke's law for Voigt dashpot model. Right: representation of time causal model with function $r(t)$ replacing the dashpot as the viscous element.

tions will be shown to agree well with both longitudinal and shear wave loss and dispersion data in Sec. III.

I. ABSORPTION AND VELOCITY DISPERSION

For a medium with absorption losses, the speed of sound is dispersive in the most general case. A material transfer function (MTF) for a viscoelastic medium is

$$MTF(\omega) = \exp[\gamma_M(\omega)z], \quad (4)$$

where a complex wave number or propagation factor γ can be written as a function of ω , angular frequency, $\omega = 2\pi f$, for each elastic mode M (here, $M=S$ denotes a shear mode and $M=L$, longitudinal mode),

$$\gamma_M(\omega) = -\alpha_M(\omega) + i\beta_M(\omega), \quad (5a)$$

$$\gamma_M(\omega) = -\alpha_M(\omega) + i[\beta_{OM}(\omega) + \beta_{EM}(\omega)], \quad (5b)$$

where V_{OM} is a constant speed of sound for each mode, $\beta_{OM} = \omega/V_{OM}$. $\beta_{EM}(\omega)$ is an excess dispersion term needed to satisfy causality so that for each mode, the overall velocity varies with frequency as

$$V_M(\omega) = \omega/\beta_M(\omega) = [1/V_{OM} + \beta_{EM}(\omega)/\omega]^{-1}. \quad (6)$$

The absorption power law is recast for viscoelastic modes,

$$\alpha_M = \alpha_{OM} + \alpha_{1M}|\omega|^y, \quad (7)$$

where α_{OM} is a constant and

$$\alpha_{1M} = \frac{\eta_{IJ}}{2c_{IJ}V_{OM}}, \quad (8)$$

in which, in terms of the abbreviated subscripts of the Appendix, $M=S$, $IJ=44$ for the shear case; and for the longitudinal case, $M=L$, and $IJ=11$. For example, as shown in the Appendix, if $y=2$, there is no dispersion, or $\beta_{EM}=0$, and $\alpha_M = \alpha_{1M}\omega^2$. For loss with values of y which are not even integers, $\beta_{EM}(\omega) \neq 0$.

These results provide the structure needed for determination of wave equations and solutions for the general power-law case of attenuation, Eq. (7), where y can be any positive, finite (reasonably valued) value and for the derivation of a more comprehensive form of Hooke's law, Eq. (3). The necessary derivations are outlined in the Appendix and they draw on a time-domain statement of causality used in an earlier work on losses in fluids, Szabo (1994) and a smallness approximation (see the end of the Appendix). Dispersion is found to be of the following forms:

for y as an even integer or noninteger,

$$\beta_{EM}(\omega) = \alpha_{1M} \tan(\pi y/2) \omega |\omega|^{y-1}, \quad (9)$$

and for y as an odd integer,

$$\beta_{EM}(\omega) = -(2/\pi) \alpha_{1M} \omega^y \ln|\omega|. \quad (10)$$

Versions of these equations more appropriate for velocity data are the following:

For y as an even integer or noninteger,

$$1/V_M(\omega) = 1/V_M(\omega_0) - \alpha_{1M} \tan(\pi y/2) [|\omega|^{y-1} - |\omega_0|^{y-1}], \quad (11)$$

in which $V_M(\omega_0)$ is the speed of sound at a reference frequency ω_0 . For y as an odd integer,

$$1/V_M(\omega) = 1/V_M(\omega_0) - (2\alpha_{1M}\omega^{y-1}/\pi)(\ln|\omega| - \ln|\omega_0|). \quad (12)$$

II. HOOKE'S LAW FOR POWER-LAW ABSORPTION

From these expressions for attenuation and dispersion, explicit solutions for $r(t)$ for a more comprehensive Hooke's

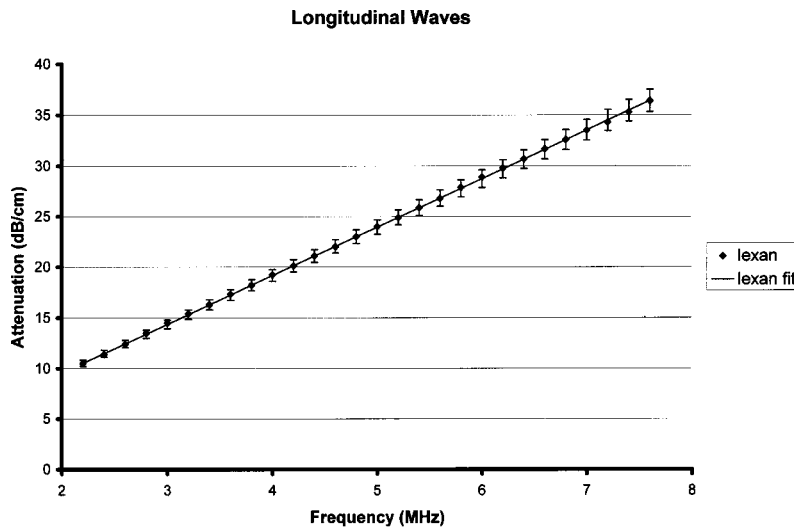


FIG. 4. Absorption loss data with error bars and frequency power law for longitudinal waves in Lexan.

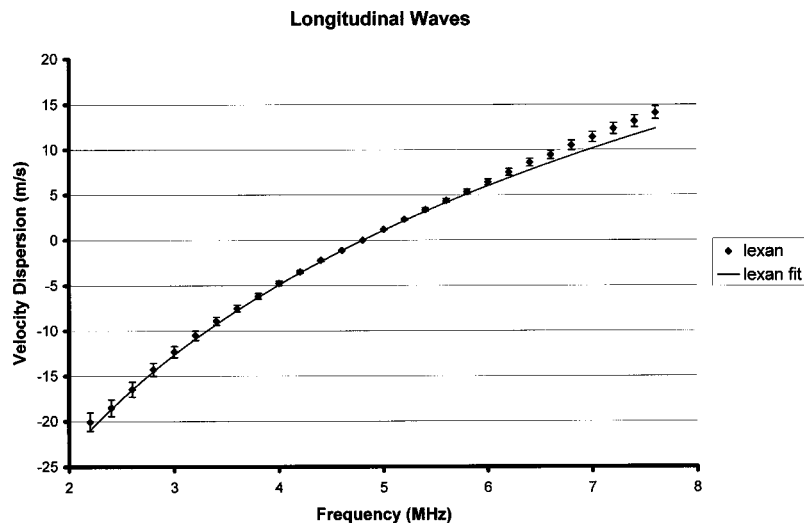


FIG. 5. Velocity dispersion predicted from time causal model compared to velocity data referenced to 4.8-MHz value with error bars for longitudinal waves in Lexan.

law, Eq. (3), are given in the Appendix. Different forms of this Hooke's law depend on the value of y . For y as an even integer,

$$T(t) = c:S(t) - (-1)^{y/2} \eta: \frac{\partial^{y-1} S(t)}{\partial t^{y-1}}, \quad (13)$$

and for the common case, $y=2$, Eq. (13) reduces to the Voigt model, Eq. (2). When y is an odd integer,

$$T(t) = c:S(t) + \eta: \left[\frac{2}{\pi} (y-1)! (-1)^{(y+1)/2} \frac{H(t)}{t^y} * S(t) \right], \quad (14)$$

in which $H(t)$ is the step function (Bracewell, 1968).

For the important case of $y=1$,

$$T(t) = c:S(t) + \eta: \left[\frac{-2}{\pi} \frac{H(t)}{t} * S(t) \right]. \quad (15)$$

Finally, for the most general case of y as a noninteger,

$$T(t) = c:S(t) + \eta: \left[\frac{-2}{\pi} \Gamma(y) \sin(\pi y/2) \frac{H(t)}{|t|^y} * S(t) \right], \quad (16)$$

where Γ is a gamma function. For y as a noninteger which is slightly greater than 1, $y = 1 + \varepsilon$,

$$T(t) \approx c:S(t) + \eta: \left[\frac{-2}{\pi} \frac{H(t)}{|t|^{1+\varepsilon}} * S(t) \right]. \quad (17)$$

With the help of generalized functions, the Fourier transform of the general stress Eq. (3) provides stress as a function of frequency,

$$T(\omega) = c:S(\omega) + \eta:R(S, \omega). \quad (18)$$

For the commonly used Voigt model, the Fourier transform of Eq. (2) leads to the following:

$$T(\omega) = c:S(\omega) + i\omega\eta:S(\omega). \quad (19)$$

Equation (19) provides the basis for the concept of a complex elastic constant, $c + i\omega\eta$ (Auld, 1990) for materials fitting a Voigt model. A complex elastic constant would not be appropriate for power-law absorption in general because of the more complicated relation that is expressed by Eq. (18).

Bagley and Torvik (1983) developed an empirical model for a Hooke's law based on the relation

$$T(\omega) = E_0:S(\omega) + (i\omega\eta)^{y-1}E_1:S(\omega), \quad (20)$$

where the constants E_0 and E_1 are determined empirically from data. Using fractional derivatives, they derived a model

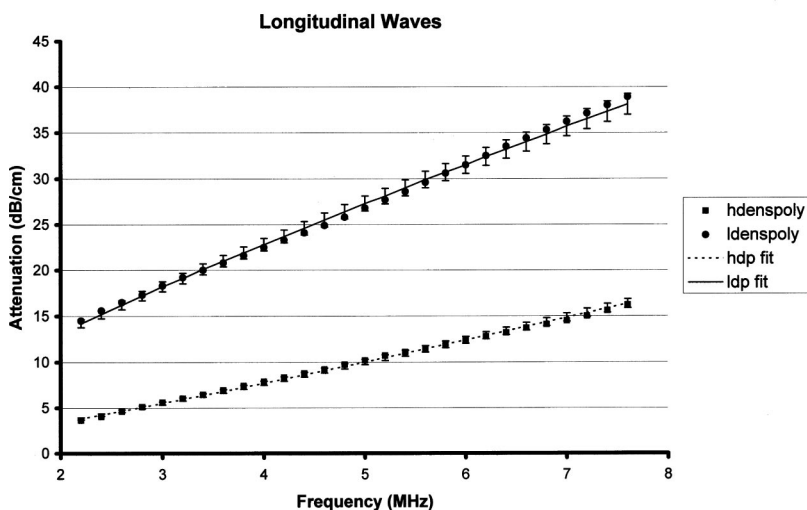


FIG. 6. Absorption loss data with error bars and frequency power law for longitudinal waves in low-density polyethylene and high-density polyethylene.

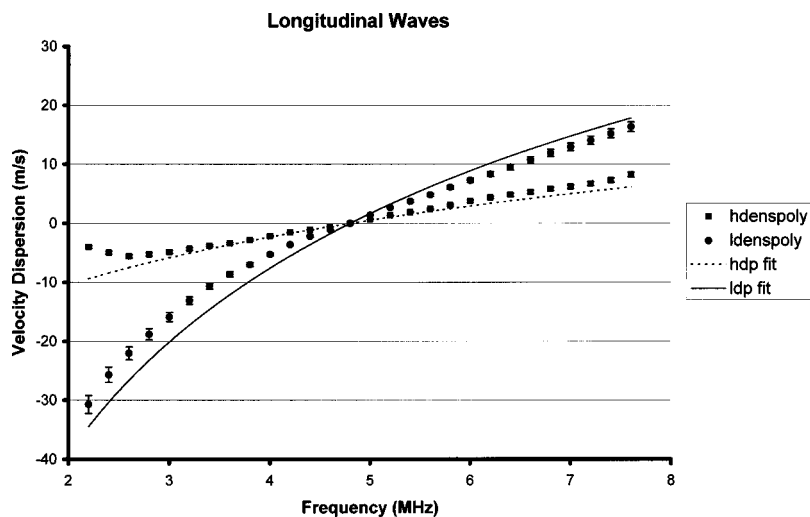


FIG. 7. Velocity dispersion predicted from time causal model compared to velocity data referenced to 4.8-MHz value with error bars for longitudinal waves in low-density polyethylene and high-density polyethylene.

for the range $1 < y < 2$. Their time-domain version of Eq. (20) is similar in form to Eq. (16) for the noninteger values of y in that it depends on a step function divided by t raised to a power of y ; however, it is missing the explicit dependence on y in Eq. (16) which is traceable to absorption. The sine term in Eq. (16) indicates that the factor should be maximum at $y=1$ (or odd integers) and diminish as y approaches 0 or 2, a characteristic that is seen in absorption data (Szabo, 1995). The experimental support for the empirical model is based on fits to complex moduli of the form of Eq. (20) for materials with $y \sim 1.5$, so it is not possible to determine how their data vary with y . The important cases of integral values of y , Eqs. (13) and (14) in the present model, are not covered by Bagley and Torvik's empirical model. Related approaches are discussed in Tschoegl (1989).

III. DATA

A. Experiments

For a given data set, the determination of α_{0M} , α_{1M} and y can be found from a least-squares power-law fit. Conversely, if a power-law fit were applied to velocity data, the absorption could be determined; however, this approach has been found to be less reliable experimentally. Typically, the velocity change with frequency is small, on the order of one part in 10^3 , so that it is more difficult to measure than attenuation (He, 1998b).

The shear and longitudinal wave data were obtained by the broadband substitution method. This technique, described in detail elsewhere (Wu, 1996), provides a means of determining both attenuation and phase velocity simulta-

neously over a broad bandwidth. In brief, the transmitted pulse between two broadband transducers in a water tank is recorded and digitized; then, a sample is inserted between them and the waveform transmitted through the specimen is also digitized. A complex spectrum is obtained from the test sample waveform and divided by the complex spectrum of the reference water waveform to calculate the attenuation and velocity of the sample. Angular rotation of the sample is used to cause complete mode conversion to obtain the shear wave characteristics. Data are corrected for impedance discontinuities at the sample interfaces by Fresnel equations.

B. Discussion

Longitudinal data are plotted in Figs. 4 and 5 for Lexan (a commercial name of a polymer plastic). The presentation of the data is the following: absorption data points are shown in Fig. 4 along with a least-squares power-law fit; second, the speed of sound data is plotted along with the velocity predicted from the absorption fit and Eq. (12) in Fig. 5. For Lexan, the longitudinal attenuation fit is a nearly linear fit to frequency; the velocity change is logarithmic. Similar results are given for two types of polyethylene in Figs. 6 and 7, where the absorption is no longer linear with frequency and Eq. (11) is used. Detailed coefficients and velocities are summarized in Tables I and II. Data were taken at room temperature (22°C). Error bars in the plots indicate the variation for both the attenuation and velocity data as well as the estimated error in the velocity prediction.

Shear wave data in Figs. 8–11 are different in several ways from the longitudinal data for these materials. The data for these plastics, even after the usual correction for impedance discontinuities at the sample interfaces, indicate that shear wave absorption is nearly an order of magnitude (in dB/cm) greater than the longitudinal cases. For shear waves,

TABLE I. Absorption coefficients ($\text{dB}/(\text{MHz})^y\text{-cm}$) and exponential powers y .

Material	Longitudinal α_{1L}	Longitudinal y	Shear α_{1S}	Shear y
Lexan	4.783	1.001	41.84	0.695
Low density polyethylene	7.577	0.796	28.64	0.815
High density polyethylene	1.522	1.171	22.80	0.950

TABLE II. Reference speed of sound @4.8 MHz ($\times 10^3$ m/s).

Material	Longitudinal	Shear
Lexan	2.195	0.943
Low-density polyethylene	2.566	1.273
High-density polyethylene	2.380	0.987

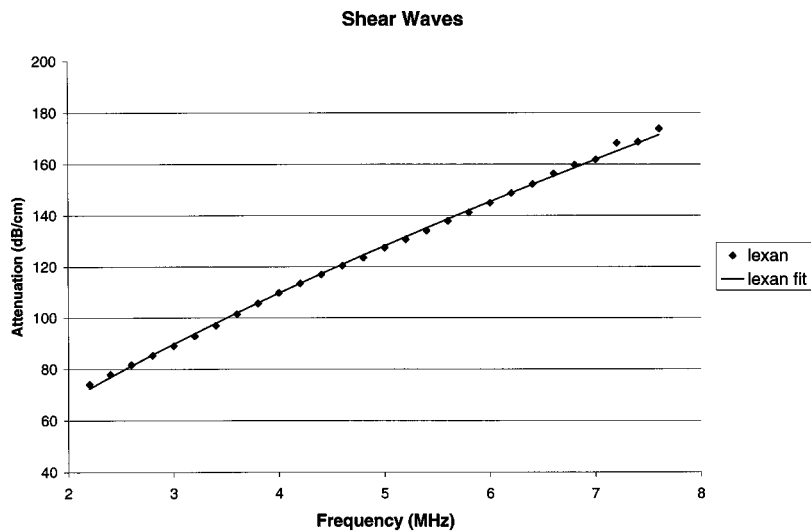


FIG. 8. Absorption loss data with error bars and frequency power law for shear waves in Lexan.

the absorption coefficient α_{0S} is usually nonzero, unlike the longitudinal absorption for these materials. Factors contributing to higher shear wave losses are the lower values of shear wave velocities and shorter wavelengths which lead to more loss per wavelength.

The relative magnitude of shear to longitudinal mode loss depends on material structure. Ferry (1980, pp. 34–37) identifies eight types of polymer structures. For example, polyethylenes fall into a cross-linked category of highly crystalline polymers. In a single-crystal rigid solid (nonpolymer), however, shear wave absorption can be less than longitudinal loss as illustrated by Fig. 1 for YIG.

IV. DISCUSSION

The dominance of the relaxation model for losses in acoustics, as originally derived by Stokes and Kirchhoff for fluids, is matched by a similar situation in electromagnetics where the Debye model has held an analogous position (Jonscher, 1977). The Debye model, originally formulated for noninteracting charged dipoles in a viscous medium, had been the main paradigm for explaining dielectric loss mechanisms. The real, X' , and imaginary susceptance, X'' , for this model, plotted in Fig. 12, are similar to absorption, α/ω , and

dispersion, β/ω , in the acoustic relaxation model of Fig. 2. Both the viscoelastic model and the Debye models are based on independent relaxation events.

Jonscher (1977), in examining dielectric losses for many solids over several decades of frequency, concluded that a fractional frequency power law was “the universal dielectric response.” Later, Hill and Jonscher (1983) developed a more general susceptance model that included a resonance and reduced to different fractional power laws at frequencies above and below the loss peak. They derived their model based on configurational quantum-mechanical tunneling and local fluctuations. Their empirical model is a Gaussian hypergeometric function of normalized frequency, ω/ω_p , where ω_p is the loss peak frequency, and empirically determined indices m and n (similar in function to y of present model), both less than 1. Well below ω_p , this function is close to a frequency raised to a fractional power m ; far above ω_p , it reduces to frequency raised to a fractional power $n-1$. Furthermore, the short-time transition time behavior is proportional to t^{-n} and long time behavior falls as $t^{-(1+m)}$.

In contrast, for acoustic waves in viscoelastic media, the frequency exponent is most often not fractional, but has been found to vary from 0 to 2. Longitudinal mode absorption

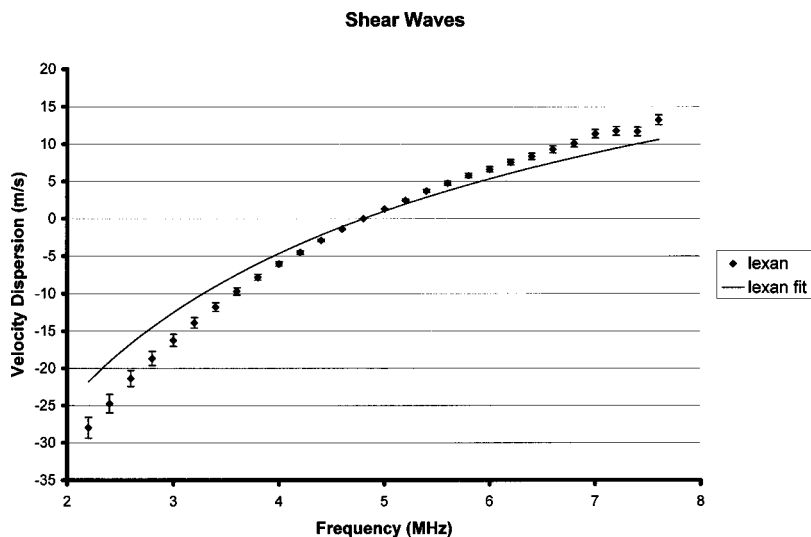


FIG. 9. Velocity dispersion predicted from time causal model compared to velocity data referenced to 4.8-MHz value with error bars for shear waves in Lexan.

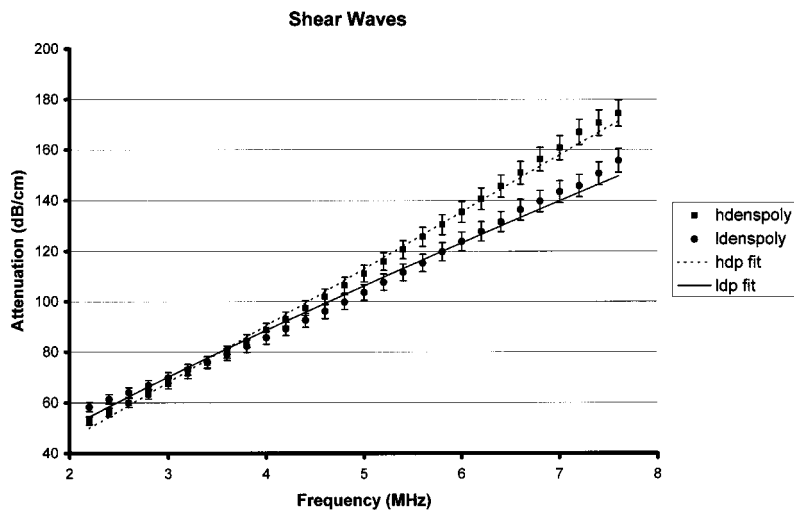


FIG. 10. Absorption loss data with error bars and frequency power law for shear waves in low-density polyethylene and high-density polyethylene.

obeys a power law with an exponent most frequently in the range of $1 < \gamma < 2$ for a large number of materials, both fluid and solid [Duck (1990); Zeqiri (1988); Bamber (1986); Szabo (1993, 1994, 1995); O'Donnell (1981), He (1998a, 1998b, 1999)]. For the shear wave losses in this work, $\gamma < 1$. In the present model, the function $r(t)$ varies inversely with a power of t like the transition times of the Hill and Jonscher dielectric model. This time dependence is in contrast with that of the relaxation model $\exp(-\omega\tau)$. Despite the fact that the acoustic model is for propagating waves, whereas the dielectric theory is for dynamic nonpropagating losses, there are differences as well as similarities.

Hill and Jonscher (1983) explained that their many-body interaction approach accounted for cooperative interactions in solids, and they obtained excellent agreement with data for a wide range of solids. If many-body interactions are occurring for acoustic and elastic waves, the appropriate underlying physical mechanisms are different than the dielectric case. For the plastic materials measured in this study, an extensive body of knowledge exists on the viscoelastic properties of polymers. Ferry (1980), in his text on polymers, distinguishes the character of polymers from the local viscous effects occurring in the deformation of hard solids, "In a polymer, on the other hand, each flexible threadlike mol-

ecule pervades an average volume much greater than atomic dimensions and is continually changing the shape of its contour as it wriggles and writhes with its thermal energy." Complicated interactions between macromolecules include entanglements, branching and cross-linking. In these cases, isolated relaxation mechanisms for acoustic losses seem an unlikely explanation for many materials, just as the isolated polar molecules of the Debye model fail in explaining measured dielectric losses. Bagley and Torvik (1983) and Ferry (1980) have derived macromolecular theories for specific cases. More work is necessary to uncover physical principles which cause the observed absorption to occur for any specific type of material; the role of cooperative interactions appears to be a promising direction for further understanding.

V. CONCLUSIONS

In summary, a comprehensive model accounting for absorption losses of the frequency power-law type has been presented in both time-domain and frequency-domain versions. In general, wave equations for these kinds of loss include convolution operators for both absorption and dispersion. Only for the simple cases of constant loss with

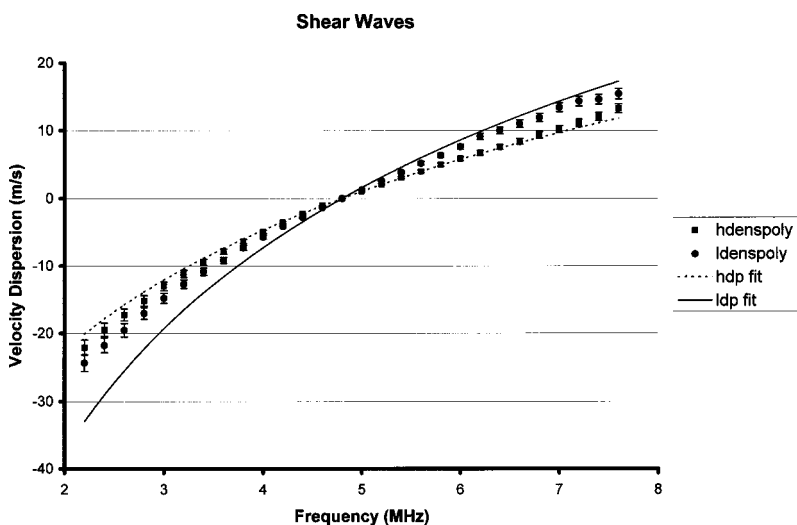


FIG. 11. Velocity dispersion predicted from time causal model compared to velocity data referenced to 4.8-MHz value with error bars for shear waves in low-density polyethylene and high-density polyethylene.

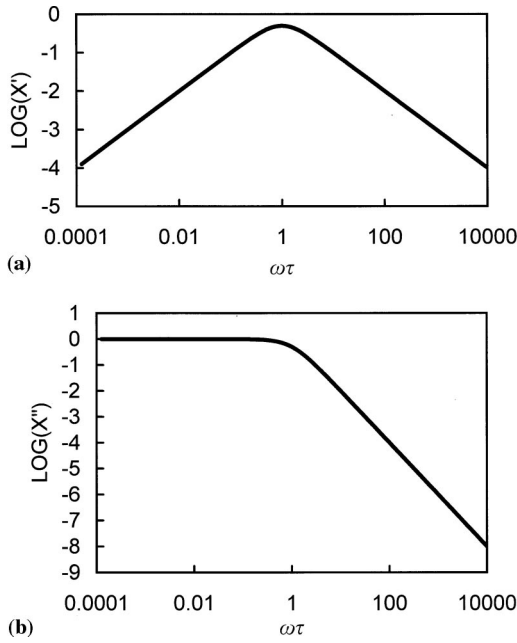


FIG. 12. (a) Real susceptibility X' vs angular frequency ω times τ for the Debye model. Loss peak at $\omega_p = 1/\tau$, where τ is relaxation time. (b) Imaginary susceptibility X'' vs frequency.

frequency or frequency-squared loss do the wave equations reduce to familiar derivative operators for absorption. A consequence of the model is a more comprehensive Hooke's law which reduces to the Voigt model Hooke's law under a low-frequency approximation. Data for losses of shear waves are consistent with the model presented and are much greater than those for longitudinal waves.

APPENDIX: DERIVATION OF LOW FREQUENCY APPROXIMATION OF VOIGT DASHPOT WAVE EQUATION

To derive our model, we follow the steps and the dyadic abbreviated subscript tensor notation used by Auld (1990) for the wave equation of the Voigt model. A particle velocity vector is defined by

$$v = \frac{\partial u}{\partial t}, \quad (\text{A1})$$

where u is a particle displacement vector. The strain-velocity relation is

$$\frac{\partial S}{\partial t} = \nabla_s v. \quad (\text{A2})$$

In order to find wave equations and solutions for the general power-law case of attenuation, Eq. (7), where y can be any positive, finite (reasonably valued) number, a more comprehensive form of Hooke's law is needed

$$T = c:S + \eta:r(S, t), \quad (\text{A3})$$

where c and η are stiffness and viscosity tensors, respectively, S is strain (a second-rank tensor) and in which $r(S, t)$ is yet to be determined. Except for Eq. (A3), the following steps are like those in Auld (1990):

(1) Differentiate Eq. (A3) with respect to time.

(2) Substitute relation (A2) in (A3) to obtain an expression in terms of v .

(3) Differentiate the equation of motion,

$$\nabla \cdot T = \rho \frac{\partial v}{\partial t} - F \quad (\text{A4})$$

(in which ρ is density and F is an applied force, a second-rank tensor) with respect to time and substitute the result for $\partial T / \partial t$ to obtain a wave equation,

$$\nabla \cdot c : \nabla_s v + \nabla \cdot \eta : r(\nabla_s v) = \rho \frac{\partial^2 v}{\partial t^2} - \frac{\partial F}{\partial t}. \quad (\text{A5})$$

In the case of a Voigt solid, $r(S, t) = \partial S / \partial t$, so that for a plane wave propagating along the z axis of an isotropic medium, Eq. (A5) breaks down into two separable modified wave equations for shear and longitudinal waves, respectively,

$$c_{44} \frac{\partial^2 v_x}{\partial z^2} + \eta_{44} \frac{\partial^3 v_x}{\partial z^2 \partial t} = \rho \frac{\partial^2 v_x}{\partial t^2}, \quad (\text{A6})$$

and

$$c_{11} \frac{\partial^2 v_z}{\partial z^2} + \eta_{11} \frac{\partial^3 v_z}{\partial z^2 \partial t} = \rho \frac{\partial^2 v_z}{\partial t^2}. \quad (\text{A7})$$

Phase velocities are defined as longitudinal, $V_L = \sqrt{c_{11}/\rho}$ and shear, $V_S = \sqrt{c_{44}/\rho}$. A solution of the form

$$v_M(z, \omega) = v_{OM}(0, \omega) e^{-i\omega t + \gamma z}, \quad (\text{A8})$$

in which M denotes shear or longitudinal mode, $M = S$ or L , ω is angular frequency, $\omega = 2\pi f$, and γ is the wave number or propagation factor as defined by Eq. (5), can be substituted in Eqs. (A6) and (A7) to obtain the wave-number dispersion relations used to calculate the curves of Fig. 2 for the Voigt model. When the loss per wavelength is small, the smallness approximation $(\alpha \bar{\lambda}_0)^2 \ll 1$, where $\bar{\lambda}_0 = V_{OM}/\omega$, can be applied to Eqs. (A6) and (A7). As shown in Szabo (1993, 1994), this criteria is equivalent to applying a plane wave approximation to the small viscosity terms in the above equations,

$$\frac{\partial v}{\partial t} = -\frac{1}{V} \frac{\partial v}{\partial z}, \quad (\text{A9})$$

therefore, Eqs. (A6) and (A7) become

$$\frac{\partial^2 v_x}{\partial z^2} - \frac{1}{V_S^2} \frac{\partial^2 v_x}{\partial t^2} + \frac{\eta_{44}}{c_{44} V_S^2} \frac{\partial^3 v_x}{\partial t^3} = 0, \quad (\text{A10})$$

$$\frac{\partial^2 v_z}{\partial z^2} - \frac{1}{V_L^2} \frac{\partial^2 v_z}{\partial t^2} + \frac{\eta_{11}}{c_{11} V_L^2} \frac{\partial^3 v_z}{\partial t^3} = 0. \quad (\text{A11})$$

This completes the low frequency approximation of the Voigt model.

In order to derive a more general viscoelastic wave equation, we return to the wave equation containing r , Eq. (A5), which, under the smallness approximation, leads to the counterparts of Eqs. (A10) and (A11) for the isotropic case

$$\frac{\partial^2 v_x}{\partial z^2} - \frac{1}{V_S^2} \frac{\partial^2 v_x}{\partial t^2} + \frac{\eta_{44}}{c_{44} V_S^2} \frac{\partial^2 r(v_x)}{\partial t^2} = 0, \quad (\text{A12})$$

$$\frac{\partial^2 v_z}{\partial z^2} - \frac{1}{V_L^2} \frac{\partial^2 v_z}{\partial t^2} + \frac{\eta_{11}}{c_{11} V_L^2} \frac{\partial^2 r(v_z)}{\partial t^2} = 0. \quad (\text{A13})$$

These last two equations can be represented by the general form

$$\frac{\partial^2 v}{\partial z^2} - \frac{1}{V_{0M}^2} \frac{\partial^2 v}{\partial t^2} - L_\gamma^* \frac{\partial^2 v}{\partial t^2} = 0, \quad (\text{A14})$$

where $L_\gamma(t)$ is a propagation operator. Equations of the same form as Eq. (A14) have been solved for frequency power-law attenuation (Szabo, 1993, 1994). The propagation operator is based on the constraint that causality imposes on Eqs. (A12)–(A14). This constraint, often expressed as a Kramers–Kronig relation, requires that the excess dispersion and attenuation be a Hilbert transform of each other

$$\beta_{EM}(\omega) = \frac{-1}{\pi \omega} * [-\alpha_{1M}(\omega)]. \quad (\text{A15})$$

The extra absorption term α_{0M} is a constant with frequency; therefore, dispersion does not depend on it. Two drawbacks of Eq. (A15) are that the attenuation must be known at all frequencies to determine the dispersion, and that the convolution diverges for values of $y \geq 1$. Appropriate convergent forms of the Kramers–Kronig relations can, however, be found by the method of subtractions and they will satisfy the Paley–Wiener theorem. Through the use of generalized functions, Szabo (1993, 1994) has derived a time causal counterpart of Eq. (A15)

$$L_{\beta_{EM}}(t) = -\text{sgn}(t) * L_\alpha(t), \quad (\text{A16})$$

where $\text{sgn}(t)$ is a signum function (Bracewell, 1968) and $L_{\beta_{EM}}$ and β_{EM} as well as L_α and α_{1M} are Fourier transform pairs. The overall propagation operator can be written as

$$L_\gamma(t) = L_\alpha(t) + iL_{\beta_{EM}}(t) = 2H(t)L_\alpha(t), \quad (\text{A17})$$

where $H(t)$ is the step function (Bracewell, 1968).

The function $r(t)$ can be found from the results previously derived for L_γ by equating Eqs. (A12) and (A13) to Eq. (A14)

$$-L_\gamma^* v = \frac{\eta_{IJ}}{c_{IJ} V_{0M}^2} \frac{\partial^2 r(v)}{\partial t^2} = \frac{2\alpha_{0M}}{V_{0M}} \frac{\partial^2 r(v)}{\partial t^2}. \quad (\text{A18})$$

The explicit results for $r(t)$ depend on whether y is an even, $r(t) = r_e(t)$, or an odd integer, $r(t) = r_o(t)$, or a noninteger $r(t) = r_{ni}(t)$. For y as an even integer,

$$r_e(t) = -(-1)^{y/2} \frac{\partial^{y-1}}{\partial t^{y-1}}. \quad (\text{A19})$$

Note for $y=2$, $r_e(t) = \partial/\partial t$ and so Eq. (3) becomes Eq. (2). For y as an odd integer,

$$r_o(t) = \frac{2}{\pi} (y-1)! (-1)^{(y+1)/2} \frac{H(t)}{t^y}. \quad (\text{A20})$$

For y as a noninteger,

$$r_{ni}(t) = \frac{-2}{\pi} \Gamma(y) \sin(\pi y/2) \frac{H(t)}{|t|^y}. \quad (\text{A21})$$

The above results are valid for a wide frequency range determined by the smallness approximation. The smallness approximation is $(\alpha \bar{\lambda}_0)^2 \ll 1$, where $\bar{\lambda}_0 = V_{0M}/\omega$. For absorption of the form of Eq. (1), a conservative upper-frequency limit f_{LIM} can be estimated (Szabo, 1993, 1994) by setting $\alpha \bar{\lambda}_0 = 0.1$ and solving for f

$$f_{\text{LIM}} = 0.1/(\alpha_{1M} 2\pi V_{0M})^{1/(y-1)}. \quad (\text{A22})$$

In most practical cases, this frequency far exceeds the finite bandwidth of acoustic waveforms of interest. In addition, values of attenuation at very high frequencies usually imply extremely high and unmeasurable values of loss. For example, the upper-frequency limit for the longitudinal YIG loss of Fig. 1 is 3.8×10^{12} Hz, corresponding to a loss of 2.9×10^7 dB/cm. For this conservative condition, $\alpha \bar{\lambda}_0 = 0.1$, the binomial expansion error for the propagation factor is 5×10^{-6} .

- Ahuja, A. S. (1979). "Tissue as a Voigt body for the propagation of ultrasound," *Ultrason. Imaging* **1**, 136–143.
- Aki, K., and Richards, P. G. (1980). *Quantitative Seismology* (Freeman, San Francisco), Vol. 1, Chap. 5.
- Auld, B. A. (1990). *Acoustic Fields and Waves in Solids*, 2nd ed. (Krieger, Malabar, FL), Vol. 1, Chap. 3.
- Bagley, R. L., and Torvik, P. J. (1983). "A theoretical basis for the application of fractional calculus to viscoelasticity," *J. Rheol.* **27**, 201–210.
- Bamber, J. C. (1986). *Physical Principles of Medical Ultrasonics*, edited by C. R. Hill (Wiley, Chichester), pp. 118–199.
- Bracewell, R. N. (1978). *The Fourier Transform and its Applications* (McGraw-Hill, New York).
- Collins, F., and Lee, C. C. (1956). "Seismic attenuation characteristics from pulse experiments," *Geophysics* **1**, 16–40.
- Duck, F. A. (1990). *Physical Properties of Tissue* (Academic, New York).
- Ferry, J. D. (1980). *Viscoelastic Properties of Polymers* (Wiley, New York), pp. 1–33.
- Goss, S. A., Frizzell, L. A., and Dunn, F. (1979). "Ultrasonic absorption and attenuation in mammalian tissues," *Ultrasound Med. Biol.* **5**, 181–186.
- Guess, J. F., and Campbell, J. S. (1995). "Acoustic properties of some biocompatible polymers at body temperature," *Ultrasound Med. Biol.* **21**, 273–277.
- He, P. (1998a). "Simulation of ultrasound pulse propagation in lossy media obeying a frequency power law," *IEEE Trans. Ultrason. Ferroelectr. Freq. Control* **45**, 114–125.
- He, P. (1998b). "Determination of ultrasonic parameters based on attenuation and dispersion measurements," *Ultrason. Imaging* **20**, 275–287.
- He, P. (1999). "Experimental verification of models for determining dispersion from attenuation," *IEEE Trans. Ultrason. Ferroelectr. Freq. Control* **46**, 706–714.
- Herzfeld, K. F., and Litovitz, T. A. (1959). *Absorption and Dispersion of Ultrasonic Waves* (Academic, New York).
- Hill, R. M., and Jonscher, A. K. (1983). "The dielectric behavior of condensed matter and its many-body interpretation," *Contemp. Phys.* **24**, 71–110.
- Horton, C. W. (1959). "A loss mechanism for the Pierre shale," *Geophysics* **24**, 667–680.
- Jonscher, A. K. (1977). "The universal dielectric response," *Nature (London)* **267**, 673–679.
- Kinsler, L. E., Frey, A. R., Coppens, A. B., and Sanders, J. V. (1982). *Fundamentals of Acoustics*, 3rd ed. (Wiley, New York).
- Kudo, N., Kamataki, T., Yamamoto, K., Onozuka, H., Mikami, T., Kitabatake, A., Ito, Y., and Kanda, H. (1997). "Ultrasound attenuation measurement of tissue in frequency range 2.5–40 MHz using a multi-resonance transducer," 1997 Proc. IEEE Ultrasonics Symp., 1181–1184.
- Nachman, A. I., Smith, J. F. III, and Waag, R. C. (1990). "An equation for acoustic propagation in inhomogeneous media with relaxation losses," *J. Acoust. Soc. Am.* **88**, 1584–1595.

- O'Donnel, M., Jaynes, E. T., and Miller, J. G. (1981). "Kramers–Kronig relationship between ultrasonic attenuation and velocity," *J. Acoust. Soc. Am.* **69**, 696–701.
- Polhammer, J. D., Edwards, C. A., and O'Brien, W. D. (1981). "Phase insensitive ultrasonic attenuation coefficient determination for fresh bovine liver over an extended frequency range," *Med. Phys.* **8**, 692–694.
- Postma, G. W. (1958). "Changes of shape of seismic impulses in inhomogeneous viscoelastic media," *Geophys. Prospect.* **6**, 438–455.
- Szabo, T. L. (1993). "Linear and nonlinear acoustic propagation in lossy media," Ph.D. thesis, University of Bath, U.K.
- Szabo, T. L. (1994). "Time domain wave equations for lossy media obeying a frequency power law," *J. Acoust. Soc. Am.* **96**, 491–500.
- Szabo, T. L. (1995). "Causal theories and data for acoustic attenuation obeying a frequency power law," *J. Acoust. Soc. Am.* **97**, 14–24.
- Tschoegl, N. W. (1989). *Phenomenological Theory of Linear Viscoelastic Behavior* (Springer, New York).
- White, J. E. (1965). *Seismic Waves Radiation, Transmission and Attenuation* (McGraw-Hill, New York).
- Wu, J. (1996). "Determination of velocity and attenuation of shear waves using ultrasonic spectroscopy," *J. Acoust. Soc. Am.* **99**, 2871–2875.
- Zeqiri, B. (1988). "An intercomparison of discrete-frequency and broadband techniques for the determination of ultrasonic attenuation," *Physics in Medical Ultrasound*, edited by D. H. Evans and K. Martin (IPSM, London), pp. 27–35.

Weakly localized waves on the corrugated surface of a medium of arbitrary anisotropy

A. N. Darinskii^{a)}

Institute of Crystallography, Academy of Sciences of Russia, Leninskii pr. 59, Moscow 117333, Russia

(Received 27 April 1999; accepted for publication 21 January 2000)

The existence of weakly localized surface and pseudosurface waves on the corrugated surface of a medium of arbitrary symmetry is studied. The localized solution originates from an exceptional bulk wave associated with a transonic state of positive or negative curvature. It is shown that the period of corrugation or the frequency of the wave can universally be chosen such that the solution will exist. The range of permissible periods (frequencies) is different for positive and negative curvature transonic states, it being wider near transonic states of positive curvature. The imaginary component of the pseudosurface wave velocity is found to vary as the height of grooves to period of corrugation ratio raised in fourth or sixth power, depending on the transonic state with which the exceptional wave is associated. A simple analytical expression is derived allowing estimations to be made of the penetration depth and the velocity of the localized wave. © 2000 Acoustical Society of America. [S0001-4966(00)05404-7]

PACS numbers: 43.35.Cg [HEB]

INTRODUCTION

A pure shear bulk wave with group velocity parallel to the smooth surface of a solid is known to meet the condition of a mechanically free surface both in isotropic and anisotropic elastic media. However, a slight perturbation of the conditions of propagation can cause localization of this shear wave. In particular the shear wave gives rise to a localized wave because of corrugation of the surface.¹⁻⁴

In elastically anisotropic media, bulk waves that satisfy the condition of free surface, so-called exceptional bulk waves, need not be pure shear polarized. The exceptional bulk waves appear in crystals of arbitrary symmetry and can be associated with a branch of bulk waves of any type (quasi-transverse or quasi-longitudinal). The sets of orientations allowing their existence form continuous one-dimensional subspaces (lines) in the three-dimensional space of angles that specify the orientation of the normal to the surface and the direction of propagation in it.⁵⁻⁷ Note that the polarization vector of an exceptional wave is necessarily parallel to the surface of the substrate.

It is natural to think that corrugation of the surface, like in the case of shear waves,¹⁻⁴ will cause transformation of an exceptional bulk wave into a surface or pseudosurface wave. In the present paper we consider this problem assuming that the substrate is of arbitrary symmetry. Our investigation will be based on the Stroh formalism^{8,9} and we begin with a brief discussion of basic relations to be used.

I. PLANE MODES IN ANISOTROPIC MEDIA

A plane mode

$$\mathbf{u}_\alpha(\mathbf{r}, t) = \mathbf{a}_\alpha(y) e^{i[kx - \omega t]}, \quad x = \mathbf{m}\mathbf{r}, \quad y = \mathbf{n}\mathbf{r}, \quad (1)$$

produces a force with the amplitude $\mathbf{f}_\alpha = -ik\mathbf{l}_\alpha(y)$ per unit area of a surface $y = \text{const}$; here \mathbf{m} and \mathbf{n} are unit vectors (Fig. 1),

$$\mathbf{l}_\alpha(y) = - \left[(nm)\mathbf{a}_\alpha - ik^{-1}(nn) \frac{d\mathbf{a}_\alpha}{dy} \right], \quad (2)$$

the symbols of the type (ab) stand for the 3×3 matrix $(ab)_{ij} = C_{kijl}a_k b_l$, C_{kijl} are the elastic-stiffness moduli of the medium, and a_k and b_l are components of the three-component vectors \mathbf{a} and \mathbf{b} , respectively.

According to Refs. 8 and 9, the vectors $\mathbf{a}_\alpha(y)$ and $\mathbf{l}_\alpha(y)$ corresponding to the plane wave solutions of the wave equation can be found from a set of six first-order differential equations,

$$\hat{\mathbf{N}} \xi_\alpha = \frac{1}{ik} \frac{d\xi_\alpha}{dy}, \quad (3)$$

where $\xi_\alpha(y) = (\mathbf{a}_\alpha(y), \mathbf{l}_\alpha(y))^t$ is a six-component vector column (the symbol t denotes transposition) and $\hat{\mathbf{N}}$ is a 6×6 real matrix,

$$\hat{\mathbf{N}} = - \begin{Bmatrix} \hat{\mathbf{N}}_{11} & \hat{\mathbf{N}}_{12} \\ \hat{\mathbf{N}}_{21} + \rho v^2 \hat{\mathbf{I}} & (\hat{\mathbf{N}}_{11})^t \end{Bmatrix}. \quad (4)$$

In (4),

$$\begin{aligned} \hat{\mathbf{N}}_{11} &= (nn)^{-1}(nm), \quad \hat{\mathbf{N}}_{12} = (nn)^{-1}, \\ \hat{\mathbf{N}}_{21} &= (mn)(nn)^{-1}(nm) - (mm), \end{aligned} \quad (5)$$

and $\hat{\mathbf{I}}$ is the 3×3 unit matrix, $(nn)^{-1}$ is the inverse of the matrix (nn) , ρ is the density, and $v = \omega/|k|$ is the phase speed of the mode (1) along the axis x (Fig. 1).

The matrix $\hat{\mathbf{N}}$ (Stroh's matrix) is generally nondegenerate and all six solutions of (3) have the form $\xi_\alpha(y) = \xi_\alpha e^{ikp_\alpha y}$, $\alpha = 1, \dots, 6$, where the constant vectors $\xi_\alpha = (\mathbf{A}_\alpha, \mathbf{L}_\alpha)^t$ and decay parameters p_α are found from the eigenvalue problem

^{a)}Electronic mail: adar@ns.crys.ras.ru

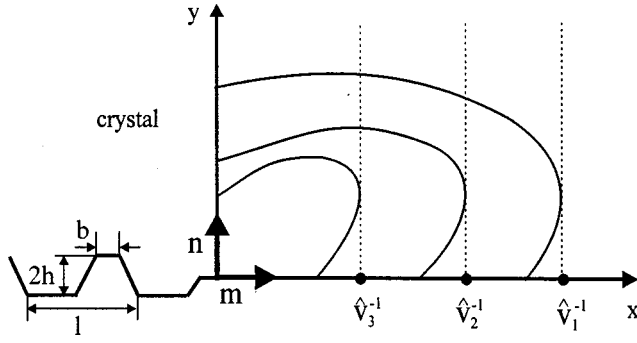


FIG. 1. Slowness curves ($1/v$) of the bulk waves in the substrate. Dashed lines are tangent to the slowness curve. \hat{v}_1 , \hat{v}_2 , and \hat{v}_3 are limiting speeds; only positive curvature transonic states are shown.

$$\hat{\mathbf{N}}\xi_\alpha = p_\alpha \xi_\alpha, \quad \alpha = 1, \dots, 6. \quad (6)$$

With regard to (4) and (5), Stroh's matrix obeys the symmetry relation

$$(\hat{\mathbf{T}}\hat{\mathbf{N}})' = \hat{\mathbf{T}}\hat{\mathbf{N}}, \quad \hat{\mathbf{T}} = \begin{pmatrix} \hat{\mathbf{O}} & \hat{\mathbf{I}} \\ \hat{\mathbf{I}} & \hat{\mathbf{O}} \end{pmatrix}, \quad (7)$$

where $\hat{\mathbf{O}}$ is the 3×3 zero matrix, where it follows that

$$\xi_\alpha \cdot \hat{\mathbf{T}}\xi_\beta = 0, \quad p_\alpha \neq p_\beta. \quad (8)$$

We shall assume, unless otherwise stated, that the vectors ξ_α are Stroh normalized,

$$\xi_\alpha \cdot \hat{\mathbf{T}}\xi_\alpha = 1. \quad (9)$$

Note that the projection of the energy flux of a bulk wave onto the normal \mathbf{n} can be written as $E_\alpha = -\omega k \xi_\alpha \cdot \hat{\mathbf{T}}\xi_\alpha^*/4$. Hence, the Stroh-normalized vector ξ_α corresponding to the reflected (incident) bulk wave is purely imaginary (real) when the wave propagates along the direction \mathbf{m} ($k > 0$) and vice versa when it propagates along the direction $-\mathbf{m}$ ($k < 0$).

The bulk waves with group velocity parallel to a plane $y = \text{const}$ (limiting bulk waves) are represented by points at which vertical lines touch slowness curves of bulk waves (Fig. 1). The situation $v = \hat{v}$, where \hat{v} is the limiting wave phase speed (limiting speed), is called a transonic state. In the present study we consider transonic states of the most common type, type 1 transonic states.⁹ At a type 1 transonic state there is only one limiting wave and the curvature of the slowness curve at the tangency point is not equal to zero (Fig. 1).

As v tends to \hat{v} , two eigensolutions, for example, (ξ_3, p_3) and (ξ_6, p_6) , coalesce to one eigensolution (ξ_{d3}, p_{d3}) that corresponds to the limiting wave (p_{d3} is real). The matrix $\hat{\mathbf{N}}$ becomes non-semi-simple degenerate⁹ and solutions to Eq. (3) at $v = \hat{v}$ are of the form $\xi_\alpha(y) = \xi_\alpha \exp(ikp_\alpha y)$, $\alpha = 1, 2, 4, 5, d3$, $\xi_{d6}(y) = (\xi_{d6} - ik y \xi_{d3}) \times \exp(ikp_{d3} y)$, where (ξ_α, p_α) , $\alpha = 1, 2, 4, 5, d3, d6$, are determined from

$$\begin{aligned} \hat{\mathbf{N}}\xi_\alpha &= p_\alpha \xi_\alpha, \quad \alpha = 1, 2, 4, 5, \quad \hat{\mathbf{N}}\xi_{d3} = p_{d3} \xi_{d3}, \\ \hat{\mathbf{N}}\xi_{d6} &= p_{d3} \xi_{d6} - \xi_{d3}. \end{aligned} \quad (10)$$

In view of (7)

$$\begin{aligned} \xi_\alpha \cdot \hat{\mathbf{T}}\xi_\beta &= 0, \quad \xi_{d3} \cdot \hat{\mathbf{T}}\xi_{d3} = 0, \\ \alpha &= 1, 2, 4, 5, \quad \beta = d3, d6, \end{aligned} \quad (11)$$

and one can arrange such that

$$\xi_{d3} \cdot \hat{\mathbf{T}}\xi_{d6} = 1, \quad \xi_{d6} \cdot \hat{\mathbf{T}}\xi_{d6} = 0. \quad (12)$$

It can be shown that the vectors ξ_{d3} and ξ_{d6} obeying Eq. (12) are purely real at positive curvature transonic states. At negative curvature transonic states they are purely imaginary.^{10–12}

In the vicinity of \hat{v} the eigenvectors $\xi_{3,6}(v)$ and the eigenvalues $p_{3,6}(v)$ have the form^{10–12}

$$\begin{aligned} \xi_3(v) &= \xi_{d3} - \Delta p \xi_{d6} + O(\Delta p^2), \\ p_3(v) &\approx p_{d3} + \Delta p + O(\Delta p^2), \\ \xi_6(v) &= \xi_{d3} + \Delta p \xi_{d6} + O(\Delta p^2), \\ p_6(v) &\approx p_{d3} - \Delta p + O(\Delta p^2), \end{aligned} \quad (13)$$

where

$$\Delta p = \sqrt{\rho} |\mathbf{A}_{d3}| f(v) \quad \text{at positive curvature transonic states,} \quad (14)$$

$$\Delta p = i\sqrt{\rho} |\mathbf{A}_{d3}| f(v) \quad \text{at negative curvature transonic states,} \quad (15)$$

$$f(v) = \sqrt{v^2 - \hat{v}^2}, \quad v > \hat{v}, \quad \text{and} \quad f(v) = i\sqrt{\hat{v}^2 - v^2}, \quad v < \hat{v}; \quad (16)$$

the vectors ξ_3 and ξ_6 are sought subject to the normalization condition

$$\xi_3 \cdot \hat{\mathbf{T}}\xi_3 = -2\Delta p, \quad \xi_6 \cdot \hat{\mathbf{T}}\xi_6 = 2\Delta p. \quad (17)$$

One can check that the subscripts $\alpha = 3$ and $\alpha = 6$ are used such that for real v and positive k the mode $\alpha = 3$ will be either an inhomogeneous “physical” mode (i.e., its amplitude decreases with distance from the surface) or a bulk reflected mode, whereas the mode $\alpha = 6$ will be either an inhomogeneous “nonphysical” mode (its amplitude increases with distance from the surface) or a bulk incident mode.

On the other hand, when the speed changes by δv , the nondegenerate eigensolutions of (10) [and of (6)] receive increments that vary as δv ,

$$\delta \xi_\alpha \propto \delta v, \quad \delta p_\alpha \propto \delta v, \quad (18)$$

rather than as $\sqrt{\delta v}$, like ξ_3 and ξ_6 near \hat{v} .

II. LOCALIZED WAVES ON THE CORRUGATED SURFACE

We assume that the limiting wave (the mode $\alpha = 3$ at $v = \hat{v}$) satisfies the condition of mechanically free surface on the plane $y = 0$, i.e., $\mathbf{L}_3(\hat{v}) \equiv \mathbf{L}_{d3} = 0$; the transonic state \hat{v} is then referred to as an exceptional transonic state. Now let the surface be corrugated. Its form is specified by the function

$$y(x) = \frac{\varepsilon}{Q} \sum_{n \neq 0} a_n \exp(inQx), \quad (19)$$

where $Q = 2\pi/l$, l is the period of corrugation, $\varepsilon \propto h/l$ is the small parameter, h is the depth of grooves, and a_n is the “amplitude” of the n th harmonic ($a_n = a_n^*$). For instance,

if the groove has the form of symmetric trapezoid (Fig. 1) then (see Ref. 4, Chap. 8)

$$\varepsilon = \frac{4h}{l} \frac{\sin \alpha}{\alpha}, \quad a_n = \sin \left(n \frac{\pi}{2} \right) \frac{\sin(n\alpha)}{n^2 \sin \alpha},$$

$$\alpha = \frac{\pi}{2} \left(1 - \frac{2b}{l} \right).$$
(20)

The solution to the boundary-value problem on the corrugated surface is sought in the form of a linear combination $\mathbf{u}(\mathbf{r}, t) = (\sum_n \mathbf{u}^{(n)}(\mathbf{r})) \exp(-i\omega t)$ involving the mode $\alpha=3$ with wave number k close to $\hat{k} = \omega/\hat{v}$, to which other inhomogeneous “physical” modes and reflected modes having wave numbers $k+nQ$, $n = \dots, -1, 0, 1, \dots$, are added:

$$\mathbf{u}^{(0)}(\mathbf{r}) = \left\{ \mathbf{A}_3 e^{ikp_3 y} + \sum_{\alpha=1}^2 b_\alpha \mathbf{A}_\alpha e^{ikp_\alpha y} \right\} e^{ikx},$$

$$\mathbf{u}^{(n)}(\mathbf{r}) = \sum_{\alpha=\alpha_1, \alpha_2, \alpha_3} c_\alpha^{(n)} \mathbf{A}_\alpha^{(n)} e^{i(k+nQ)(x+p_\alpha^{(n)}y)}.$$
(21)

For the sake of definiteness it is assumed that the mode $\alpha=3$ travels to the right (in Fig. 1), i.e., $\text{Re}(k) > 0$. The speed $\hat{v} + \delta v = \omega/k$ together with the amplitude coefficients b_α and $c_\alpha^{(n)}$ are to be found from the boundary conditions.

If all modes incorporated in (21) are inhomogeneous, we obtain a pure surface solution. However, if (21) involves reflected modes, then the solution will be leaky, in which case the speed will be complex, $\hat{v} + \delta v = v_l' - i v_l''$, where the imaginary component v_l'' describes radiative loss of the leaky wave.

We assume that the speed ranges of our concern are far from the branch points of all functions $p_\alpha(v)$, $\alpha=1, 2$, and $p_\alpha^{(n)}(v)$ except for $p_3(v)$ in (13). Since the mode $\alpha=3$ must become either “physical” inhomogeneous or bulk reflected as v_l'' tends to zero, it is expedient to consider the function $f(v)$ in (16) for complex speeds as the analytical continuation of (16) onto the complex plane with cutoff branch along $[\hat{v}, \hat{v} - i\infty)$ and take $\arg(v)$ equal to $3\pi/2$ and $-\pi/2$ on the left and right edges, respectively. The argument of Δp in (14) and (15) then ranges within the intervals $[-\pi/4, 3\pi/4]$ and $[\pi/4, 5\pi/4]$, respectively. However, as a matter of fact, the argument of Δp corresponding to a solution of the boundary-value problem may belong only to the intervals $[-\pi/4, 0)$ and $[\pi/2, 3\pi/4]$ (positive curvature transonic states) or to the intervals $[\pi/4, \pi/2]$ and $(\pi, 5\pi/4]$ (negative curvature transonic states), because otherwise the existence of the solution would be in contradiction with the law of energy conservation.

The boundary conditions make it possible to derive an approximate expression for the decay parameter Δp of the mode $\alpha=3$. Multiplying mechanical stresses that a plane mode produces on the surface in (19),

$$\sigma_{ij} = ik C_{ijkl} (m_k + p n_k) A_l e^{ikpy(x)},$$
(22)

by the unit normal to this surface gives the force \mathbf{f} , with which the mode acts upon the surface. To estimate Δp in the lowest approximation relative to ε , it is sufficient to take \mathbf{f} in the form

$$\mathbf{f} \approx -ik \left\{ \mathbf{L} + i\varepsilon \sum_{n \neq 0} a_n e^{inQx} \left\{ n[(mm) + p(mn)] \mathbf{A} + \frac{k}{Q} p \mathbf{L} \right\} \right\}$$

$$= -ik \left\{ \mathbf{L} - i\varepsilon \sum_{n \neq 0} a_n e^{inQx} \left\{ \frac{nQ+k}{Q} \right. \right.$$

$$\left. \times [\hat{\mathbf{N}}_{21} \mathbf{A} + (\hat{\mathbf{N}}_{11})^t \mathbf{L}] + \frac{k}{Q} \rho v^2 \mathbf{A} \right\},$$
(23)

where we have eliminated $p\mathbf{A}$ and $p\mathbf{L}$ with the aid of (6).

Using (23), taking into account (13) and (18), and remembering that $\mathbf{L}_{d3}=0$, and that then, due to (10) and (4),

$$\hat{\mathbf{N}}_{21} \mathbf{A}_{d3} = -\rho v^2 \mathbf{A}_{d3},$$
(24)

we obtain the following approximate expressions for the forces $\mathbf{f}^{(0)}(x)$ and $\mathbf{f}^{(n)}(x)$ associated with $\mathbf{u}^{(0)}(\mathbf{r})$ and $\mathbf{u}^{(n)}(\mathbf{r})$ in (21), respectively:

$$\mathbf{f}_0(x) \approx -i\hat{k} \left\{ \left\{ -\Delta p \mathbf{L}_{d6} + \sum_{\alpha=1}^2 b_\alpha \mathbf{L}_\alpha \right\} e^{ikx} \right.$$

$$\left. + i\varepsilon \rho \hat{v}^2 \mathbf{A}_{d3} \left\{ \sum_{n \neq 0} a_n n e^{i(k+nQ)x} \right\} \right\},$$
(25)

$$\mathbf{f}^{(n)}(x) \approx -i(\hat{k} + nQ) \left\{ \mathbf{L}^{(n)} e^{i(k+nQ)x} \right.$$

$$\left. - i\varepsilon a_{-n} \left\{ \frac{\hat{k}}{Q} \{ \hat{\mathbf{N}}_{21} \mathbf{A}^{(n)} + (\hat{\mathbf{N}}_{11})^t \mathbf{L}^{(n)} \} \right. \right.$$

$$\left. + \frac{\hat{k} + nQ}{Q} \rho v_n^2 \mathbf{A}^{(n)} \right\} e^{ikx},$$
(26)

where $\mathbf{L}^{(n)} = \sum_\alpha c_\alpha^{(n)} \mathbf{L}_\alpha^{(n)}$, $\mathbf{A}^{(n)} = \sum_\alpha c_\alpha^{(n)} \mathbf{A}_\alpha^{(n)}$, and $v_{(n)} = \omega/|\hat{k} + nQ|$. The vectors $\mathbf{A}^{(n)}$ and $\mathbf{L}^{(n)}$ are computed at $v = v_{(n)}$ and $\mathbf{L}_{1,2}$ in (25) at $v = \hat{v}$.

Combining Eqs. (25) and (26), equating the preexponential factors to zero yields

$$-\Delta p \mathbf{L}_{d6} + \sum_{\alpha=1}^2 b_\alpha \mathbf{L}_\alpha - i\varepsilon \sum_{n \neq 0} a_{-n} \frac{\hat{k} + nQ}{\hat{k}} \left\{ \frac{\hat{k}}{Q} \{ \hat{\mathbf{N}}_{21} \mathbf{A}^{(n)} \right.$$

$$\left. + (\hat{\mathbf{N}}_{11})^t \mathbf{L}^{(n)} \} + \frac{\hat{k} + nQ}{Q} \rho v_n^2 \mathbf{A}^{(n)} \right\} = 0,$$
(27)

$$\mathbf{L}^{(n)} = -i\varepsilon a_n n \rho \hat{v}^2 \mathbf{A}_{d3} \frac{\hat{k}}{\hat{k} + nQ}.$$
(28)

Let $|\hat{k} + nQ| \neq 0$, \hat{k} , and $k_R = \omega/v_R$, where v_R is the velocity of the Rayleigh wave on the flat surface. Equation (28) then allow the coefficients $c_\alpha^{(n)}$ in (21) to be found. These coefficients are small in the measure of the smallness of ε .

Further, substitution of $\mathbf{L}^{(n)}$ as given by (28) in Eq. (27) leads to

$$-\Delta p \mathbf{L}_{d6} + \sum_{\alpha=1}^2 b_{\alpha} \mathbf{L}_{\alpha} - i\varepsilon \sum_{n \neq 0} a_{-n} \left\{ \frac{\hat{k} + nQ}{Q} \hat{\mathbf{N}}_{21} \mathbf{A}^{(n)} + \rho \hat{v} v_Q \mathbf{A}^{(n)} \right\} = 0, \quad (29)$$

where $v_Q = \omega/Q$. Relations (11) and (12) implying $\mathbf{A}_{d3} \cdot \mathbf{L}_{\alpha} = 0$, $\alpha = 1, 2$, and $\mathbf{A}_{d3} \cdot \mathbf{L}_{d6} = 1$ if $\mathbf{L}_{d3} = 0$, we can multiply (29) by the vector \mathbf{A}_{d3} to obtain, accounting for the symmetry of the matrix $\hat{\mathbf{N}}_{21}$ and Eq. (24),

$$\Delta p \approx i\varepsilon \rho \hat{v}^2 \mathbf{A}_{d3} \cdot \left(\sum_{n \neq 0} a_{-n} n \mathbf{A}^{(n)} \right). \quad (30)$$

From (29) it is also follows that the coefficients b_{α} , $\alpha = 1, 2$, vary as ε^2 because $\mathbf{A}^{(n)} \propto \varepsilon$. Note also that the expansion of Δp with respect to ε involves only terms proportional to even powers of ε , since Δp cannot depend on the sign of ε .

III. EXISTENCE CONDITIONS FOR SURFACE AND LEAKY WAVES

Let us analyze expression (30). It is seen that $\Delta p \propto \varepsilon^2$ so that the correction δv to \hat{v} will be of order ε^4 . However, it should be emphasized that (30) viewed as an equation for δv need not have a solution. Remembering the permissible range of Δp , we conclude that (30) is solvable providing its right-hand side, which we denote by Ξ , satisfies the conditions

$$\text{Re}(\Xi) \leq 0, \quad \text{Im}(\Xi) > 0, \quad \text{Im}(\Xi) \geq |\text{Re}(\Xi)| \quad (31a)$$

or

$$\text{Re}(\Xi) \geq 0, \quad \text{Im}(\Xi) < 0, \quad \text{Re}(\Xi) \geq |\text{Im}(\Xi)| \quad (31b)$$

(positive curvature transonic states); or

$$\text{Re}(\Xi) \leq 0, \quad \text{Im}(\Xi) < 0, \quad |\text{Re}(\Xi)| \geq |\text{Im}(\Xi)| \quad (32a)$$

or

$$\text{Re}(\Xi) \geq 0, \quad \text{Im}(\Xi) > 0, \quad \text{Im}(\Xi) \geq \text{Re}(\Xi) \quad (32b)$$

(negative curvature transonic states).

Suppose that v_n , $n = \dots, -1, 1, \dots$, lie in the subsonic speed interval, or $v_{(n)}$ are less than the limiting speed \hat{v}_1 (Fig. 1). All the modes involved in $\mathbf{u}^{(n)}(\mathbf{r})$ in (21) are then inhomogeneous. Since $\hat{k} - nQ < 0$, the negative-traveling modes incorporated into $\mathbf{u}^{(-n)}(\mathbf{r})$, $n = 1, \dots$, must have the decay parameters $p_{\alpha}^{(-n)}$ with negative imaginary parts for their amplitudes to decrease with depth into the substrate. On the other hand, $\mathbf{u}^{(n)}(\mathbf{r})$ involves positive-traveling modes having $\text{Im}(p_{\alpha}^{(n)}) > 0$ because $\hat{k} + nQ > 0$.

According to Refs. 13 and 14, the traction \mathbf{L}_{α} and the polarization vector \mathbf{A}_{α} of a plane mode can be interrelated with the help of the admittance matrix $\hat{\mathbf{Y}}$ independent of the index α as follows: $\mathbf{A}_{\alpha} = i\hat{\mathbf{Y}}\mathbf{L}_{\alpha}$ if $\text{Im}(p_{\alpha}) > 0$, and $\mathbf{A}_{\alpha} = -i\hat{\mathbf{Y}}\mathbf{L}_{\alpha}$ if $\text{Im}(p_{\alpha}) < 0$. As a result, due to (28)

$$\mathbf{A}^{(n)} = \varepsilon \rho \hat{v} v_n a_n n \hat{\mathbf{Y}}^{(n)} \mathbf{A}_{d3}, \quad n = \pm 1, \pm 2, \dots, \quad (33)$$

where $\hat{\mathbf{Y}}^{(n)} = \hat{\mathbf{Y}}(v_n)$.

The admittance $\hat{\mathbf{Y}}$ is a Hermitian matrix. Hence, $\text{Im}(\hat{\mathbf{Y}})$ is an antisymmetric matrix. As a result, $\mathbf{A}_{d3} \cdot \text{Im}(\hat{\mathbf{Y}}) \mathbf{A}_{d3} = 0$ because the vector \mathbf{A}_{d3} is either purely real or purely imaginary depending on whether the curvature of the transonic state is positive or negative, respectively. Equation (30) then reduces to

$$\Delta p \approx i\varepsilon^2 \rho^2 \hat{v}^3 \mathbf{A}_{d3} \cdot \left\{ \sum_{n \neq 0} v_n |a_n|^2 n^2 \text{Re}[\hat{\mathbf{Y}}^{(n)}] \right\} \mathbf{A}_{d3}. \quad (34)$$

First of all, from (34) one sees that Δp will be purely imaginary, $\Delta p \approx \Xi = i \text{Im}(\Xi)$. Let us discuss the sign of Ξ . Consider the positive curvature transonic state (\mathbf{A}_{d3} is real).

As shown in Refs. 13 and 14, $\hat{\mathbf{B}} = (\text{Re}[\hat{\mathbf{Y}}])^{-1}$ is a positive definite matrix at $v = 0$. Its eigenvalues λ_i , $i = 1, 2, 3$, are positive at $v = 0$ and steadily decrease with increasing v from 0 to the limiting velocity \hat{v} bounding from above the subsonic interval (\hat{v}_1 in Fig. 1). At the Rayleigh wave velocity v_R two of these eigenvalues, e.g., $\lambda_{1,2}$, vanish. The third eigenvalue λ_3 becomes zero at \hat{v}_1 if the limiting wave at \hat{v}_1 is exceptional. Otherwise λ_3 remains positive up to \hat{v}_1 . If the subsonic Rayleigh wave does not exist, then the eigenvalues $\lambda_{1,2}$ are positive up to \hat{v}_1 .

Accordingly the eigenvalues $\tilde{\lambda}_i = 1/\lambda_i$, $i = 1, 2, 3$, of $\text{Re}[\hat{\mathbf{Y}}]$ behave as follows. The eigenvalues are positive at $v = 0$ and steadily increase with v up to \hat{v}_1 . At $v = v_R$ the functions $\tilde{\lambda}_{1,2}(v)$ have discontinuity, $\tilde{\lambda}_{1,2}(v)$ tending to $+\infty$ and $-\infty$ as v tends to $v_R - 0$ and $v_R + 0$, respectively. The eigenvalue $\tilde{\lambda}_3$ is positive up to \hat{v}_1 . It tends to $+\infty$ if the limiting wave is exceptional.

As a result, the contraction

$$\mathbf{A}_{d3} \cdot \text{Re}[\hat{\mathbf{Y}}] \mathbf{A}_{d3} = \sum_{i=1}^3 \tilde{\lambda}_i(v) |e_i(v) \cdot \mathbf{A}_{d3}|^2, \quad (35)$$

where $\mathbf{e}_i(v)$ are the eigenvectors of $\text{Re}[\hat{\mathbf{Y}}(v)]$, will be positive below v_R (or below \hat{v}_1 when the subsonic Rayleigh wave does not exist) and negative within a velocity interval above v_R .

Thus, with regard to the definition of $\text{Im}(\Xi)$ we conclude that this quantity will be necessarily positive within the interval $v_{-1} < v_R$ or within the whole subsonic interval providing the Rayleigh wave does not appear there. However, if $v_{-1} > v_R$, then $\text{Im}(\Xi)$ can be negative.

The inequality $\text{Im}(\Xi) > 0$ implies that the corrugation of the surface causes localization of the exceptional bulk wave. From (34) it follows that in the first approximation the speed of the wave propagating along the corrugated surface will be real and smaller than \hat{v} . If the modes $\alpha = 1, 2$ in $\mathbf{u}^{(0)}(\mathbf{r}, t)$ in (21) are inhomogeneous near \hat{v} , i.e., if \hat{v} is the slowest transonic state, \hat{v}_1 in Fig. 1, then the higher-order corrections to the speed will also be real and the wave will be pure surface quasi-bulk. Note that the speed cannot appear complex, since a complex speed would correspond to a leaky solution; but in the case under discussion, the linear superposition (21) does not involve homogeneous modes which could carry energy from the surface into the bulk of the substrate.

At the same time if \hat{v} is the intermediate or highest transonic state, \hat{v}_2 or \hat{v}_3 in Fig. 1, then one of the modes

$\alpha=1,2$, or both of them, will be homogeneous reflected, so that the solution will be leaky. When evaluated with higher accuracy, Δp appears to be complex, $\text{Re}(\Delta p) \propto \varepsilon^4$, and the imaginary part v_l'' of the leaky wave speed will be proportional to ε^6 . To estimate v_l'' , one can also take advantage of the fact that the coefficients b_α vary as ε^2 and the mode $\alpha=3$ penetrates the substrate a distance of order $(\Delta p)^{-1} \propto \varepsilon^2$; correspondingly $v_l'' \propto |b_\alpha|^2 \Delta p \propto \varepsilon^6$.

When $\text{Im}(\Xi) < 0$, the equation $\Delta p = i \text{Im}(\Xi)$ is not solvable with respect to v , implying that the boundary-value problem does not possess a ‘‘physical’’ solution. In this case one can, however, replace the mode $\alpha=3$ with the mode $\alpha=6$ to obtain a formal ‘‘nonphysical’’ solution of the boundary-value problem. Such a solution will be ‘‘nonphysical’’ because the amplitude of the mode $\alpha=6$ grows with depth into the medium.

Summed up, it is seen that if the transonic state is of positive curvature, then by choosing properly the period of corrugation or the frequency of the wave one can arrange such that a ‘‘physical’’ solution (pure surface or leaky) will exist.

Let us discuss how corrugation modifies the exceptional bulk wave associated with negative curvature transonic states. Remembering the above-indicated properties of $\text{Re}[\hat{\mathbf{Y}}]$, taking into account that \mathbf{A}_{d3} is a purely imaginary vector, we conclude that $\text{Im}(\Xi)$ is necessarily negative when $v_{-1} < v_R$ and $\text{Im}(\Xi)$ is allowed to be positive when $v_{-1} > v_R$. In view of (16) and (32), the ‘‘physical’’ solution will exist provided $\text{Im}(\Xi) > 0$. Hence, in distinction to the previous case, the ‘‘physical’’ solution does not exist when $v_{-1} < v_R$.

The $\text{Im}[\Xi]$ value becomes positive at least within an interval $v_R < v_{-1} < v_c \leq \hat{v}_1$, where v_c is a certain critical speed such that $\text{Im}[\Xi] = 0$ at $v_{-1} = v_c$ (if $v_c < \hat{v}_1$), unless \mathbf{A}_{d3} is orthogonal to the eigenvectors of $\hat{\mathbf{B}}$ associated to the vanishing eigenvalues. This means that in the case of negative curvature transonic state corrugation also makes the exceptional wave localized. However, now the interval of permissible Q or ω is bounded both from below and above. Besides, the solutions will be necessarily leaky because a negative curvature transonic state cannot be the slowest transonic state so that at least one of the modes $\alpha=1,2$ turns out to be a bulk reflected mode. The imaginary component of the leaky wave speed varies as ε^6 . Note also that the real part of the speed will be larger than \hat{v} .

Consider now the case when some of the modes involved in $\mathbf{u}^{(n)}(\mathbf{r})$ in (21) are bulk reflected at real v . For the sake of definiteness we suppose that there is only one bulk mode $\alpha=r$ which travels with the speed v_{-1} along the direction $-\mathbf{m}$. The right-hand side of Eq. (30) then appears to be complex,

$$\begin{aligned} \text{Re}(\Xi) &\approx \varepsilon \rho \hat{v}^2 \text{Im}[a_1 \mathbf{A}_{d3} \cdot \mathbf{A}^{(-1)}], \\ \text{Im}(\Xi) &\approx \varepsilon \rho \hat{v}^2 \text{Re}\left[\mathbf{A}_{d3} \cdot \sum_{n \neq 0} a_{-n} n \mathbf{A}^{(n)}\right]. \end{aligned} \quad (36)$$

Let us determine the sign of $\text{Re}(\Xi)$. To this end, using (28) we form the vector

$$\xi = \sum_{\alpha_1, \alpha_2, r} c_\alpha^{(-1)} \xi_\alpha^{(-1)} = (\mathbf{A}^{(-1)}, -i \varepsilon \rho \hat{v} v_{-1} a_{-1} \mathbf{A}_{d3})^t, \quad (37)$$

where the vectors $\xi_\alpha^{(-1)} = (\mathbf{A}_\alpha^{(-1)}, \mathbf{L}_\alpha^{(-1)})^t$ are associated with the modes α_1, α_2 , and r involved in $\mathbf{u}^{(-1)}(\mathbf{r})$ in (21). By our assumption, α_1 and α_2 are inhomogeneous modes so that the vectors $\xi_{\alpha_1}^{(-1)}$ and $\xi_{\alpha_1}^{(-1)*}$ as well as $\xi_{\alpha_2}^{(-1)}$ and $\xi_{\alpha_2}^{(-1)*}$ correspond to different (complex conjugate) eigenvalues of Stroh’s matrix. Hence, due to (8), $\xi_{\alpha_1}^{(-1)} \cdot \hat{\mathbf{T}} \xi_{\alpha_1}^{(-1)*} = \xi_{\alpha_2}^{(-1)} \cdot \hat{\mathbf{T}} \xi_{\alpha_2}^{(-1)*} = 0$.

On the other hand, the vector $\xi_r^{(-1)}$ associated with the reflected negative-traveling wave obeys the relation $\xi_r^{(-1)} \cdot \hat{\mathbf{T}} \xi_r^{(-1)*} = \xi_r^{(-1)} \cdot \hat{\mathbf{T}} \xi_r^{(-1)} = 1$ [see the discussion around Eqs. (8) and (9)]. With this in mind, we can multiply the vector ξ in (37) by the vector $\hat{\mathbf{T}} \xi^*$ to obtain

$$\xi \cdot \hat{\mathbf{T}} \xi^* = |c_r^{(-1)}|^2 = \mp 2 \varepsilon \rho \hat{v} v_{-1} \text{Im}[a_1 \mathbf{A}_{d3} \cdot \mathbf{A}^{(-1)}], \quad (38)$$

where it follows that

$$\text{Re}(\Xi) = \mp 0.5 |c_r^{(-1)}|^2 \frac{\hat{v}}{v_{-1}}, \quad (39)$$

depending on whether the transonic state is of positive or negative curvature, respectively.

Thus in the case under discussion corrugation in a general manner causes transformation of an exceptional bulk wave into a leaky wave. With regard to (39) the solution exists provided the condition (31a) or (32b) is fulfilled. Note that both the real and imaginary components of the leaky wave speed will be of order ε^4 .

IV. EXAMPLE

Consider the propagation of shear waves on the corrugated surface (19). We assume that the substrate is of monoclinic symmetry $2/m$ (the least-symmetric case allowing the existence of pure shear modes; notice that the increase of material symmetry does not simplify either intermediate evaluations or final expressions). Let the plane of symmetry coincide with the plane (\mathbf{n}, \mathbf{m}) (Fig. 1). It is apparent that (21) will incorporate only shear modes. Explicitly $p_\alpha, \mathbf{A}_\alpha$, and \mathbf{L}_α corresponding to a pair of shear modes read as

$$\begin{aligned} p_{3,6} &= -\frac{c_{45}}{c_{44}} \pm \Delta p, \quad \Delta p = \sqrt{\frac{\rho}{c_{44}}} \sqrt{v^2 - \hat{v}^2}, \\ \hat{v} &= \sqrt{\frac{c_{44} c_{55} - c_{45}^2}{c_{44} \rho}}, \\ A_{3z} &= \frac{i}{\sqrt{2 c_{44} \Delta p}}, \quad L_{3z} = -i \sqrt{\frac{c_{44} \Delta p}{2}}, \\ A_{6z} &= \frac{1}{\sqrt{2 c_{44} \Delta p}}, \quad L_{6z} = \sqrt{\frac{c_{44} \Delta p}{2}}, \end{aligned} \quad (40)$$

where c_{ij} are the elastic moduli with respect to the coordinate system xyz (Fig. 1) and \hat{v} is the limiting speed of shear bulk waves. The vectors \mathbf{A}_α and \mathbf{L}_α are Stroh normalized and directed along the axis z . The transonic state is of positive

curvature; $p_{d3} = -c_{45}/c_{44}$, $A_{d3z} = 1/\sqrt{c_{44}}$, and $L_{d6z} = \sqrt{c_{44}}$ ($L_{d3z} = A_{d6z} = 0$). The subscripts $\alpha = 3, 6$ are assigned in such a way that the positive-traveling mode $\alpha = 3$ will be either inhomogeneous with the amplitude decaying with depth into the substrate or the reflected mode, whereas $\alpha = 6$ will be either the “physical” inhomogeneous mode or the reflected mode for the opposite direction of propagation.

Let $Q > 2\hat{k}$. In this instance all shear modes involved in the problem will be inhomogeneous and the wave on the corrugated surface will be pure surface. Using (40) we obtain via Eqs. (28) and (30)

$$\Delta p \approx i\varepsilon^2 \left(\frac{\rho\hat{v}^2}{c_{44}} \right)^{3/2} \left\{ \sum_{n \neq 0} |a_n|^2 n^2 \frac{v_n}{\sqrt{\hat{v}^2 - v_n^2}} \right\}. \quad (41)$$

In the case under discussion $\Delta p/i$ is positive irrespective of the relationship between v_n and v_R since due to symmetry shear modes are not coupled with sagittally polarized modes, from which the Rayleigh wave is constructed. Notice also that for sinusoidal relief $y(x) = (\varepsilon/Q)\sin(Qx)$ in the limit $Q \gg \hat{k}$ in (41) reduces to the expression for the decay factor derived in Ref. 4, Chap. 6.

Now let $\hat{k} < Q < 2\hat{k}$. Then the mode with the speed v_{-1} will be the bulk reflected mode, so that the solution, if it exists, must be of the leaky type. Computations give

$$\begin{aligned} \text{Re}(\Xi) &\approx -\varepsilon^2 \left(\frac{\rho\hat{v}^2}{c_{44}} \right)^{3/2} \frac{|a_1|^2 v_{-1}}{\sqrt{v_{-1}^2 - \hat{v}^2}}, \\ \text{Im}(\Xi) &\approx \varepsilon^2 \left(\frac{\rho\hat{v}^2}{c_{44}} \right)^{3/2} \left\{ \sum_{n \neq 0, -1} |a_n|^2 n^2 \frac{v_n}{\sqrt{\hat{v}^2 - v_n^2}} \right\}. \end{aligned} \quad (42)$$

The leaky solution exists provided $|\text{Re}(\Xi)| \leq \text{Im}(\Xi)$ [Eq. (31a)]. This inequality may or may not hold. In particular, it is not difficult to check that on the sinusoidal surface $|\text{Re}(\Xi)| > \text{Im}(\Xi)$ and, hence, the equation $\Delta p - \Xi + \dots = 0$ has no solution, that is, leaky shear waves do not appear.

V. CONCLUSION

The analysis performed in the study indicates that three factors control the existence of surface (pseudo-surface) waves originating from an exceptional bulk wave on the corrugated surface of an anisotropic medium. These are (1) the relationship between the wave number of the exceptional bulk wave $\hat{k} = \omega/\hat{v}$ and the wave number Q of corrugation; (2) the sign of the curvature of the slowness curve at the exceptional transonic state; and (3) the number of homogeneous and inhomogeneous modes at the exceptional transonic state.

When the exceptional bulk wave is associated with the positive curvature transonic state, it is always possible to choose the period of corrugation (or the frequency ω) such that the solution of the boundary-value problem will exist. Specifically, the condition $v_{-1} = \omega/|\hat{k} - Q| < v_R$ guarantees the existence of the wave solution on the corrugated surface. In the lowest approximation with respect to the height-to-period of corrugation ratio, this wave is constructed from the weakly localized mode $\alpha = 3$, to which combinations of three

localized modes having the wave number $k + nQ$, where k is the wave number of the mode $\alpha = 3$, are generally adjacent for the boundary conditions to be fulfilled.

The penetration depth of the mode $\alpha = 3$ varies as $\varepsilon^{-2} \gg 1$ and the speed of the solution is smaller than the exceptional bulk wave speed \hat{v} by a quantity of the order of ε^4 . The amplitudes $c_\alpha^{(n)}$ of the adjacent modes, where n is the number of a harmonic involved in (19) ($a_n \neq 0$), are of order ε . The amplitude of the other harmonics involved in (21), including the modes $\alpha = 1, 2$ with the same wave number as the mode $\alpha = 3$, will be of higher order of the smallness. However, the latter two modes can play an important role although their amplitudes are smaller than $c_\alpha^{(n)}$ in (21), $b_\alpha \propto \varepsilon^2 \ll c_\alpha^{(n)}$. The matter is that if the exceptional bulk wave is associated with the intermediate or highest transonic state, one of the modes $\alpha = 1, 2$, or both them, will be bulk reflected so that, once the corrugation “mixes” the modes $\alpha = 3$ and $\alpha = 1, 2$, the exceptional bulk wave appears to be transformed into the leaky wave. The imaginary component of the leaky wave speed will be of order ε^6 . However, if the transonic state is the slowest one and, hence, the modes $\alpha = 1, 2$ are inhomogeneous, then the wave will be pure surface and these modes can be viewed just as higher-order corrections to $\mathbf{u}^{(n)}(\mathbf{r})$ in (21).

When $v_{-1} > v_R$, the quasi-bulk wave (surface or pseudo surface) may or may not exist.

Quasi-bulk solutions near negative curvature exceptional transonic states do not exist providing $v_{-1} < v_R$ and can appear on the condition that $v_{-1} > v_R$. Note that such a solution will be leaky since a negative curvature transonic state cannot be the slowest transonic state. Accordingly at least one of the modes $\alpha = 1, 2$ will be bulk reflected.

One can arrange such that some of the modes involved in $\mathbf{u}^{(n)}(\mathbf{r})$ in (21) will be bulk reflected for real v . Then the exceptional bulk wave in a general manner gives rise to the leaky wave, the imaginary component of its speed varying as ε^4 . For the leaky solution to exist the material constants of the medium must obey certain conditions of the type of inequalities. These inequalities do not hold necessarily. For example, leaky SH waves of the type under discussion do not exist on the surface of sinusoidal form.

ACKNOWLEDGMENTS

The author is grateful to V. I. Alshits and V. N. Lyubimov for fruitful discussions and to the Russian Foundation for Basic Research (Grant No. 98-02-16077) for partial financial support. Also, the author thanks the reviewer for correcting the English in the initial version of the paper.

¹B. A. Auld, J. J. Gagnepain, and M. Tan, “Horizontal shear waves on corrugated surfaces,” *Electron. Lett.* **12**(24), 650–651 (1976).

²Yu. V. Gulyaev and V. P. Plessky, “‘Slow’ acoustic surface waves in solids,” *Sov. Tech. Phys. Lett.* **3**(3), 87–88 (1977).

³Yu. V. Gulyaev and V. P. Plessky, “Propagation of acoustic surface waves in periodic structures,” *Sov. Phys. Usp.* **32**(1), 51–74 (1989).

⁴S. V. Biryukov, Yu. V. Gulyaev, V. V. Krylov, and V. P. Plessky, *Surface Acoustic Waves in Inhomogeneous Media* (Springer, Berlin, 1995).

⁵V. I. Alshits and J. Lothe, “Elastic waves in triclinic crystals. III. The problem of existence and some general properties of exceptional surface waves,” *Sov. Phys. Crystallogr.* **24**, 1122–1130 (1979).

- ⁶V. I. Alshits and J. Lothe, "Comments on the relation between surface wave theory and the theory of reflection," *Wave Motion* **3**, 297–310 (1981).
- ⁷V. I. Alshits, V. N. Lyubimov, N. F. Naumenko, N. V. Perelomova, and A. L. Shuvalov, "Exceptional elastic body waves in crystals of various symmetries," *Sov. Phys. Crystallogr.* **30**, 123–126 (1985).
- ⁸A. N. Stroh, "Steady state problems in anisotropic elasticity," *J. Math. Phys.* **41**, 77–103 (1962).
- ⁹P. Chadwick and G. D. Smith, "Foundation of the theory of surface waves in anisotropic elastic media," *Adv. Appl. Mech.* 17, edited by C.-S. Yih, (Academic, New York, 1977), pp. 303–376.
- ¹⁰A. N. Darinskii, "Quasi-bulk Rayleigh waves in semi-infinite media of arbitrary anisotropy," *Wave Motion* **27**(1), 79–93 (1998).
- ¹¹A. N. Darinskii, "Leaky waves and the elastic wave resonance reflection on a crystal-thin solid layer interface. II. Leaky waves given rise to by an exceptional bulk wave," *J. Acoust. Soc. Am.* **103**, 1845–1854 (1998).
- ¹²A. N. Darinskii, V. I. Alshits, and J. Lothe, "Simple reflection and leaky waves in the vicinity of a line of exceptional bulk waves," *Wave Motion* **30**, 253–274 (1999).
- ¹³J. Lothe and D. M. Barnett, "Further development of the theory for surface waves in piezoelectric crystals," *Phys. Norv.* **8**(4), 239–254 (1976).
- ¹⁴D. M. Barnett and J. Lothe, "Free surface (Rayleigh) waves in anisotropic half-spaces: the surface impedance method," *Proc. R. Soc. London, Ser. A* **402**, 135–152 (1985).

The quasistatic approximation for a cracked interface between a layer and a substrate

Claudio Pecorari

Institute for Advanced Materials, European Commission, PO Box 2, 1755 ZG Petten, The Netherlands

Piarras A. Kelly

Department of Engineering Science, University of Auckland, Private Bag 92019, Auckland, New Zealand

(Received 7 January 1999; accepted 3 February 2000)

Heuristic in nature, the quasistatic approximation (QSA) describes the interaction of ultrasonic waves with imperfect interfaces by modeling the interfacial imperfection as distributions of springs and masses. The QSA does not provide any relationship between the interfacial stiffness constants and the micromechanics of the defects. The aims of this paper are threefold. First, a derivation from first principles of the QSA boundary conditions on a cracked interface is presented. Relationships linking the interfacial constants to the mechanical and geometrical properties of the distributed cracks are also obtained. Second, the stiffness dependence of a cracked interface between a layer and a substrate on the layer thickness is investigated. It is shown that the interfacial stiffnesses cannot be regarded as intrinsic properties of the interface, but they may also depend on the structural properties of the hosting system. Finally, the effect of the thickness dependence of the interfacial stiffnesses on the phase velocity of the lowest mode supported by the layered structure is investigated. © 2000 Acoustical Society of America. [S0001-4966(00)01905-6]

PACS numbers: 43.35.Cg, 43.35.Zc [HEB]

INTRODUCTION

For the last two decades the quasistatic approximation (QSA) has been the approach most commonly used to describe the interaction of ultrasonic waves with imperfect interfaces. The QSA is a low-frequency approximation and it can be used when the thickness of the interface and the extension of the interfacial defects are much smaller than the wavelength of the wave used to inspect the interface. The most complete formulation of the QSA has been presented by Baik and Thompson,¹ and models the real interfacial imperfections as continuous, uniform distributions of springs along the interface plane (see Fig. 1).

The mathematical formulation of the QSA is provided by the modified boundary conditions enforced at the interface plane. Following Baik and Thompson,¹ the QSA boundary conditions can be written as follows,

$$\begin{aligned} \frac{1}{2}[\sigma_{33}(x, z=0^+) + \sigma_{33}(x, z=0^-)] \\ = K_N[u_3(x, z=0^+) - u_3(x, z=0^-)], \end{aligned} \quad (1a)$$

$$\begin{aligned} \frac{1}{2}[\sigma_{31}(x, z=0^+) + \sigma_{31}(x, z=0^-)] \\ = K_T[u_1(x, z=0^+) - u_1(x, z=0^-)], \end{aligned} \quad (1b)$$

$$\sigma_{33}(x, z=0^+) = \sigma_{33}(x, z=0^-), \quad (1c)$$

$$\sigma_{31}(x, z=0^+) = \sigma_{31}(x, z=0^-). \quad (1d)$$

In these equations K_N and K_T are the stiffness constants of the distributed springs, and relate the discontinuity of the displacement components to the corresponding components of the stress applied to the interface. The QSA boundary conditions are not derived from first principles. Rather, they are heuristic in nature.¹

The QSA does not provide any way to correlate the values of the spring stiffness constants K_N and K_T to the micromechanics of the defects. Baik and Thompson¹ defined the spring constants as the ratio between the stress applied at “infinity,” $\Sigma_{N,T}$, and the extra displacement, $\Delta_{N,T}$, measured at a location far from the interface,

$$K_{N,T} = \frac{\Sigma_{N,T}}{\Delta_{N,T}}. \quad (2)$$

The extra displacement, $\Delta_{N,T}$, is not zero when the interface contains imperfections that alter its elastic properties. A few simple cases, such as that of an interphase layer embedded between two infinite half-spaces, and those of periodic distributions of one- and two-dimensional cracks, have been considered in detail,¹ and expressions for the spring constants have been presented in terms of the geometrical and mechanical properties of the distributed defects.

The QSA boundary conditions [Eqs. (1a) and (1d)] have been widely applied to isolated interfaces¹ as well as to interfaces in layered media,^{2,3} or between fibers and a matrix⁴. In connection to systems featuring a characteristic length, such as the thickness of a layer, or the radius of a fiber, the legitimate question arises whether the system's structure affects the elastic properties of the imperfect interface.

The objectives of this work are threefold. The first objective is to present a derivation from first principles of the QSA boundary conditions for a randomly cracked interface. In so doing, expressions for the spring constants that link the values of these quantities to the structural and micromechanical properties of the crack distribution are obtained. Second, the spring constants for a cracked interface between a layer and its substrate is evaluated numerically. To this end, the crack opening displacement (COD) of an isolated

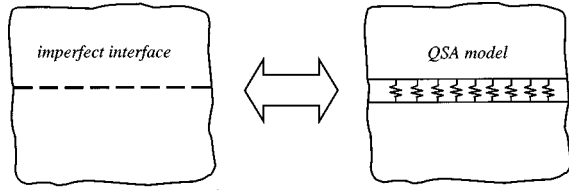


FIG. 1. Imperfect interface and its model according to the QSA.

interfacial crack undergoing an external uniform load is obtained by numerically solving a system of integral equations for the dislocation densities associated with the components of the COD. The dislocation densities are equivalent to the crack surface displacement gradients. The introduction of the crack compliance tensor connecting the COD to the applied stress leads to the evaluation of the spring constants. The third objective is to assess the influence of the structural properties of the system on the interfacial stiffness and, consequently, on the dispersion of the modes supported by the layered structure.

I. THE QSA BOUNDARY CONDITIONS ON A CRACKED INTERFACE

In this section the cracks are assumed to be one-dimensional. However, the derivation presented here, as well as the relationship between the spring stiffness constants and the properties of the distributed imperfections, can be extended to distributions of two-dimensional cracks in a straightforward manner.

Consider a random distribution of cracks at the interface between two media. Let $z=0$ be the interface plane, and $u_i^+(x)$ and $u_i^-(x)$ be the i th component of the displacement just above and just below the interface, respectively. The average displacement discontinuity at the interface can be written as

$$\langle u_i^+(x) - u_i^-(x) \rangle = \langle \Delta u_i(x) \rangle = \frac{1}{L} \int_L \Delta u_i(x) dx, \quad i=x, z, \quad (3)$$

where L is the representative length of the crack distribution. The displacement discontinuity $\Delta u_i(x)$ is zero except at the locations of the cracks, and there it is equal to the COD, $b_i(x)$. Then, Eq. (3) can be written as

$$\langle \Delta u_i \rangle = \frac{1}{L} \sum_{k=1}^N \int_{a_k} b_i^k(x) dx. \quad (4)$$

In Eq. (4), N is the number of cracks in L , and a_k is the length of the k th crack. For the sake of simplicity, let the distribution be uniform and the cracks identical to each other. Then Eq. (4) becomes

$$\langle \Delta u_i \rangle = \nu a \langle b_i \rangle, \quad (5)$$

where $\nu = N/L$ is the crack density, a is the crack length, and $\langle b_i \rangle$ is the average i th component of the COD. The crack compliance tensor, S_{ij} , that relates the average displacement components to the average stress, $\langle \sigma_{ij} \rangle$, applied to the crack faces, can now be introduced:⁵

$$\langle b_i \rangle = a S_{ij} \langle \sigma_{jz} \rangle n_z, \quad i, j = x, z. \quad (6)$$

Note that here the z axis is assumed normal to the interface and, consequently, only the components of the stress tensor with indexes “ jz ” appear in this equation. The z -component of the unit vector n is equal to 1 and, therefore, it will be omitted hereafter. By introducing Eq. (6) into Eq. (5) the latter becomes

$$\langle \Delta u_i \rangle = \nu a^2 S_{ij} \langle \sigma_{jz} \rangle. \quad (7)$$

On the assumption that the cracks do not interact with each other, it can be shown numerically that the average normal (tangential) displacement component due to a uniform shear (normal) stress field is zero. Thus, only one term remains on the right-hand side of Eq. (7). By inverting Eq. (7), the QSA boundary condition for a cracked interface can be obtained,

$$\langle \sigma_{xz} \rangle = K_{xx} \langle \Delta u_x \rangle, \quad (8a)$$

$$\langle \sigma_{zz} \rangle = K_{zz} \langle \Delta u_z \rangle, \quad (8b)$$

where

$$K_{xx} = K_T = \frac{1}{\nu a^2 S_{xx}}. \quad (9a)$$

$$K_{zz} = K_N = \frac{1}{\nu a^2 S_{zz}}. \quad (9b)$$

Equations (9a) and (9b) show that K_T and K_N are inversely proportional to the crack density, to the square of the crack length, and to the crack compliance. Therefore, both the geometrical and the micromechanical properties of the interfacial defects are included in the definition of the macroscopic interfacial properties. Note that the definition of K_T and K_N involves quantities that are intrinsic properties of the distribution, a and ν , or that are derived from average values of the stress field and COD, S_{xx} and S_{zz} .

Disregarding the mutual interaction between neighboring cracks of a planar distribution seems to be a reasonable simplification that is supported by the behavior of the total stress field in the neighborhood of a one-dimensional crack subjected to a static stress. In fact, as shown by Kachanov,⁵ although the total stress field is amplified in the plane containing the crack, the regions where the amplification effect occurs have an extent at most of the order of the crack length. Therefore, the range of validity of the independent crack approximation can be thought to reach values of the normalized crack density, $a\nu$, up to 0.5. At larger crack densities crack interaction begins to occur, and off-diagonal terms in the boundary conditions are expected to play a significant role.

II. MICROMECHANICS OF A CRACKED INTERFACE

A. Isolated interfacial crack

In this subsection the mechanical response to an external load of a crack located at the interface between a layer and a substrate is examined. To evaluate the COD of such a crack, a system of integral equations for the unknown dislocation densities is solved.⁶⁻⁸ The problem is formulated in its full complexity, so that the crack closure near the crack tips due to the different elastic properties of the two media is correctly described by the solution.⁶ However, to avoid large-

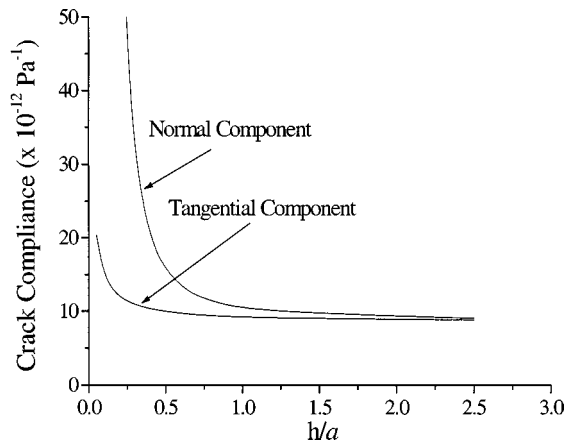


FIG. 2. Compliance of an isolated crack at the interface between a copper layer and a steel substrate versus the ratio of the layer thickness to the crack length, h/a .

scale crack closure during a compressive cycle,⁹ a tensile stress, T_o , is assumed to maintain the crack open all the times (save for the small-scale crack-tip closure). The same system of integral equations is solved twice. The first time, the crack is subjected to the tensile stress T_o only, while the second time, a normal, ΔT , or a tangential, ΔS , stress is superimposed onto T_o . The magnitude of the stresses ΔT and ΔS are chosen to be one order of magnitude smaller than that of T_o . Once the dislocation densities (i.e., crack surface displacement gradients) are known, the components of the COD can be evaluated by integrating them over the crack extension. The difference between the two CODs obtained with and without the stress perturbation is eventually evaluated and used to obtain the effective average COD due to the stress wave. Finally, Eq. (6) is used to calculate the components of the crack compliance tensor, S_{ij} .

Consider a system consisting of a layer of copper on a steel substrate, and let h be the thickness of the layer. Figure 2 illustrates the behavior of the normal and tangential component of the crack compliance as a function of the ratio h/a . The plots show that both components tend to infinity as the layer thickness decreases, while they approach the same limit as the layer approximates a half-space. It is worth noting that the normal compliance is always larger than the transverse compliance. This fact can be easily understood in terms of the amount of material surrounding the crack that opposes the crack deformation.

B. Cracked interface between dissimilar materials

1. Cracked interface between two identical half-spaces

Baik and Thompson¹ gave the solution for the normal spring constant for an interface consisting of a periodic array of one-dimensional cracks between two identical half-spaces. In the limit $a/\Lambda \ll 1$, where a is the length of the crack and Λ is the period of the distribution, they found

$$K = \frac{2}{\pi} \frac{E}{1 - \eta^2} \frac{1}{\nu a^2}, \quad (10)$$

where E and η are the Young modulus and the Poisson ratio of the material, respectively, and $\nu = \Lambda^{-1}$. When the two

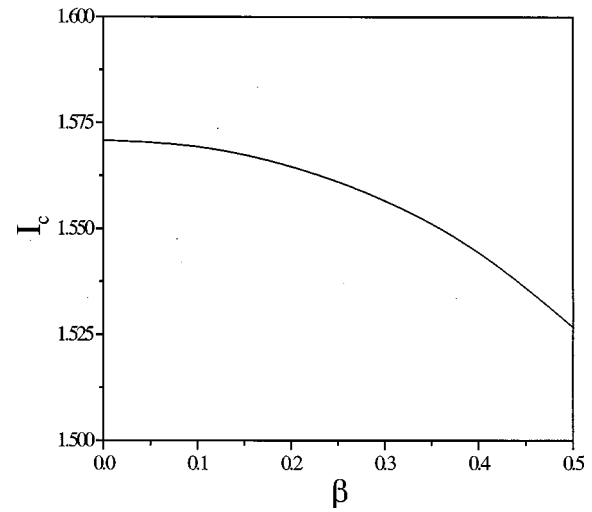


FIG. 3. Plot of the integral I_c versus the composite material parameter β .

half-spaces are dissimilar, the closed form solution for the COD of an interface crack can be obtained using the well-known “open” interface crack model.⁷ This model, unlike the “closed” model used elsewhere in this study, does not assume crack closure near the crack-tips. The COD solution can be used to show that Eq. (9) leads to the following result,⁸

$$K_N = K_T = \Gamma \sqrt{1 - \beta^2} \frac{1}{\nu a^2 I_c}. \quad (11)$$

In Eq. (11), Γ is defined by the following expression,

$$\Gamma = \frac{2\mu^{(l)}(1 - \alpha)}{(1 + \kappa^{(l)})(1 - \beta^2)}, \quad (12)$$

where μ is the shear modulus of medium, K is Kolosov's constant, equal to $3 - 4\eta$ in plane strain, and the superscript (l) refers to the layer. The functions α and β in Eq. (12) are the well-known Dundurs' composite parameters,^{7,10}

$$\alpha = \frac{\mu^{(l)}(\kappa^{(s)} + 1) - \mu^{(s)}(\kappa^{(l)} + 1)}{\mu^{(l)}(\kappa^{(s)} + 1) + \mu^{(s)}(\kappa^{(l)} + 1)}, \quad (13)$$

$$\beta = \frac{\mu^{(l)}(\kappa^{(s)} - 1) - \mu^{(s)}(\kappa^{(l)} - 1)}{\mu^{(l)}(\kappa^{(s)} + 1) + \mu^{(s)}(\kappa^{(l)} + 1)},$$

wherein the superscript (s) refers to the substrate. The quantity I_c represents the following integral,

$$I_c = \int_{-1}^{+1} \sqrt{1 - t^2} \cos \left(\epsilon \ln \left| \frac{t - 1}{t + 1} \right| \right) dt, \quad (14)$$

where $\epsilon = (1/2\pi) \ln |(\beta + 1)/(\beta - 1)|$. The integral I_c is a function of β only (see Fig. 3), and can be evaluated numerically. However, a good approximation to I_c can be obtained by neglecting the cosine term, that is, by setting $\beta = 0$ in which case $I_c = \pi/2$. The maximum difference between the actual value of I_c and $\pi/2$ is obtained when β is maximum, i.e., $\beta = 0.5$, and is equal to 0.044. When the materials are similar, $\beta = 0$, $I_c = \pi/2$, $\Gamma = E/[4(1 - \eta^2)]$, and Eq. (11) reduces to the result of Baik and Thompson, Eq. (10). In the follow-

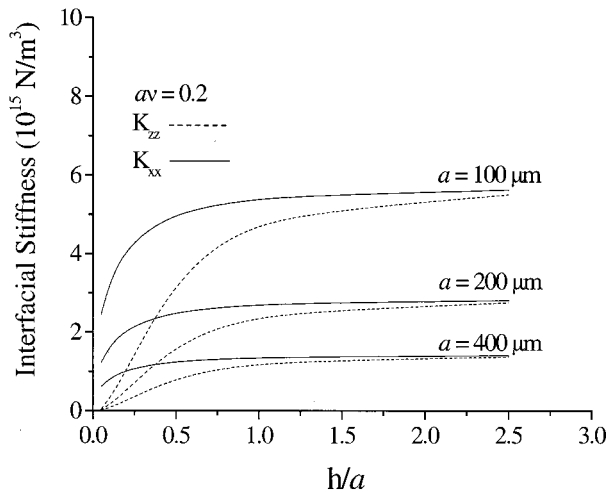


FIG. 4. Interfacial stiffness constants versus h/a for three values of the crack length: $a = 100 \mu\text{m}$, $a = 200 \mu\text{m}$, and $a = 400 \mu\text{m}$. The normalized crack density is constant, $av = 0.2$.

ing sections, the case of a crack between a layer and a substrate is considered. Equations (11) and (14) can be used when the thickness of the layer is infinite.

2. Cracked interface between a layer and a substrate

Having assumed that the cracks do not interact with each other, and having evaluated the compliance tensor of an isolated crack as a function of the ratio h/a , it is now possible to investigate the effect of the geometrical properties of the layered structure on the interfacial stiffness.

Figure 4 presents plots of the normal and transverse interfacial stiffness constants versus the ratio h/a for three values of the crack length a , again for a copper layer on a steel substrate. The normalized crack density, av , is equal to 0.2. The plots show that the interface becomes more compliant as the layer thickness decreases, and, as expected, the larger the crack length, the more compliant the interface. As for the isolated crack, both spring constants approach the same limit as the thickness of the layer increases. Because of the inverse proportionality between the stiffness constants and the crack compliance, K_N is always smaller than the transverse spring constant. No physical interphase layer model of an imperfect interface could simulate such a property of a cracked interface.

Figure 5 shows plots of the interfacial constants versus the ratio h/a for three values of the layer thickness, h . The normalized crack density is again equal to 0.2. Here, the behavior of the spring constants is markedly different from that shown in Fig. 4. Such dependence of the spring constants is explained as follows. If h is constant, an increase of the ratio h/a is obtained by decreasing the crack length, a . In order to maintain the *normalized* crack density, av , constant, the crack density, ν must be increased accordingly. Thus, as h/a increases the crack distribution changes its properties, becoming progressively a more dense distribution of smaller and smaller cracks. Figure 5, therefore, shows that among different interfaces with crack distributions having the same normalized crack density or, equivalently, the same cracked area, those having the smaller cracks are the stiffer. Results

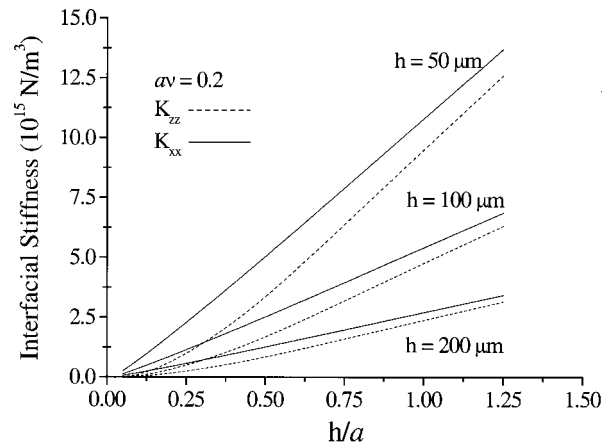


FIG. 5. Interfacial stiffness constants versus h/a for three values of the layer thickness: $h = 50 \mu\text{m}$, $h = 100 \mu\text{m}$, and $h = 200 \mu\text{m}$. The normalized crack density is constant, $av = 0.2$.

similar to those presented in Figs. 2, 4, and 5 have also been obtained for a system consisting of a stiffening layer on a substrate.⁸

In an attempt to identify the nature of interfacial imperfections, Nagy¹¹ considered the ratio between the reflection coefficients of longitudinal and shear waves, R_L and R_S , respectively, at normal incidence and at low frequencies, ω . He focused on interfaces between samples with identical material properties, and showed that the ratio $r = R_L/R_S$ is proportional to the ratio between the transverse and normal interfacial constants,

$$r = \lim_{\omega \rightarrow 0} \frac{R_L(\omega)}{R_S(\omega)} = \frac{K_T V_L}{K_L V_S}. \quad (15)$$

In Eq. (15), the symbols V_L and V_S are the phase velocity of the longitudinal and shear waves, respectively. Nagy found that r is smaller than 1 for distributions of volumetric imperfections, while it is larger than 1.5 for cracked interfaces. The extension of Nagy's analysis to interfaces between different materials and to structures more complex than that of an isolated interface should be straightforward. Although this is not a primary objective of this work, the behavior of the ratio K_T/K_N is presented as a function of h/a in Fig. 6. The

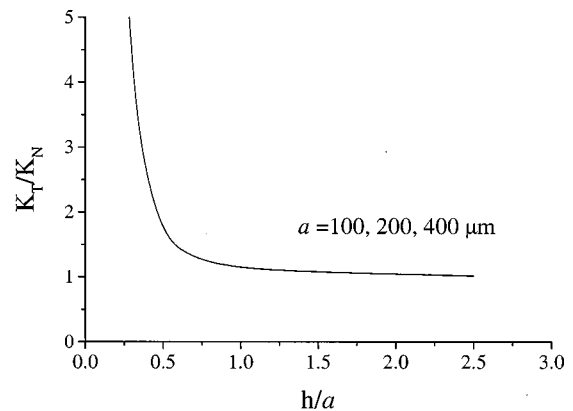


FIG. 6. Stiffness ratio versus h/a . The normalized crack density, av , is equal to 0.2.

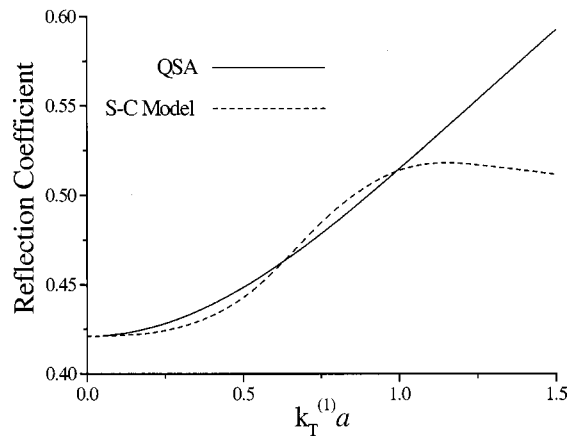


FIG. 7. Reflection coefficient from a cracked interface between two different media versus normalized crack dimension. The normalized crack density is $av=0.3$. QSA: —, self-consistent model: ---.

plot shows that the ratio K_T/K_N is always greater than one, and progressively increases as the layer thickness decreases.

III. ULTRASONIC WAVES AND CRACKED INTERFACES

A. Cracked interface between two different half-spaces

Qu¹² presented a theoretical model to evaluate the normal reflectivity of a cracked interface between two materials having different elastic properties. His approach was based on a differential self-consistent scheme and was used to obtain the reflection coefficient of a longitudinal wave as a function of the normalized crack length. The predictions of the self-consistent approach are reported in Fig. 7.¹³ Qu used the following values for the phase velocities, V_L and V_T , modulus of rigidity, μ , and the Poisson ratio, η , of the two materials:

Material 1: $V_L^1 = 6300$ m/s, $V_T^1 = 3100$ m/s,

$\mu_1 = 26.1$ GPa, $\eta_1 = 0.34$,

Material 2: $V_L^2 = 4600$ m/s, $V_T^2 = 2300$ m/s,

$\mu_2 = 48.3$ GPa, $\eta_2 = 0.33$.

The incident wave propagates in the half-space 1, and the normalized crack density is $av=0.3$. Figure 7 also presents the values of the same reflection coefficient according to the QSA. The agreement between the two approaches is reasonable up to values of $a/\lambda_L \sim 0.2$, where λ_L is the wavelength of the longitudinal incident wave. This relationship between the crack size and the wavelength of the interrogating wave will be used hereafter as an upper limit for the range of validity of the QSA.

B. Cracked interface between a layer and a substrate

In several applications the nondestructive assessment of the interface bond between a layer and its substrate is performed by using the lowest mode supported by the system. In this section the effect of a crack distribution on the phase velocity of this mode is briefly considered. In particular, the

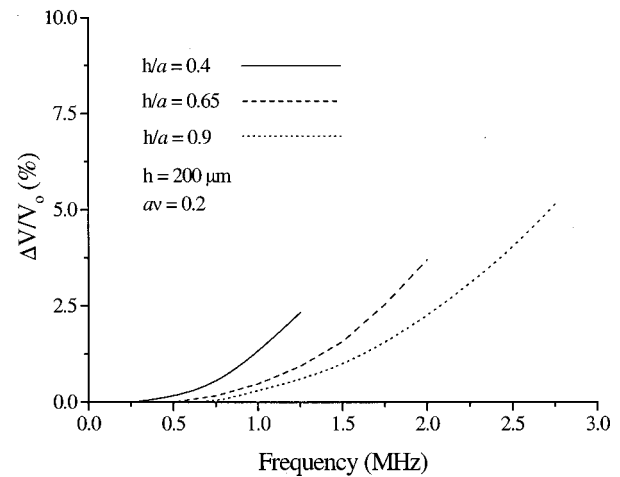


FIG. 8. Relative variation of the phase velocity of the first mode propagating along a copper layer on a steel substrate versus the ultrasonic frequency for three values of the ratio h/a : 0.4, 0.65, and 0.9. The layer thickness is $h=200$ μm . The normalized crack density is $av=0.2$.

investigation focuses on the relationship between the layer thickness, h , and the mode dispersion when the values of K_N and K_T depend on the ratio h/a .

Figure 8 shows the relative variation of the lowest mode's phase velocity as a function of the frequency for a system consisting of a copper layer on a steel substrate. The normalized crack density is $av=0.2$, and the thickness of the layer is $h=200$ μm . Three values for the ratio h/a are chosen: 0.4, 0.65, and 0.9. They correspond to three interfaces with increasing values of the spring constants. The reference velocity is that of the same mode propagating along the surface of a system having a perfectly bonded interface. The interfacial constants used to generate these plots are those presented in Fig. 5. The plots stop at the frequency where the product $a/\lambda_{\text{SAW}} \sim 0.2$, where λ_{SAW} is the wavelength of the propagating surface acoustic wave. The use of the QSA beyond this limit leads to considerable errors in the evaluation of the phase velocity of the guided mode. Figure 8 shows that the QSA can describe interfaces with crack distributions that may cause relative phase velocity variations not greater than 10%.

Figure 9 reports the percentage error in the predicted phase velocity caused by the use of the spring constants of an isolated interface between two infinite half-spaces in place of those considered in Fig. 8. The layered system is that considered in the previous figure. The error increases with the frequency and reaches the value of about 1% at the upper end of the range of validity of the QSA. Although a maximum error of 1% may be acceptable in some experimental situation, careful consideration must be given to a notable exception that is provided by line-focus beam acoustic microscopy of layered systems. This technique, in fact, allows measuring surface acoustic wave velocity with relative accuracy lower than 0.1%.¹⁴ Theoretical systematic errors greater than this limit may considerably affect the validity of any comparison between measured and theoretically predicted values of the phase velocity.

Finally, a system consisting of a stiffening NiT layer on

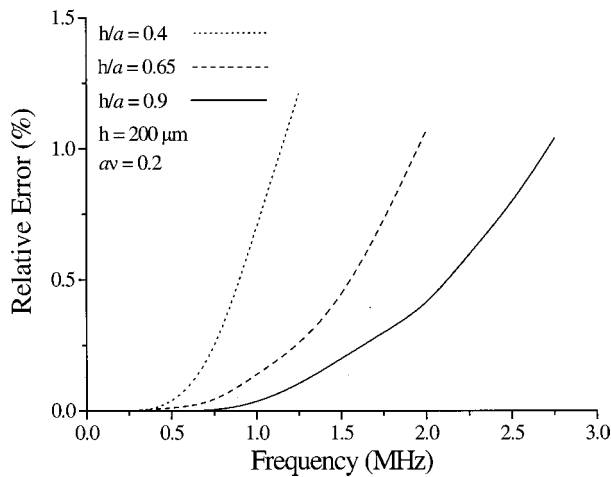


FIG. 9. Relative error versus ultrasonic frequency for the layered system of Fig. 7.

a steel substrate is considered. The elastic properties of these materials are the following:¹⁵

NiT: $\rho = 5200 \text{ Kg/m}^3$, $V_T = 6000 \text{ m/s}$, $V_L = 11021 \text{ m/s}$,

steel: $\rho = 7932 \text{ Kg/m}^3$, $V_T = 3260 \text{ m/s}$, $V_L = 5960 \text{ m/s}$.

Note that the ratio between the shear phase velocity of the layer and of the substrate is greater than 2. A well-known consequence of this fact¹⁶ is that there exists a cutoff frequency for the lowest mode supported by this system. At the cutoff frequency the velocity of the mode is equal to the velocity of a shear wave propagating in the substrate (horizontal line in Fig. 10). Figure 10 shows the three dispersion curves of the lowest mode for three interfaces. The first interface is characterized by a perfect bond, the second by a distribution of cracks with linear dimension $a \sim 150 \mu\text{m}$, and the third one by a distribution of cracks with length $a \sim 290 \mu\text{m}$. The normalized crack density is $av = 0.2$, and the layer thickness is $h = 100 \mu\text{m}$. The cut-off frequency for the lowest mode propagating along the perfect interface is $f_{\text{cutoff}} = 3.7$

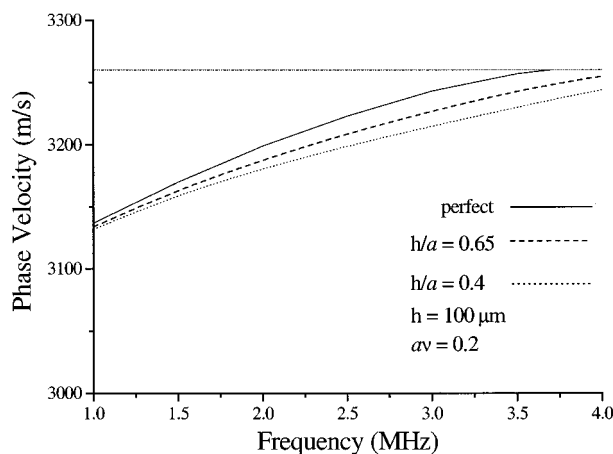


FIG. 10. Dispersion of the first mode supported by a stiffening NiT layer on a steel substrate versus ultrasonic frequency for a perfectly bonded interface (—), an interface with the ratio $h/a = 0.4$ (---), and an interface with the ratio $h/a = 0.65$ (----). The thickness of the layer is $h = 100 \mu\text{m}$, and the normalized crack density is $av = 0.2$.

MHz. The most interesting feature of this figure is that the lowest mode propagating along an interface affected by some degree of damage continues to propagate even at frequencies higher than 3.7 MHz.

IV. SUMMARY

The QSA boundary conditions for a cracked interface were derived from first principles. The derivation leads to a definition of the interfacial stiffness constants in terms of local quantities that describe the geometrical and the micro-mechanical properties of the distributed cracks.

In the case of an interface between different materials, it has been shown that the QSA provides an accurate description of the wave-interface interaction for ultrasonic frequencies such that $a/\lambda \leq 0.2$, and for values of the normalized crack density smaller than 0.5. In this work, the stiffness constants have been obtained under the assumption that the distributed cracks do not interact with each other. The extension of this modeling to include the interaction among first neighbors is conceptually straightforward. Multiple crack interaction is expected to introduce cross terms in the boundary conditions.

The spring constants of an interface between a layer and a substrate have been shown to depend, in general, on the layer thickness. Therefore, the interfacial constants can no longer be regarded as intrinsic properties of the interface, and determined only by the distributed imperfections. Neglecting their dependence of the structural properties of the hosting system may lead to relative errors on the predicted surface acoustic wave velocity of the order of 1%, which in some case may be not acceptable.

Finally, this investigation has shown that the first mode supported by a stiffening layer continues to propagate beyond its characteristic cutoff frequency when the interface contains a crack distribution.

APPENDIX A: AN INTERFACIAL CRACK BETWEEN A LAYER AND A SUBSTRATE

The solution to the problem of a crack lying along the interface between a layer and a substrate may be solved by considering first the stress due to a single interfacial displacement discontinuity, or *dislocation*.⁷ The magnitude of the displacement discontinuity is called the *Burgers vector*, b . Let the resolved components of the Burgers vector be b_y in the direction normal to the interface (into the layer) and b_x along the interface. The normal (N) and shear (S) stress σ arising at position x along the interface, due to an interfacial dislocation at position ξ , may then be written as¹⁷

$$\frac{\pi}{\Gamma} \begin{Bmatrix} \sigma_N(x) \\ \sigma_S(x) \end{Bmatrix} = \frac{b_x}{h} \begin{Bmatrix} \tilde{G}_{xN} + \hat{G}_{xN} \\ \tilde{G}_{xS} + \hat{G}_{xS} \end{Bmatrix} + \frac{b_y}{h} \begin{Bmatrix} \tilde{G}_{yN} + \hat{G}_{yN} \\ \tilde{G}_{yS} + \hat{G}_{yS} \end{Bmatrix}, \quad (\text{A1})$$

where h is the thickness of the layer, and the parameter Γ is defined above. The influence functions \tilde{G} in Eq. (A1) give the stress due to a dislocation between two half-planes, and are

$$\tilde{G}_{xN} = -\beta\pi\delta(\bar{x}-\bar{\xi}), \quad \tilde{G}_{xS} = 1/(\bar{x}-\bar{\xi}), \quad (\text{A2})$$

$$\tilde{G}_{yN} = 1/(\bar{x}-\bar{\xi}), \quad \tilde{G}_{yS} = +\beta\pi\delta(\bar{x}-\bar{\xi}),$$

where δ is the delta function and the overbar denotes normalization with respect to the layer thickness h , i.e., $\bar{x} = x/h$, etc. The influence functions \hat{G} in Eq. (A1) describe the influence of the layer free-surface. They may be obtained using Fourier transform theory,⁸ and are

$$\begin{aligned} \hat{G}_{xN} &= (1-\alpha) \int_0^\infty F_{xN} \frac{e^{-\lambda}}{\Delta} \cos[\lambda(\bar{x}-\bar{\xi})] d\lambda, \\ \hat{G}_{xS} &= (1-\alpha) \int_0^\infty F_{xS} \frac{e^{-\lambda}}{\Delta} \sin[\lambda(\bar{x}-\bar{\xi})] d\lambda, \\ \hat{G}_{yN} &= (1-\alpha) \int_0^\infty F_{yN} \frac{e^{-\lambda}}{\Delta} \sin[\lambda(\bar{x}-\bar{\xi})] d\lambda, \\ \hat{G}_{yS} &= -\hat{G}_{xN}, \end{aligned} \quad (\text{A3})$$

where

$$\begin{aligned} F_{xN} &= 2[(1+\beta)^2\lambda^2 + \beta]e^{+\lambda} - \beta(1+\alpha)e^{-\lambda}, \\ F_{xS} &= -\{2(1+\beta)\lambda[(1+\beta)\lambda - (1-\beta)] \\ &\quad + (1+\beta^2)\}e^{+\lambda} + (\alpha+\beta^2)e^{-\lambda}, \\ F_{yN} &= -\{2(1+\beta)\lambda[(1+\beta)\lambda + (1-\beta)] \\ &\quad + (1+\beta^2)\}e^{+\lambda} + (\alpha+\beta^2)e^{-\lambda}, \end{aligned} \quad (\text{A4})$$

and

$$\begin{aligned} \Delta &= (1-\beta^2)e^{2\lambda} + (\alpha^2 - \beta^2)e^{-2\lambda} \\ &\quad + 4(1+\beta)(\beta-\alpha)\lambda^2 - 2(\alpha-\beta^2). \end{aligned} \quad (\text{A5})$$

Note that these equations correspond to Eqs. (1)–(9) of Ref. 18, but with the misprints in Ref. 18 corrected. The integrals in Eq. (A3) may be evaluated numerically (replacing the upper infinite limit of integration with a value of about 15 will provide convergence and accuracy).

Consider now an interface crack lying along $[-L, +L]$. One assumes that the crack faces make contact near the crack-tips,⁶ and that the crack is open along $[-a, +b]$. Suppose also that the crack is loaded by a tension $T_o + \Delta T$ and a shear ΔS . The opening displacements of the crack are modeled as a continuous distribution of dislocations along its length. The boundary conditions along the crack are that the net shear traction at any point along the crack, due to the load and due to dislocations, must be zero (assuming that the crack-faces make frictionless contact in the contact-zones), and that the net normal traction must be zero along the open portion $[-a, +b]$. These conditions then lead, integrating Eq. (A1), to the dual integral equations

$$\begin{aligned} -\frac{T_o + \Delta T}{\Gamma} &= -\beta B_x(x) + \frac{1}{\pi h} \int_{-L}^{+L} B_x(\xi) [\tilde{G}_{xN} + \hat{G}_{xN}] d\xi \\ &\quad + \frac{1}{\pi h} \int_{-a}^{+b} B_y(\xi) [\tilde{G}_{yN} + \hat{G}_{yN}] d\xi, \\ -\frac{\Delta S}{\Gamma} &= +\beta H(\cdot) B_y(x) + \frac{1}{\pi h} \int_{-L}^{+L} B_x(\xi) [\tilde{G}_{xS} + \hat{G}_{xS}] d\xi \\ &\quad + \frac{1}{\pi h} \int_{-a}^{+b} B_y(\xi) [\tilde{G}_{yS} + \hat{G}_{yS}] d\xi, \end{aligned} \quad (\text{A6})$$

where $H(\cdot) = H(x+a) - H(x-b)$ is the Heaviside function, and B_x, B_y are the shear and normal crack-surface displacement gradients (dislocation densities). Here, it is assumed that the contact zones are infinitesimal in length (contact lengths $< 10^{-4}L$). It can be shown that this will be true provided $(T_o + \Delta T)/\Delta S > 1.13$. The equations are solved by assuming that the shear densities B_x are square-root singular near $\pm L$ and that the normal densities B_y are so near $+a$ and $-b$ (even though the normal densities are known to be bounded there). The singularity in B_y is then driven to zero. The contact lengths are unknown: an initial guess may be made of the position of a and b by using the bonded half-plane solution.¹⁹ The equations are solved in an iterative manner using the Gauss–Chebyshev quadrature formulas (see Ref. 7, Table 2.2, case I) and Krenk's interpolation formulas (see Ref. 7, Table 2.4, case I). Full details are given elsewhere.⁸

¹J. M. Baik and R. B. Thompson, "Ultrasonic scattering from imperfect interfaces: a quasi static model," *J. Nondestruct. Eval.* **4**, 177–196 (1984).

²L. Singher, Y. Segal, E. Segal, and J. Shamir, "Considerations in bond strength evaluation by ultrasonic guided waves," *J. Acoust. Soc. Am.* **96**, 2497–2505 (1996).

³L. Adler, M. de Billy, G. Quentin, M. Talamant, and P. B. Nagy, "Evaluation of friction-welded aluminum-steel bonds using dispersive guided modes of a layered substrate," *J. Appl. Phys.* **68**, 6072–6076 (1990).

⁴W. Huang, S. Brisuda, and S. I. Rokhlin, "Ultrasonic wave scattering from fiber-matrix interphase," *J. Acoust. Soc. Am.* **97**, 807–817 (1995).

⁵M. Kachanov, "Effective elastic properties of cracked solids: critical review of some basic concepts," *Appl. Mech. Rev.* **45**, 304–335 (1992).

⁶M. Comninou, "The interface crack," *J. Appl. Mech.* **44**, 631–634 (1977).

⁷D. A. Hills, P. A. Kelly, D. N. Dai, and A. M. Korsunsky, *Solution of Crack Problems* (Kluwer Academic, Dordrecht, 1996).

⁸P. A. Kelly and C. Pecorari, "The interaction between ultrasonic waves and imperfect interfaces in layered media," Report No. OUEL 2148/97, Dept. of Engineering Science, Oxford University (1998).

⁹J. D. Achenbach and A. N. Norris, "Loss of specular reflection due to nonlinear crack-face interaction," *J. Nondestruct. Eval.* **3**, 229–239 (1982).

¹⁰J. Dundurs, "Effect of elastic constants on stress in a composite under plane deformation," *J. Cryst. Growth* **1**, 310–321 (1967).

¹¹P. B. Nagy, "Ultrasonic classification of imperfect interfaces," *J. Nondestruct. Eval.* **11**, 127–139 (1992).

¹²J. Qu, "Reflection of elastic waves by a randomly cracked interface," in *Ultrasonic Characterization and Mechanics of Interfaces*, AMD-Vol. 177, AMSE, 223–238 (1993).

¹³The data of the reflection coefficient from Ref. 11 have been obtained from Fig. 4 of that reference.

¹⁴J. Kushibiki and M. Arakawa, "A method for calibrating the line-focus-beam acoustic microscopy system," *IEEE Trans. Ultrason. Ferroelectr. Freq. Control* **45**, 421–430 (1998).

- ¹⁵O. Lefeuvre, P. Zinin, G. A. D. Briggs, and A. G. Every, "Surface wave dispersion beyond cutoff for a fast layer on a slow substrate," *Appl. Phys. Lett.* **72**, 856–857 (1998).
- ¹⁶G. W. Farnell and E. L. Adler, "Elastic wave propagation in thin layers," in *Physical Acoustics*, edited by W. P. Mason and R. N. Thurston (Academic, New York, 1972), pp. 35–127.
- ¹⁷P. A. Kelly, J. J. O'Connor, and D. A. Hills, "The stress field due to a dislocation in layered media," *J. Phys. D* **28**, 530–534 (1995).
- ¹⁸M. Comninou and J. Dundurs, "Partial closure of cracks at the interface between a layer and a half-space," *Eng. Fract. Mech.* **18**, 315–323 (1983).
- ¹⁹A. K. Gautesen and J. Dundurs, "The interface crack under combined loading," *J. Appl. Mech.* **55**, 580–586 (1988).

Integration of modeling and acoustic microscopy measurements for thin films

Zhiqi Guo and J. D. Achenbach^{a)}

Center for Quality Engineering and Failure Prevention, Northwestern University, Evanston, Illinois 60208

Anita Madan, Keith Martin, and M. E. Graham

Advanced Coating Technology Group, Northwestern University, Evanston, Illinois 60208

(Received 27 April 1999; accepted for publication 21 January 2000)

A model for measuring the $V(z)$ curve by line-focus acoustic microscopy contains the reflectance function of the specimen as a principal component. In this paper the reflectance function has been analyzed for multilayered thin films on a substrate for both fast-on-slow and slow-on-fast systems. The phase velocities of modes of surface acoustic wave propagation and their associated mode reflection coefficients can be obtained from the reflectance function. This information can be used together with estimates of the elastic constants to determine suitable frequency ranges for measuring the $V(z)$ curve. Minimization of the difference between phase velocities obtained from measured and calculated $V(z)$ curves is used to determine the elastic constants. Results are presented for TiN films on M2 high-speed steel substrates. © 2000 Acoustical Society of America.

[S0001-4966(00)05504-1]

PACS numbers: 43.35.Ns [HEB]

INTRODUCTION

Thin film coatings are widely used in surface engineering to extend the life or enhance the performance of components. For example, transition metal carbides and nitrides of the IV to VI groups of the periodic table have extremely high melting points, extreme hardness, excellent high-temperature strength and good corrosion resistance. They are widely used for thin film coatings to produce wear resistant surfaces, to provide corrosion protection against harsh environments, and for other applications.¹

Basically, a coated component, here referred to as a film/substrate system, is a layered structure, with a single- or multiple-layered thin film deposited on a substrate. Film/substrate systems are usually prepared by different deposition techniques, such as physical vapor deposition and chemical vapor deposition. The mechanical properties of the thin film, and the adhesion between coating layers and between the coating and the substrate, determine its functional characteristics. These properties are very sensitive to a number of factors, which are determined by the deposition technique and by processing parameters. The thin-film properties may be quite different from bulk properties. Therefore there is a need for a suitable nondestructive technique to characterize these thin-film coatings.

For acoustic characterization, film/substrate systems are commonly classified into two categories. If the coating layer has larger shear wave velocity than the substrate, it is a stiffening layer and it is referred to as a fast-on-slow system. The opposite case is referred to as a slow-on-fast system. Until now most research has been concerned with slow-on-fast systems. There has been less work on fast-on-slow systems and surface wave propagation for this case is less well understood.^{2,3} In certain frequency ranges it has been diffi-

cult to obtain a useful $V(z)$ curve for thick or very stiff thin films.⁴

This paper has a theoretical and an experimental part. A systematic theoretical study is carried out on surface acoustic waves (SAW's) propagating in isotropic thin-film systems with different material combinations, such as single- or multiple-layered polycrystalline ceramic thin-film coatings on an isotropic substrate. The frequency dependence of the phase velocities of SAW modes has been investigated for the purpose of improving the efficiency of experimental characterization of isotropic film/substrate systems by line-focus acoustic microscopy (LFAM). Corresponding experimental studies have been carried out.

A theoretical measurement model for the $V(z)$ curve, measured by line-focus acoustic microscopy, contains the reflectance function of the film/substrate system as a principal component. By using the theoretical reflectance function model for parametrical studies, the phase velocities of the possible SAW modes and their corresponding mode reflection coefficients can be predicted for an arbitrary film/substrate system. Therefore, the reflectance function can be used as a predictive tool to estimate in which frequency range a wave mode may be picked up by acoustic microscopy. By applying the $V(z)$ measurement model, the phase velocities can be calculated for a film/substrate system at different frequencies. For trial values of the unknown elastic constants of the thin film, these theoretically calculated phase velocities have been used in an iterative manner together with phase velocities obtained from $V(z)$ measurements to inversely determine the elastic constants of thin films by minimization of the deviation between theoretical and experimental results. The method has been applied to determine elastic constants for TiN thin-film coatings on M2 high-speed steel substrates.

^{a)}Electronic mail: achenbach@nwu.edu

I. THEORETICAL ANALYSIS

A. Reflectance function

Figure 1 shows the configuration of a layered medium with N layers on a substrate. A time-harmonic plane wave is incident from a coupling fluid, at arbitrary incident angle θ . The response of the layered medium to the incident time-harmonic plane wave can be expressed in terms of a reflectance function. This reflectance function carries information on the layered medium, including its mechanical properties, the thickness of each layer, and the interface conditions. A general theoretical model has been developed which quantitatively relates the reflectance function to these characteristic properties of the layered specimen. The model has been developed by completing the following steps:

(I) Find general wave field solutions for each layer to relate the state vector at an arbitrary location to the material properties and the geometrical parameters. The state vector consists of the displacement and stress components that will be used in the boundary conditions. For isotropic layers it is defined as

$$[S(x_3)] = [u_1 \ u_3 \ \sigma_{33} \ \sigma_{31}]^T.$$

By considering displacement solutions for a homogeneous layer as the superposition of downgoing and upgoing bulk waves within the layer, i.e., as

$$u_i = \sum_{m=1}^4 U_m \phi_m^i \exp[ik_x(x_1 + \alpha_m x_3 - ct)],$$

where $\alpha_1 = \alpha$, $\alpha_2 = -\alpha$, $\alpha_3 = \beta$, $\alpha_4 = -\beta$, ϕ^i is the corresponding eigenvector of the displacement component u_i , $\phi^i = (\phi_1^i, \phi_2^i, \phi_3^i, \phi_4^i)$, and U_m ($m=1,2,3,4$) are unknown constants. The state vector can be derived as

$$[S(x_3)] = [D(x_3)] \cdot [U],$$

where $[D(x_3)]$ is defined as

$$[D(x_3)] = \begin{bmatrix} 1 & 1 & 1 & 1 \\ \alpha & -\alpha & \frac{-1}{\beta} & \frac{1}{\beta} \\ i\gamma\mu k_x & i\gamma\mu k_x & -2i\mu k_k & -2i\mu k_x \\ 2i\mu\alpha k_x & -2i\mu\alpha k_x & \frac{i\gamma\mu k_x}{\beta} & \frac{-i\gamma\mu k_x}{\beta} \end{bmatrix} \times \begin{bmatrix} e^{ik_x\alpha x_3} & 0 & 0 & 0 \\ 0 & e^{-ik_x\alpha x_3} & 0 & 0 \\ 0 & 0 & e^{ik_x\beta x_3} & 0 \\ 0 & 0 & 0 & e^{-ik_x\beta x_3} \end{bmatrix}. \quad (1)$$

In these equations

$$\alpha = \left(\frac{c^2}{c_L^2} - 1 \right)^{1/2}; \quad \beta = \left(\frac{c^2}{c_T^2} - 1 \right)^{1/2}; \quad \gamma = \frac{c^2}{c_T^2} - 2;$$

$$c_L^2 = \frac{\lambda + 2\mu}{\rho}; \quad c_T^2 = \frac{\mu}{\rho}.$$

(II) For each layer derive the transfer matrix $[T_n]$, which relates the state vector at one side of the layer, say $x_3=0$, $[S_n^-]$, to that at the other side, say $x_3=h_n$, $[S_n^+]$, i.e.,

$$[S_n^-] = [T_n] \cdot [S_n^+],$$

where

$$[T_n] = [D_n(x_3=0)] \cdot [D_n(x_3=h_n)]^{-1} \quad (n=1,2,\dots,N).$$

(III) Model a nonperfect interface by introducing characteristic parameters to represent the properties of the interface, and find its transfer matrix $[T_{\text{int}}]$ which relates the state vector at one side of the interface to that at the other side:

$$[S_{\text{int}}^-] = [T_{\text{int}}] \cdot [S_{\text{int}}^+].$$

(IV) Apply the transfer matrix method to construct a transfer matrix $[G]$ to relate the state vector at the top of the N -layer medium (at $x_3=0$) to that at the bottom (at $x_3=d$) by invoking the continuity conditions on both sides of each layer, to obtain

$$[S(x_3=0)] = [G] \cdot [S(x_3=d)],$$

where

$$[G] = [T_1] \cdot [T_2] \cdots [T_{\text{int}}] \cdots [T_N].$$

Here only one nonperfect interface within the N layers is considered. $[G]$ is the transfer matrix of the N -layer medium. (V) Incorporate the derived state vector in the coupling fluid at $x_3=0$, $[S^f(x_3=0)]$,

$$[S^f(x_3=0)] = \begin{bmatrix} 1 & 1 \\ \alpha_f & -\alpha_f \\ ik_x\rho_f c^2 & ik_x\rho_f c^2 \\ 0 & 0 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ R \end{bmatrix},$$

where $\alpha_f = \left(\frac{c^2}{c_f^2} - 1 \right)^{1/2}$, (2)

and in the substrate at $x_3=d$, $[S^s(x_3=d)]$,

$$[S^s(x_3=d)] = [T_{\text{sub}}] \cdot [U_1 \ 0 \ U_3 \ 0]^T,$$

where $[T_{\text{sub}}]$ is the transfer matrix of the substrate,

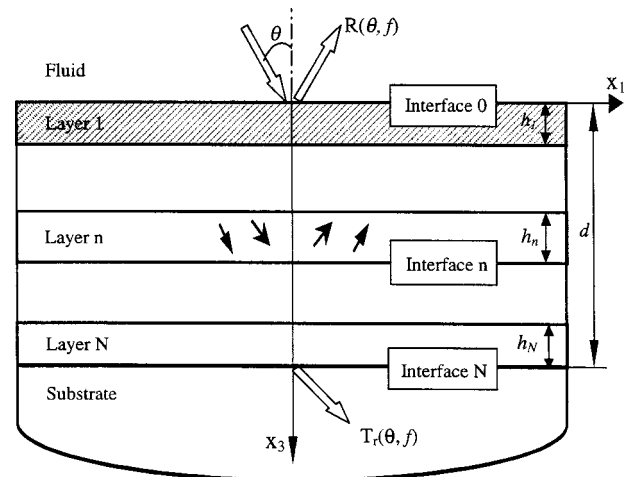


FIG. 1. Schematic configuration of a multiple-layered medium with N layers on a substrate.

$$[T_{\text{sub}}] = [D_{\text{sub}}(x_3 = 0)].$$

Next enforce the boundary conditions at these two extreme interfaces to obtain the following coupling equations:

$$\begin{bmatrix} T_{21} & T_{23} & \alpha_f \\ T_{31} & T_{33} & -ik_x \rho_f c^2 \\ T_{41} & T_{43} & 0 \end{bmatrix} \cdot \begin{bmatrix} U_1 \\ U_3 \\ R \end{bmatrix} = \begin{bmatrix} \alpha_f \\ ik_x \rho_f c^2 \\ 0 \end{bmatrix}. \quad (3)$$

Finally, solve these equations to obtain the reflectance function $R(\theta, f)$ for a system of isotropic layers as

$$R(\theta, f) = \frac{A - \beta B}{A + \beta B}, \quad (4)$$

where

$$\beta = \frac{i \rho_f \omega^2}{\alpha_f k_x}, \quad A = \begin{vmatrix} T_{31} & T_{33} \\ T_{41} & T_{43} \end{vmatrix} = T_{62}^*,$$

$$B = \begin{vmatrix} T_{21} & T_{23} \\ T_{41} & T_{43} \end{vmatrix} = T_{52}^*.$$

Here ρ_f and c_f are the density and the velocity of the coupling fluid, and f and ω are the frequency and the angular frequency of the incident wave with incident angle θ , and k_x is the tangential component of the wave vector, where

$$k_x = k_f \sin \theta, \quad k_f c_f = k_x c = \omega, \quad (5)$$

$$k_z = \sqrt{k_f^2 - k_x^2}, \quad c = \frac{c_f}{\sin \theta}.$$

Here, T_{ij} are the elements of the global transfer matrix $[T]$,

$$[T] = [G] \cdot [T_{\text{sub}}].$$

For the numerical calculations, the delta-matrix⁵ reformulation is introduced to avoid the “precision problem.” For a 4×4 matrix $[T]$, there are a total of 36 such 2×2 subdeterminants that can themselves be arranged as a 6×6 matrix, which is referred to as the delta matrix of $[T]$ and denoted as $[T]^*$. Elements of the delta matrix are free from the numerical precision difficulties. The symbolic calculations can be implemented by using the Mathematica 3.0 software package. We have

$$[T]^* = [G]^* \cdot [T_{\text{sub}}]^*$$

$$= [T_1]^* \cdot [T_2]^* \cdots [T_{\text{int}}]^* \cdots [T_N]^* \cdot [T_{\text{sub}}]^*.$$

The reflectance function given by Eq. (4) can be used for wave mode identification. The vanishing of its denominator, i.e.,

$$A + \beta B = 0,$$

is exactly the characteristic equation of the layered medium. Solutions for the velocity and the amplitude of the possible wave modes can be obtained by numerically solving this equation. Setting $\beta = 0$ takes out the effects of the fluid loading, and thus the equation

$$A = 0$$

can be used to identify the free wave modes.

For a given frequency, the reflectance function is a function of the incident angle, θ . The reflectance function shows distinct behavior when the incident angle equals the critical angle associated with a possible wave mode. By virtue of the relation between the incident angle θ and the phase velocity c shown in Eq. (5), the reflectance function can be used to obtain the phase velocities of the wave modes for the layered system at a fixed frequency. The reflectance function undergoes a -2π phase shift and its real part undergoes a jump from around $+1$ towards -1 as the incident angle θ passes a critical angle θ_{cr} , which corresponds to a wave mode. Using this criterion to determine the critical angles, the phase velocity of a wave mode can be calculated using Eq. (6), as

$$c = \frac{c_f}{\sin \theta_{\text{cr}}}. \quad (6)$$

The magnitude of the reflectance function at the critical angle θ_{cr} indicates how strongly the wave mode will be reflected. For convenience, that magnitude will be referred to as the mode reflection coefficient. A strongly reflected SAW mode with a large mode reflection coefficient can be picked up easily by acoustic microscopy. The results show that the theoretical mode reflection coefficient of a true surface wave mode is unity and that of a pseudo-SAW mode that radiates energy into the substrate is somewhere between 0 and 1. A low mode reflection coefficient indicates that the wave mode is weakly reflected and will be more difficult to be picked up by the LFAM for generation of a $V(z)$ curve.

In this paper, the reflectance function model has been employed to investigate some configurations and to perform detailed parametrical studies. The model can be used to predict the SAW modes that can be generated, and to calculate their velocities, as well as their mode reflection coefficients for an arbitrarily configured layered medium with perfect or imperfect interfacial conditions. These calculations can be done for varying frequencies to obtain the theoretical dispersion curves of the SAW modes.

B. $V(z)$ measurement model

It has been shown that quantitative acoustic microscopy is a suitable technique for nondestructive characterization of thin films. The technique is based on the measurement of the so-called $V(z)$ curve, which is a recording of the voltage output of the transducer when the distance between the lens and the specimen is decreased. Oscillations appear in the $V(z)$ curves due to interference between the rays specularly reflected by the specimen and the rays associated with propagation of a leaky SAW on the specimen. The phase velocity of a leaky surface wave, c_{LSAW} , can be determined from the period, Δz , of the periodic oscillations, using a well-known relationship based on ray considerations:^{6,7}

$$c_{\text{LSAW}} = c_f \cdot \left[1 - \left(1 - \frac{c_f}{2f \cdot \Delta z} \right)^2 \right]^{-1/2} \approx \sqrt{c_f \cdot f \cdot \Delta z}, \quad (7)$$

where c_f is the wave velocity in the coupling fluid and f is the operating frequency. Equation (7) implies that the measurable Δz must satisfy

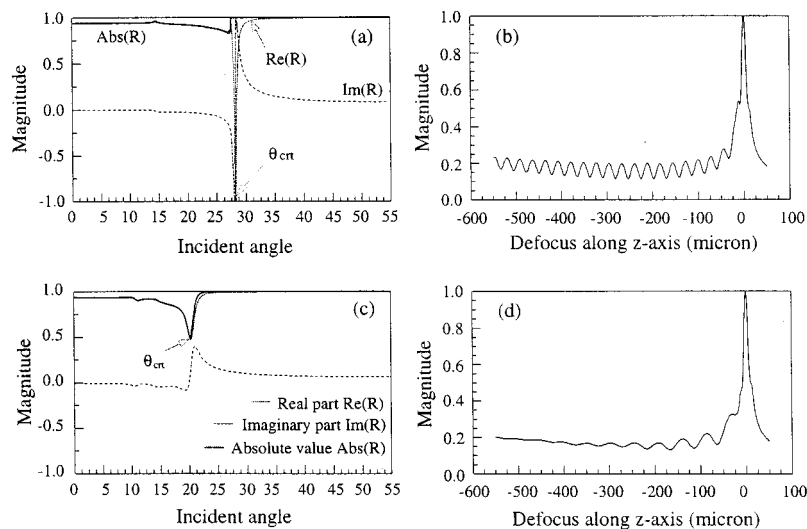


FIG. 2. (a) and (c) Calculated reflectance functions for a strongly reflected wave mode and a weakly reflected one, and (b) and (d) their corresponding calculated $V(z)$ curves.

$$\frac{c_f}{2f \cdot \Delta z} \ll 1.$$

The period Δz can be accurately extracted from the measured $V(z)$ curve by spectrum analysis, as discussed in detail in Ref. 7. Therefore, by measuring $V(z)$ curves for a specimen at different frequencies or at the same frequency but different wave propagation directions, the dispersion curves or directional variations of the leaky surface wave velocity can be measured.

By numerically simulating the $V(z)$ measurement procedure, a $V(z)$ measurement model can be developed to calculate theoretical $V(z)$ curves for a layered specimen of arbitrary configuration for use with a line-focus acoustic microscope. In this paper, we use the approach proposed by Achenbach *et al.* for synthesizing $V(z)$ curves.⁸ In that approach, the output of the line-focus acoustic probe is expressed as a Fourier integral over the product of the characteristic functions of the acoustic lens, $L_1(k_x)$ and $L_2(k_x)$, and the reflectance function of the fluid-loaded specimen $R(k_x)$, as expressed by

$$V(z) = \int_{-\infty}^{+\infty} L_1(k_x) \cdot L_2(k_x) \cdot R(k_x) \cdot \exp(2ik_z z) \cdot dk_x, \quad (8)$$

where (k_x, k_z) defines the wave vector, $k_z = \sqrt{k_f^2 - k_x^2}$, $k_f = \omega/c_f$, and $k_x = \omega/c_{\text{LSAW}}$.

Theoretical calculations and experimental observations show that LFAM cannot pick up any SAW mode with a mode reflection coefficient smaller than around 0.15. A strongly reflected SAW mode can be easily picked up by the LFAM and its velocity can be extracted with high accuracy. The $V(z)$ measurement will be less precise for a weakly reflected surface wave mode. There is a significant decrease in the number and magnitude of the oscillations of the $V(z)$ curve for a weakly reflected SAW mode.

Figures 2(a) and (c) show the calculated reflectance functions for a strongly reflected wave mode and a weakly reflected one for a specimen. Figures 2(b) and (d) show the calculated $V(z)$ curves corresponding to the reflectance functions shown in Fig. 2(a) and (c), respectively. Clearly, the

number and magnitude of the oscillations in the $V(z)$ curve for a weakly reflected SAW mode decrease significantly, and this decrease gives rise to a loss of accuracy in the extracted velocity.

C. SAW propagating in isotropic systems with different material combinations

By applying the reflectance function model and the $V(z)$ measurement model, SAW propagation in thin-film systems is investigated for different material combinations. The focus is on the frequency dependence of the phase velocity and the mode reflection coefficient associated with the possible SAW modes, and with their consequences for LFAM-based measurements of layered isotropic systems.

1. Slow-on-fast systems

Figure 3(a) shows the phase velocities of SAW modes versus the product of the frequency f and the film thickness d , calculated from the reflectance function, together with the phase velocities calculated from the simulated $V(z)$ curves for a slow-on-fast system (layer shear wave velocity V_S is smaller than the substrate shear wave velocity V_S^0). Figure 3(b) shows the calculated mode reflection coefficients associated with the SAW modes shown in Fig. 3(a). Generally, the existence of the slow layer causes a decrease of the phase velocity of the SAW with increasing $f \cdot d$. A true surface wave mode always exists, and its velocity decreases from the Rayleigh wave velocity of the substrate to the Rayleigh wave velocity of the layer. Its mode reflection coefficient is unity. Other kinds of wave modes (referred to as pre-Rayleigh waves) may exist. At small $f \cdot d$ the lowest pre-Rayleigh wave mode propagates at higher speed than the shear wave velocity of the substrate. Its mode reflection coefficient is low because it radiates energy into the substrate. The mode reflection coefficient associated with this pre-Rayleigh wave increases to unity as $f \cdot d$ increases, and the velocity decreases toward the Rayleigh wave velocity of the layer when $f \cdot d$ becomes very large and the layer functions as a half-space. Higher pre-Rayleigh wave modes show analogous behavior. Still higher wave modes propagate with velocities

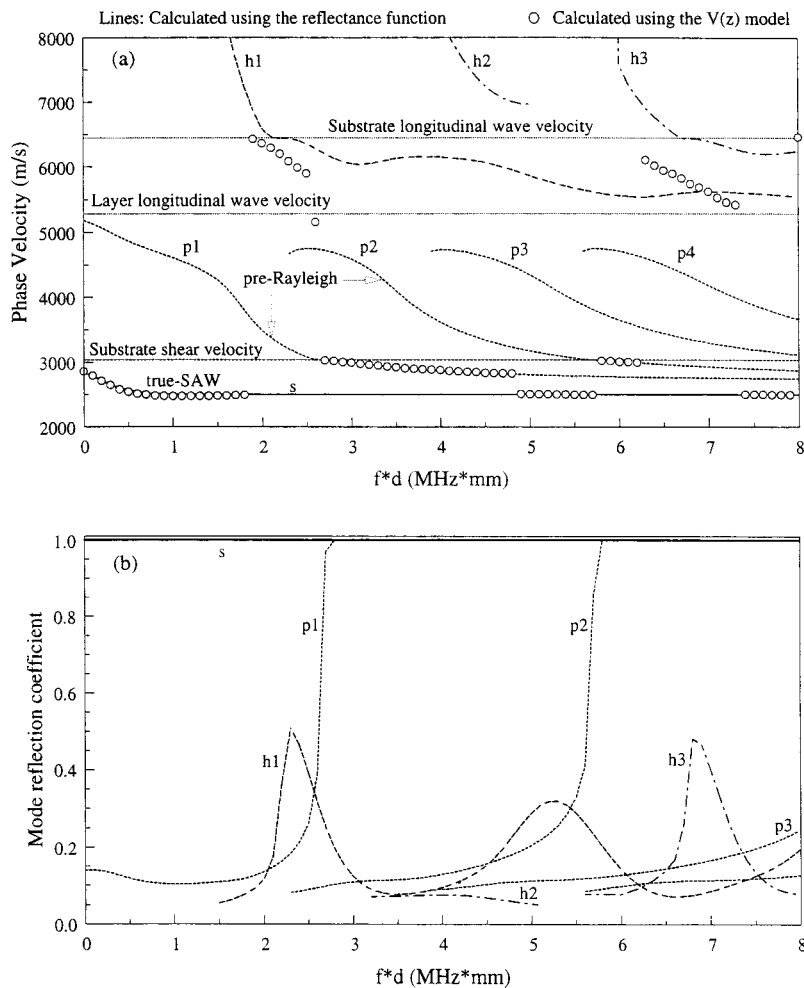


FIG. 3. (a) Phase velocities of the SAWs calculated from the reflectance function, together with phase velocities obtained from simulated $V(z)$ curves for a 225-MHz LFAM, for a slow-on-fast system, $\text{Ti}(d)/\text{Al}$. (b) Calculated mode reflection coefficients associated with the corresponding SAW modes shown in (a).

larger than the longitudinal wave velocity of the layer. These generalized Lamb wave modes, which appear after certain cutoff points, show a decreasing phase velocity and have lower mode reflection coefficients.

At small $f*d$ (here, $f*d < 2.0$), only the true SAW is reflected strong enough to be accurately picked up by the LFAM. The pre-Rayleigh wave modes can be picked up accurately only after they start to propagate at velocities lower than the substrate shear wave velocity and become sufficiently strongly reflected with reflection coefficient unity. When there are two modes with sufficiently high mode reflection coefficients the $V(z)$ curve tends to pick out the one with the higher phase velocity.

More than two modes can be picked up simultaneously by an acoustic microscope in ranges where more wave modes have sufficiently large mode reflection coefficients (> 0.15). This is, for example, happening in the ranges of $f*d = 2.2\text{--}2.8$ and $6.3\text{--}7.2$ for three sufficiently strong wave modes. For these cases, care must be taken to accurately extract multiple wave modes from the corresponding $V(z)$ curve. With the usual approach the calculated velocities are not correct as shown in Fig. 3(a). An interactive algorithm to perform $V(z)$ curve analysis to extract overlapping multiple SAW modes for such cases has been developed, but this algorithm will be described elsewhere.

2. Fast-on-slow systems

Figure 4(a) shows the phase velocities of SAW modes calculated from the reflectance function, together with the phase velocities calculated from the simulated $V(z)$ curves for three stiffening systems ($V_S > V_S^0$) with different shear wave velocity ratios (V_S/V_S^0). Figure 4(b) shows the calculated mode reflection coefficients associated with the SAW modes shown in Fig. 4(a). Generally, the existence of the stiffening layer causes an increase of the phase velocity of the SAW. There is a cutoff associated with the shear wave velocity of the substrate. Below the cutoff, a single surface wave mode exists, and as $f*d$ increases, its velocity increases from the Rayleigh wave velocity of the substrate, until the cutoff is reached at the shear wave velocity of the substrate. In this range the calculated mode reflection coefficient associated with this wave mode is unity. This means that it is a true surface wave mode and it only leaks energy into the coupling fluid and is strongly reflected. Beyond the cutoff, the SAW becomes a weakly reflected pseudo leaky SAW, and it radiates energy both into the fluid and the substrate. The mode reflection coefficient associated with the pseudo-SAW mode decreases dramatically just after the cutoff, and then again increases gradually. Simultaneously with the increase of the mode reflection coefficient, the pseudo-SAW mode transforms into a quasi-Rayleigh wave mode as $f*d$ increases. Finally, the quasi-Rayleigh wave mode

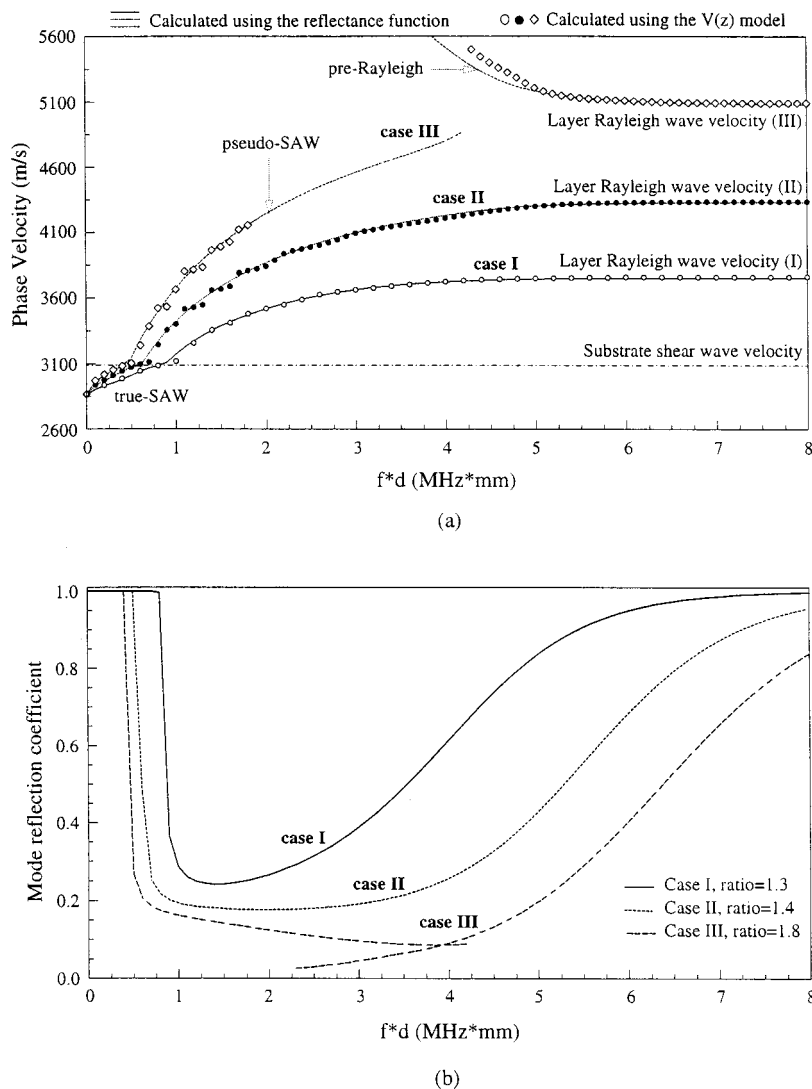


FIG. 4. (a) Phase velocities of the SAWs calculated from the reflectance function, together with phase velocities obtained from simulated $V(z)$ curves for a 225-MHz LFAM, for three fast-on-slow systems with different shear velocity ratios, $V_s/V_s^0 = 1.3$, 1.4, and 1.8, respectively. (b) Calculated mode reflection coefficients associated with the corresponding SAW modes shown in (a).

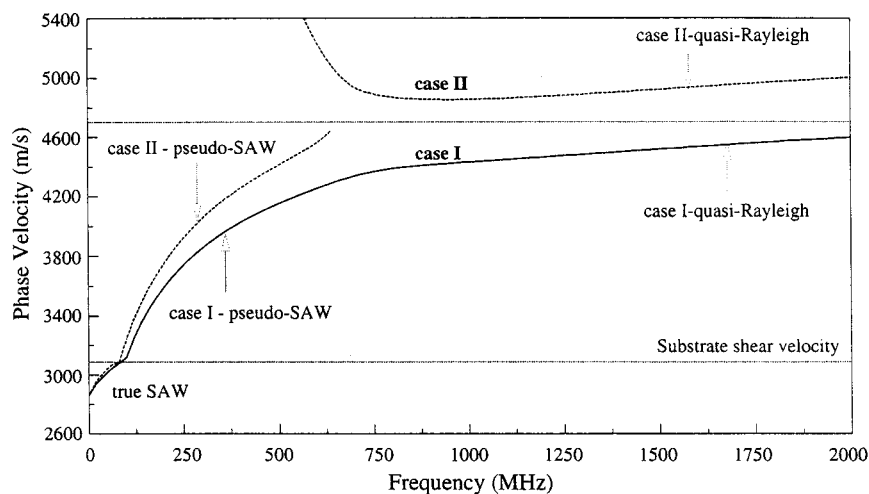
evolves into a Rayleigh wave on the layer as $f \cdot d$ becomes very large and the layer functions as a half-space. For a system with a still stiffer film (case III) the pseudo-SAW terminates. For this case a new wave mode (referred to as a pre-Rayleigh wave) occurs and higher wave modes may also appear. This pre-Rayleigh wave mode propagates initially at higher speed than the Rayleigh wave velocity of the layer, and it is weakly reflected with a low mode reflection coefficient. As $f \cdot d$ increases its mode reflection coefficient also increases, while the pre-Rayleigh wave mode transforms into a Rayleigh wave on the layer. Higher wave modes generally behave like Lamb wave modes. The behavior represented by case III generally occurs for $V_s/V_s^0 > 1.5$, which is consistent with the result of Ref. 9.

Because the pseudo-SAW has a very low mode reflection coefficient just beyond the cutoff, it may be invisible to the LFAM when the measurements are performed in this range. Even if it is possible to pick up the pseudo-SAW, the measurements are less accurate than those performed in other ranges where the SAW is strongly reflected. As indicated in Fig. 4(a), the deviation between the phase velocities calculated from the reflectance function and from the simulated $V(z)$ curves of the LFAM increases after the cutoff fre-

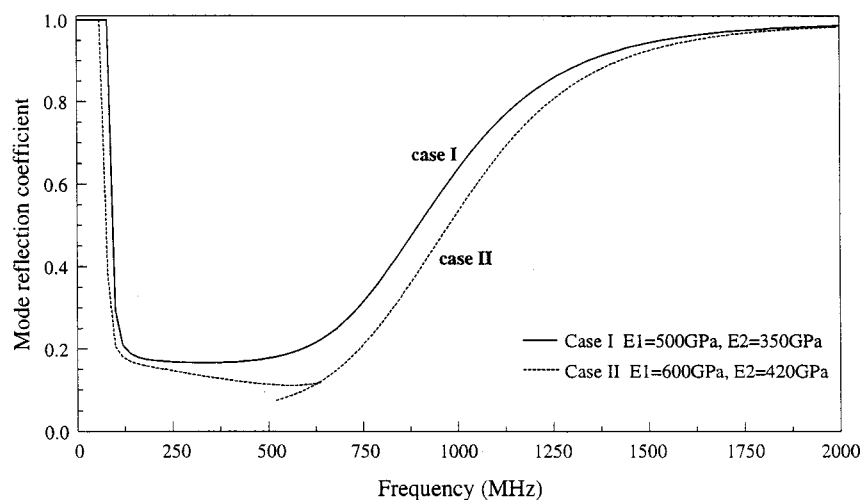
quency and within the range where the pseudo-SAW is weakly reflected.

Similar results have been obtained for multiple-layered stiffening thin-film systems. Figure 5(a) shows the phase velocities versus the frequency of the SAW modes, calculated from the reflectance function for a stiffening double-layer coating on an M2 high-speed steel (M2) substrate. The two coating layers are ZrN ($1 \mu\text{m}$, $\rho = 7.35 \text{ g/cm}^3$, $\nu = 0.20$, E_1) and ZrC ($5 \mu\text{m}$, $\rho = 6.56 \text{ g/cm}^3$, $\nu = 0.19$, E_2) with different Young's moduli, E_1 and E_2 . Figure 5(b) shows the calculated mode reflection coefficients associated with the corresponding SAW modes shown in Fig. 5(a). For case II, as the frequency becomes larger, the pseudo-SAW terminates and a pre-Rayleigh wave appears, which evolves into a wave mode which is referred to as a quasi-Rayleigh wave. This wave mode propagates at a velocity between the Rayleigh wave velocities of the two layers and is strongly reflected.

The cutoff at the shear wave velocity of the substrate is associated with a sharp decrease in the mode reflection coefficient, and thus the cutoff thickness (d) or cutoff frequency (f) can be numerically determined based on this sharp decrease. This is especially useful for multiple-layered systems where an approximate formula for the cutoff thickness as



(a)



(b)

FIG. 5. (a) Phase velocities of the SAWs calculated from the reflectance function for two layered thin-film systems with different material properties. (b) Calculated mode reflection coefficients associated with the corresponding SAW modes shown in (a).

derived by Nayfeh and Chimenti¹⁰ for a single layer is hard to determine. The determined cutoff thickness or cutoff frequency can be used to estimate the optimal frequency range.

In general, the properties of the substrate dominate the behavior of the SAW's for a thin-film system at small f^*d . The effects of the layer can be considered as a perturbation to the substrate to cause the frequency dependence of the SAW velocity. On the other hand, the properties of the layer will dominate the behavior of the SAW's at large f^*d , and the substrate can be considered as a loading of the layer, which causes a shift of phase velocity and attenuation of the wave modes. Wave mode transformations always happen when the phase velocity approaches the bulk wave velocity of either the substrate or the thin film. Significant variations in the mode reflection coefficient happen simultaneously.

The theoretical calculations shown above indicate that the reflectance function can be used as a powerful tool to predict the overall frequency dependency of the phase velocity and the mode reflection coefficient of the possible SAW modes. The mode reflection coefficient associated with a SAW mode can predict the level of difficulty of measuring the $V(z)$ curve for a specific wave mode. It can be used to optimize the measurement conditions. For example, the op-

timal frequency range for a stiffening thin-film system should be selected to avoid the weakly reflected pseudo-SAW modes. The velocities predicted by the reflectance function model for estimated elastic constants can be used as initial guesses to help extract velocities from measured $V(z)$ curves. This is very helpful especially when several wave modes may be picked up simultaneously and care must be taken to separate the different wave modes.

II. EXPERIMENTAL MEASUREMENT RESULTS AND DISCUSSIONS

A. Specimens

The Advanced Coating Technology Group (ACTG) at Northwestern University and Brycoat Inc. provided the specimens. TiN films of different thicknesses from 0.80 to 10 μm were deposited on M2 high-speed steel (M2) by dc magnetron sputtering of a TiN target. The density of the M2 substrate was determined by using an electrical balance to weigh the sample and Archimedes' principle to measure its volume. The longitudinal and Rayleigh velocities of the polished M2 substrate were measured by using the pulse-echo

TABLE I. Determined material constants of the M2 substrate and the TiN films with different thickness.

Material	Thickness d (μm)	Density ρ (g/cm^3)	Young's modulus E (GPa)	Poisson's ratio ν	Shear modulus G (GPa)
M2	6×10^3	8.07	217	0.30	83.5
TiN ^a	1.0514	5.40	445	0.20	185
TiN ^a	1.0605	5.40	472	0.20	197
TiN ^a	0.9692	5.40	500	0.20	208
TiN ^b	5	5.40	430	0.20	179
TiN ^b	10	5.40	390	0.21	161

^aDeposited at ACTG.^bDeposited at Brycoat, Inc.

technique and the LFAM, respectively. The measured values were used to determine the material constants of M2 that are listed in Table I.

B. Method

For isotropic thin-film systems the measurements of the surface acoustic waves were performed with a 225-MHz line-focus acoustic microscope. The setup has been described elsewhere.⁸ The phase velocity versus frequency of the SAW can be measured either by varying the operating frequency from 100 to 260 MHz for a single specimen, or at a fixed operating frequency (225 MHz) by using a series of specimens with different thicknesses of the thin film. The first approach is preferred because the properties may be different for thin films of different thicknesses. The LFAM system is tuned at 225 MHz, and the characteristic functions of the

line-focus lens have been modeled for this frequency. For other operating frequencies, the output of the LFAM system has been calibrated by the use of a standard specimen of known material constants. Phase velocities were obtained by measuring $V(z)$ curves at different frequencies on each specimen. The phase velocities were extracted by spectral analysis of the corresponding $V(z)$ curves to determine Δz , and by subsequent use of Eq. (7).

C. Measurement accuracy

As suggested by the reviewer, the following equation is obtained by differentiating the logarithm of Eq. (7).

$$\frac{dc_{\text{LSAW}}}{c_{\text{LSAW}}} = \frac{1}{2} \cdot \left(\frac{dc_f}{c_f} + \frac{df}{f} + \frac{d\Delta z}{\Delta z} \right).$$

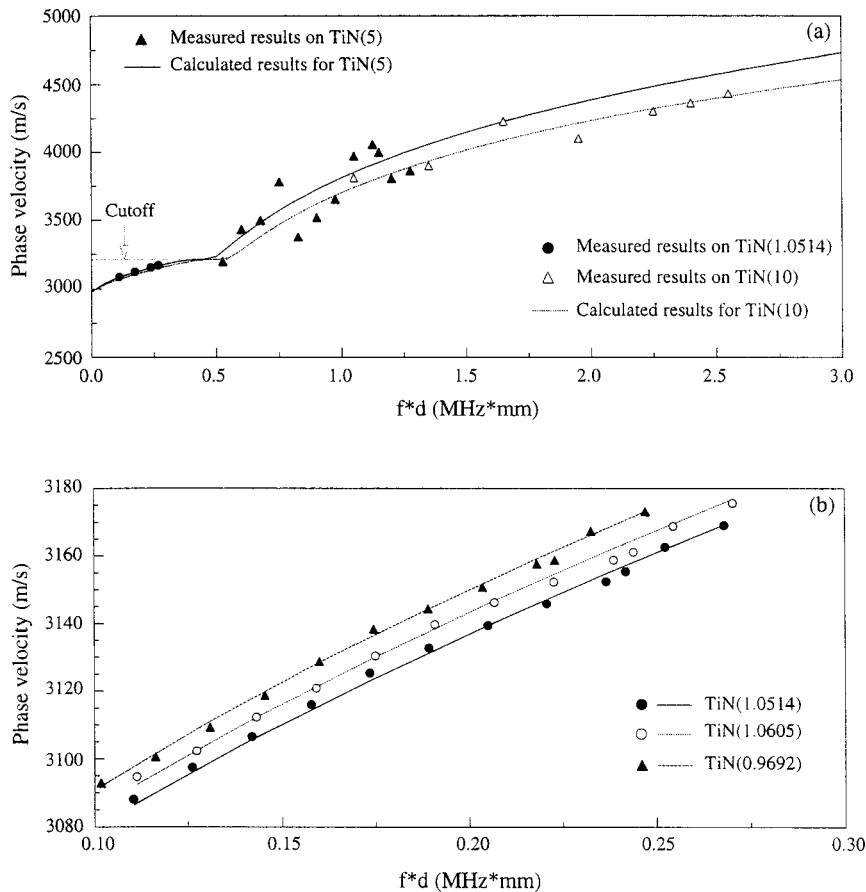


FIG. 6. Phase velocities of the SAWs calculated using the $V(z)$ model with the determined material constants (indicated by lines) plotted together with the measured data (symbols) for TiN thin films of different thicknesses on an M2 substrate.

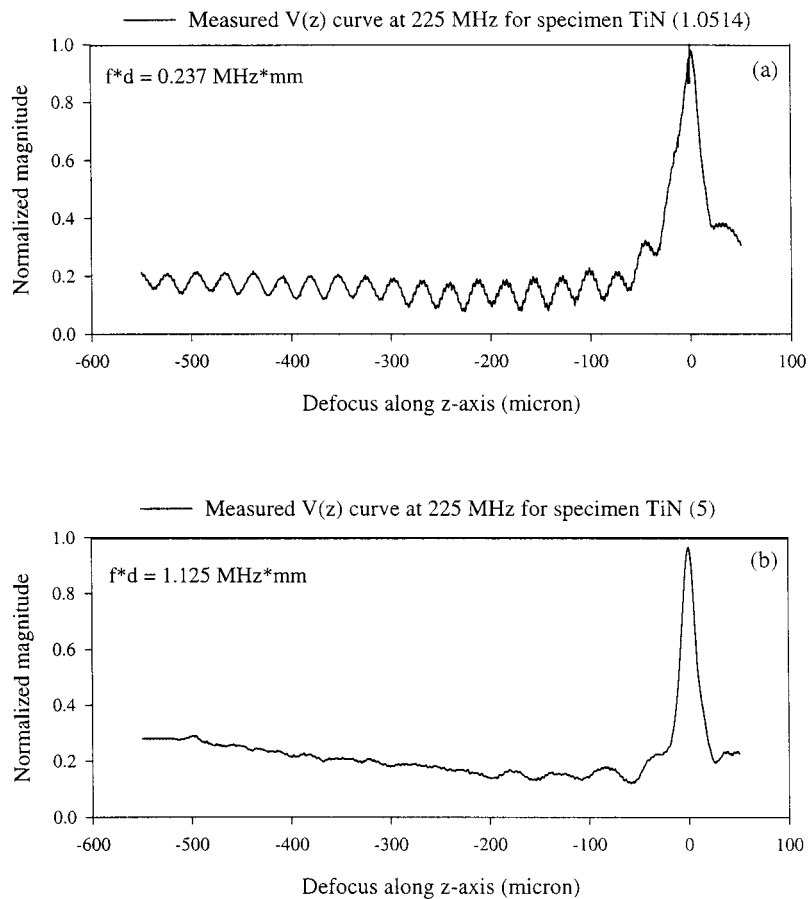


FIG. 7. Measured $V(z)$ curves using the LFAM system at 225 MHz for TiN thin films of different thicknesses on M2 substrates; (a) TiN (1.0514 μm) and (b) TiN (5 μm).

The above equation shows that the accuracy of the phase velocity c_{LSAW} is affected by the three factors c_f , Δz , and f . In other words, the error of c_{LSAW} is the addition of error contributions from the phase velocity in the coupling medium, the period, and the operating frequency.

It is known that c_f is a function of temperature. By measuring the temperature of the coupling water using a thermocouple at the same time that the $V(z)$ curve measurement is made, the velocity c_f can be accurately determined by using the predetermined relation between the wave velocity in water and the temperature of the water. In the LFAM system, a synthesized signal generator with sufficient resolution is used so that the error contributed by the frequency f can be neglected.

During $V(z)$ curve measurement, the specimen is mounted on a mechanical translation stage (x , y , z , and rotation controllable) for the necessary alignment and for the upward or downward movement along the z axis. Starting from the position of the focal line on the surface of the specimen ($z=0$), motion of the acoustic probe toward the specimen, i.e., in the negative z direction, changes the amount of defocus along the z axis to obtain the $V(z)$ curve. The period Δz is determined by spectral analysis of the corresponding $V(z)$ curve. Two major factors affect the accuracy of the period Δz . One is the precision of the mechanical translation and the alignment of the lens and the specimen. The other is the error caused by the FFT process. The translation along the z axis is driven by a stepping motor. The minimum-driving step of the z stage is 0.025 μm . The maximum error of the translation, which may be considered to be $\pm 0.1 \mu\text{m}$

along the z axis, affects the measurement accuracy of the phase velocity by less than 0.01%. Errors may be caused by the data processing. In the present FFT analysis, a total of 8192 points, including the interpolated data, are used together with the use of digital filtering techniques to obtain a sufficiently high frequency resolution. A similar spectral analysis procedure has been described in detail by Kushibiki and Chubachi.⁷ A number of experiments have been made in order to examine the accuracy of the phase velocity by LFAM. The absolute measurement accuracy of the phase velocity has been estimated to be better than $\pm 0.2\%$.

D. Data inversion for the determination of material constants of the thin films

For an arbitrarily configured thin-film system, the phase velocities of the SAWs are functions of the material properties C_{ijkl} , the interfacial parameters K , the thickness d , frequency f_n and wave propagation direction ϕ . Generally, the phase velocities can be measured and calculated with respect to both frequency and wave propagation direction. The thicknesses of the TiN thin films are known in advance. For the cases reported here perfect interfacial conditions were considered. The objective was to determine the elastic constants of polycrystalline TiN films, namely, Young's modulus E and Poisson's ratio ν . The shear modulus G is then calculated using E and ν .

For trial values of the unknown material constants the $V(z)$ measurement model was used to calculate phase velocities. A numerical iterative searching algorithm, called the

simplex method, was subsequently used to find a set of values of the unknowns that minimizes the deviation between calculated and measured velocities. The deviation function to be minimized is defined as

$$\text{Dev} = \frac{1}{N} \sum_{n=1}^N (c_n - \hat{c}_n)^2, \quad \hat{c}_n = F(C_{ijkl}, K, d, f_n, \phi), \quad (9)$$

where c_n and \hat{c}_n are the measured and calculated phase velocities at N frequencies. The results of the minimization process are listed in Table I.

After determination of the material constants as described above, the corresponding phase velocities were recalculated to verify the procedure. The calculated phase velocities are shown in Fig. 6, together with the measured ones at various frequencies. The measured phase velocities were obtained by measuring $V(z)$ curves at different frequencies on each specimen, determining Δz by spectral analysis of the corresponding $V(z)$ curves, and then by subsequent use of Eq. (7). As examples, two measured $V(z)$ curves are shown in Fig. 7. $V(z)$ curves as the one shown in Fig. 7(a) have been obtained for thin film specimens of TiN (1.0514 μm), TiN (1.0605 μm), and TiN (0.9692 μm), which have f^*d values below the cutoff value for f^*d . The $V(z)$ curves as shown in Fig. 7(b) have been obtained for thin film specimens of TiN (5 μm) and TiN (10 μm), which have f^*d values beyond the cutoff value for f^*d . There is a significant decrease in the number and magnitude of the oscillations of the $V(z)$ curve shown in Fig. 7(b) as compared with the $V(z)$ curve shown in Fig. 7(a). The decrease is, as predicted in Sec. I C 2, caused by the low mode reflection coefficient of the pseudo-SAW after the cutoff.

Excellent agreement is observed up to the cutoff value for f^*d . The measurements performed beyond the cutoff are much less accurate compared with the ones performed below the cutoff as indicated in Fig. 6(a) and (b). The lack of accuracy is due to the low mode reflection coefficient of the pseudo-SAW after the cutoff value for f^*d .

The determined elastic constants are different for TiN films of different thicknesses. Values obtained by other investigators are in the same range.^{1,11} Different deposition conditions probably lead to different values of the elastic constants for TiN films.

III. CONCLUSIONS

The reflectance function for a multiple-layered thin film on a substrate, with fluid loading of the film's surface, is an integral part of the $V(z)$ measurement model for a line-focus acoustic microscope. Using estimates for the material parameters that are to be determined, the reflectance function can provide the frequency range for optimal generation of suit-

able surface acoustic wave modes and maximum values of mode reflection coefficients, for effective use of line-focus acoustic microscopy.

Using the reflectance function, a study has been performed to investigate SAW propagation in isotropic thin-film systems with different material combinations. Distinct behaviors of the frequency dependence of the SAW phase velocity in different thin/film systems have been described, especially with regard to wave mode transformation near the cutoff.

Measurements of the $V(z)$ curves for different TiN thin films on a M2 high-speed steel substrate have been carried out. The cutoff phenomenon predicted by the model has been experimentally observed. The frequency dependency of the leaky SAW phase velocity agrees well with theoretical predictions. The phase velocities obtained from measured $V(z)$ curves have been successfully used to determine the elastic constants of TiN thin films.

ACKNOWLEDGMENT

This work was carried out in the course of research supported by the Office of Naval Research under Grant No. N00014-89-J-1362.

- ¹L. E. Toth, *Transition Metal Carbides and Nitrides*, Vol. 7 (Academic, New York, 1971).
- ²P. Z. O. Lefevre and G. A. D. Briggs, "Leaky surface waves propagation on a fast on slow system and the implications for material characterization," *Ultrasonics* **36**, 229–232 (1998).
- ³O. L. P. Zinin, G. A. D. Briggs, P. C. B. D. Zeller, A. J. Kinloch, and G. E. Thompson, "Anomalous behaviour of leaky surface waves for stiffening layer near cutoff," *J. Appl. Phys.* **82**, 1031–1035 (1997).
- ⁴R. D. Weglein and J. O. Kim, "SAW dispersion in diamond films on silicon by acoustic microscopy," in *Review of Progress in Quantitative Nondestructive Evaluation*, Vol. 11B, edited by D. O. Thompson and D. E. Chimenti (Plenum, New York, 1992), pp. 1815–1822.
- ⁵E. C. Pestel and F. A. Leckie, *Matrix Methods in Elasto-mechanics* (McGraw-Hill, New York, 1963).
- ⁶R. D. Weglein, "A model for predicting acoustic material signature," *Appl. Phys. Lett.* **34**, 179–181 (1979).
- ⁷J. Kushibiki and N. Chubachi, "Material Characterization by Line-Focus-Beam Acoustic Microscope," *IEEE Trans. Sonics Ultrason.* **SU-32**, 189–212 (1985).
- ⁸J. D. Achenbach, J. O. Kim, and Y.-C. Lee, "Measuring Thin-Film Elastic Constants by Line-Focus Acoustic Microscopy," in *Advances in Acoustic Microscopy*, Vol. 1, edited by A. Briggs (Plenum, New York, 1995), pp. 153–208.
- ⁹G. W. Farnell, "Properties of Elastic Surface Waves," in *Physical Acoustics*, Vol. VI, edited by W. P. Mason and R. N. Thurston (Academic, New York, 1970), pp. 109–165.
- ¹⁰A. H. Nayfeh and D. E. Chimenti, "Reflection of finite acoustic beams from loaded and stiffened half-spaces," *J. Acoust. Soc. Am.* **75**, 1360–1368 (1984).
- ¹¹A. G. Every, W. Pang, J. D. Comins, and P. R. Stoddart, "Brillouin scattering study of guided modes in TiN films on high-speed steel," *Ultrasonics* **36**, 223–227 (1998).

The use of partially measured source data in near-field acoustical holography based on the BEM

Sung-Chon Kang and Jeong-Guon Ih^{a)}

Center for Noise and Vibration Control, Department of Mechanical Engineering, Korea Advanced Institute of Science and Technology, Science Town, Taejeon 305-701, Korea

(Received 26 July 1999; revised 28 December 1999; accepted 9 February 2000)

In applying the conformal near-field acoustical holography (NAH) to actual source identification problems, it is often possible to determine the velocity at certain points of the source surface in advance. This partially known velocity data would reduce the problem size and permit better reconstruction accuracy. In this paper, the effectiveness of using partially measured source data in the conformal NAH is investigated, which uses the boundary element method. A vibro-acoustic transfer matrix and measured field pressure data, which is involved with the boundary integral equation, are reorganized in order to deal with the partially measured surface velocities. For a baffled vibrating panel, simulations were performed by varying the number of velocity-known nodes. In addition, the effect of measurement error is investigated for two extreme positioning methods of velocity-known nodes. Without regularization, the reconstructed error can be reduced considerably by employing some of the source data and this error can be further reduced by increasing those surface points. However, the velocity reconstruction error is not reduced substantially when the number of velocity-known nodes is less than 30%–40% of the total nodes. The reduction in the reconstruction error is not large if the regularization technique is applied to the restored field. © 2000 Acoustical Society of America. [S0001-4966(00)03105-2]

PACS numbers: 43.35.Sx, 43.40.Yq, 43.50.Yw [ANN]

INTRODUCTION

Near-field acoustical holography (NAH) is an indirect method for the identification of vibro-acoustic properties of vibrating sound sources. In this technique, the acoustic properties on the source plane can be reconstructed, using the field pressure, which is measured on the measurement or hologram plane. Of the two major techniques that have been a great stimulus to the development of NAH, one is based on the spatial Fourier transform and the other on the acoustic boundary element method (BEM). In the former method, the field pressures, which are measured on a hologram plane, are decomposed into space and wave number domains by spatial Fourier transform. The pressure decay in propagation can be compensated and the pressure on a target plane is then reconstructed by an inverse spatial Fourier transform.^{1,2} The shape of the source surface should be regular, and, if this is not the case, then a hypothetical regular plane near the actual source surface needs to be assumed for the reconstruction. In the latter approach, the geometric and vibro-acoustic relation between the sound source and the hologram or measurement plane is modeled as the vibro-acoustic transfer matrix by means of the BEM. The sound radiation and transmission between the vibrating source and the measurement field (or hologram plane) can be modeled by the vibro-acoustic transfer matrix by using boundary integral equation or its discretized form of the BEM. Consequently, the distribution of the surface velocities of the source can be reconstructed by multiplying the inverse of the calculated vibro-acoustic

transfer matrix and the measured field pressure vector at any shape of near-field plane, including the conformal one. This type of conformal NAH is capable of dealing with the complex shaped sources that cannot be described by separable coordinates.

The basic concept of the latter method was first studied by Gardner and Bernhard,³ who introduced the source identification method in the highly reactive field by utilizing direct BEM. Veronesi and Maynard⁴ utilized the singular value decomposition of the discretized direct boundary integral equation, in order to decompose the field and source properties into the wave-vector domain. They demonstrated that the suppression of the amplification effect of the measurement noise via the rejection of higher wave-vector mode components led to an improved resolution. Bai⁵ formulated the generalized holography equation based on direct BEM. He represented all possible combinations of the transfer matrix, in order to correlate the particle velocity and field pressure on the source surface and the field plane. Kim and Ih⁶ described a resolution enhancement technique using the optimal selection of measurement points and regularization of the transfer matrix. The effective independence technique, which selects the field point set by iterative computation to ensure the relatively better linear independence between the rows and columns of transfer matrix, was employed in choosing the optimal measurement points of field pressure. By utilizing a trade-off relationship between variance and bias errors, which produce the minimum reconstruction error, the optimal rank for attaining the minimum mean square error could be determined. Experimental and numerical examples for several interior and exterior acoustic problems have been reported.^{7–10}

^{a)}Currently visiting at the Department of Mathematical Sciences, Loughborough University, Loughborough, Leicestershire LE11 3TU, U.K.

The BEM-based NAH technique thus provides a good opportunity for restoring the vibro-acoustic field of many practical arbitrarily shaped sources. Its optimal feature is that only the measured field pressure is required for determining pressure, particle velocity, intensity, and power flow of the source and the domain of interest as well. However, this method has some inconvenient aspects as well. The acoustic and geometric relation of the source surface and the hologram plane should be modeled via the use of the BEM and this causes problems. A considerable number of boundary elements and nodes are ultimately required for modeling the actual source surfaces involved in a practical noise problem and the amount of field pressure data increases in parallel with that of the surface nodes. The applicable frequency range can be limited by the characteristic length of the typical element. In addition, care should be taken with respect to the inversion of the fully populated vibro-acoustic transfer matrix that has a high singularity.

In applying this technique to actual vibro-acoustic problems, it is often possible to directly and easily measure the surface velocity at some points on the source. Consequently, there is no need to reconstruct all the surface velocities and it would be expected that the value for the restored velocities would be more accurate than that obtained by restoring all the surface data by the NAH technique. This concept would be very useful when the source parameters on the partial area cannot be measured directly because of geometrical interaction, rotating parts and lined surface. This paper describes an investigation of this concept, from the viewpoint of accuracy enhancement and reduction of problem size, and the difficulties associated with its use are discussed. The transfer matrix and measured field pressure data are reorganized in order to deal with the measured surface velocities. As a test example, the sound radiation from a rectangular plate is taken and the simulations, which involve varying the number of velocity-known nodes, are performed. It should also be noted that, in the regularization process, the measurement error is likely to remain with the partial pressure vector corresponding to the velocity-unknown nodes. This error can lead to a divergence in the reconstructed field. The effect of measurement error is investigated for two extreme cases of positions, using known source data.

I. A SHORT REVIEW OF THE NAH BASED ON BEM

The relation between the measured field pressure $\{p\}_f$ and the source data on a surface, $\{p\}_s$ and $\{v\}_s$, can be represented by the matrix form of the Kirchhoff-Helmholtz integral equation as follows:^{5,6}

$$[D]_s\{p\}_s = [M]_s\{v\}_s \quad \text{on the boundary,} \quad (1)$$

$$\{p\}_f = [D]_f\{p\}_s + [M]_f\{v\}_s \quad \text{in the domain.} \quad (2)$$

Here, $[D]_s$ and $[M]_s$ indicate the dipole and monopole matrices on the surface, and $[D]_f$ and $[M]_f$ are those corresponding to field pressures, respectively. It is assumed that the domain, V , enclosed by the boundary, S_0 , is filled with an isothermal, homogenous, inviscid, compressible, and stationary fluid medium, which is disturbed by a time harmonic ($e^{j\omega t}$) acoustic field.

By substituting Eq. (1) and (2) and on the condition that $[D]_s^{-1}$ exists, the field pressure can be expressed by only one unknown boundary condition, viz., the surface velocity, as

$$\{p\}_f = ([M]_f + [D]_f[D]_s^{-1}[M]_s)\{v\}_s \equiv [G]\{v\}_s. \quad (3)$$

In Eq. (3), $[G]$ is the vibro-acoustic transfer matrix that correlates the surface normal velocity with the field pressure and contains geometric information concerning the system as well. If the field pressure is known at m points, the surface velocity at n ($\leq m$) nodes can be uniquely determined by utilizing an overdetermined least-squared solutions approach and singular value decomposition (SVD). The SVD of $[G]$ provides the acoustical modal expansion between the hologram and source field.¹¹ The transfer matrix $[G]$ can be decomposed by

$$[G] = [U][\Lambda][W]^H, \quad (4)$$

where

$$[\Lambda] = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n), \quad \lambda_1 \geq \lambda_2 \geq \dots \lambda_n \geq 0, \quad (5)$$

$$\{u_i\}^H\{u_j\} = \delta_{ij}, \quad \{w_i\}^H\{w_j\} = \delta_{ij}.$$

Here, δ_{ij} is the Kronecker delta, the elements of diagonal matrix $[\Lambda]$ are singular values λ_i , and $[U]$, $[W]$ indicate the vectors, each of which has orthogonal columns. Physically, $\{u_i\}$ and $\{w_i\}$ indicate the wave-vectors which decompose the distribution of field pressure and surface velocity on the hologram and source planes for a selected frequency. Mutually orthogonal wave-vectors constitute the eigenspace of the measurement and source field. The physical meaning of λ_i is the weighting factor for converting these field vectors from the source surface to the hologram plane. Conversely, any singular value represents the contribution of an acoustic mode on the source field to that of the measurement field. The higher order modes which correspond to the components with small singular values are nonpropagating wave components that decay out in the nearfield. By virtue of the SVD, the inverse of Eq. (3) can be expressed as follows:

$$\begin{aligned} \{v\}_s &= [G]^+\{p\}_f = ([G]^H[G])^{-1}[G]^H\{p\}_f \\ &= [W][\Lambda]^{-1}[U]^H\{p\}_f. \end{aligned} \quad (6)$$

Here, $[G]^+$ denotes the $n \times m$ pseudoinverse matrix and the superscript H signifies the Hermitian operator. Equation (6) enables the reconstruction of the velocity field of the source surface, in principle, if the near-field pressures are measured and the transfer matrix is generated by the BEM.

Generally, the source reconstruction problem involves the inverse process of an ill-conditioned matrix. This rank deficiency is due to small singular values that result in the condition number ($= \lambda_{\max}/\lambda_{\min}$) being much larger than unity. The inter-dependency between the columns or rows of the transfer matrix is one reason for this. The system characteristics, including the sensor positions, the source geometry and the source distribution, cause the matrix to be singular. In addition, errors will be generated from the measurement noise that is inevitably included in the observed field pressure and the numerical noise, which occurs in the computation. In the forward calculation, the nonpropagating wave components from the source have negligible

influence on the field pressure, because they are weighted by small singular values. However, in the reverse, the high order small singular values will amplify the nonpropagating wave components. The contaminating noise in the measured field pressure will also be amplified during the same process and this will lead to highly distorted reconstructed data.

The linear independence of vibro-acoustic transfer matrix should be first assured by the proper allocation of measurement points on the hologram plane, irrespective of its shape. This can be accomplished by selecting the required number of points with the aid of an effective independence method.^{6,12} The contribution of sensor position to the linear independence of a transfer matrix can be evaluated for a frequency range of interest and a point or set of points having the smallest value is discarded from the candidate sensor positions. Then, by repeating this process, one can define the measurement positions that are the least inter-dependent for a given number of measurement locations. However, further regularization is required, in order to deal with a large system that has high spatial frequency components. In order to suppress or eliminate the undesirable contribution from higher order singular values and incorporated noise, a proper technique of wave-vector filtering should be taken and the accuracy of the restored result can then be significantly improved. For this purpose, the mean square error of the reconstructed velocity can be represented by a combination of bias error and variance error. Here, the bias error signifies the difference in vibro-acoustic transfer matrices before and after the wave-vector filtering. This increases as the original matrix is distorted, due to filtering. The variance error is due to variations in the measurement noise of the field pressure and the latter is assumed to be in the form of an unbiased Gaussian random distribution. This value decreases as the amount of measurement noise is reduced by filtering. The total estimated mean square error has a trade-off trend with the variation in design parameters in the wave-vector filtering, including the discarding order of singular values, the regularization parameter, and the number of iterations. The optimal wave-vector filters can be determined at the condition of minimum value of the estimated total mean square error. In this study, the iterative regularization method¹³ as briefly described in the Appendix is employed in the simulation.

II. REFINEMENT OF A RECONSTRUCTED FIELD USING A PARTIALLY KNOWN SOURCE VELOCITY

If the velocity at some nodes on the source surface is known *a priori*, Eq. (3) can be modified as follows:

$$\{p\}_f = [G_k G_u] \begin{Bmatrix} v_k \\ v_u \end{Bmatrix}_s = [G_k 0] \begin{Bmatrix} v_k \\ 0 \end{Bmatrix}_s + [0 G_u] \begin{Bmatrix} 0 \\ v_u \end{Bmatrix}_s. \quad (7)$$

Here, $[G_k]$, $[G_u]$ represent the partial transfer matrices which correspond to the velocity-known and velocity-unknown nodes, and $\{v_k\}_s$, $\{v_u\}_s$ the components of the velocity vector which corresponds to the velocity-known and -unknown nodes, respectively. From Eq. (6), one can write as

$$[G_u]\{v_u\}_s = \{p\}_f - [G_k]\{v_k\}_s \equiv \{p_u\}_f, \quad (8)$$

or inversely,

$$\{v_u\}_s = [G_u]^+ \{p_u\}_f. \quad (9)$$

Equation (9) is the modified holography equation for the reconstruction of the velocity of the unknown nodes by using the partially known N_k velocities of the source.

In actual situations, the measured surface velocity and field pressure are prone to contamination by the noise. In this paper, the measured surface velocity is considered to be the true one by assuming that the noise, which is included in the measured velocity, is negligible by careful calibration and measurement. In fact, it is impossible to expect the reconstructed result to be better than the measured one. The measured field pressure, $\{\tilde{p}\}_f$, can be expressed as

$$\{\tilde{p}\}_f = \{p\}_f + \{n\}_f = [G]\{v\}_s + \{n\}_f, \quad (10)$$

where $\{n\}_f$ denotes the noise included in the field pressure measurement. The estimated velocity for the velocity-unknown nodes, $\{\hat{v}_u\}_s$, is given by

$$\{\hat{v}_u\}_s = [G_u]^+ (\{p_u\}_f + \{n\}_f) = \{v_u\}_s + [G_u]^+ \{n\}_f. \quad (11)$$

Consequently, during the reconstruction process, the contaminating measurement noise will remain in the pressure vector being used for reconstructing the unknown nodes. The second term in the right-hand side of Eq. (11) can be either smaller than or larger than $\{v_u\}_s$, depending on whether the majority of elements in $\{v_u\}_s$ are chosen from the positions near the anti-nodes or nodal points of the velocity, respectively. When this term is larger than $\{v_u\}_s$, $\{\hat{v}_u\}_s$ will diverge, although the reconstruction error in counting for all nodes will not be very large. The reason for this is that the distorted small amplitudes at the nodal points are always relatively quite small compared to those at the other nodes.

III. NUMERICAL SIMULATION

A rectangular vibrating 0.6×0.4 m steel plate, 1 mm in thickness lying on an infinite baffle, was used as a test simulation model. The boundary element model of the plate consists of 64 nodes and 96 linear elements as shown in Fig. 1. The maximum value of the characteristic length (L) of the model was 103 mm, which limited the effective high frequency to 557 Hz under the $\lambda/6$ -criterion for linear elements. Forty-eight field points were evenly distributed on the hologram plane, which is separated from the source by 20 mm. The rigid boundary condition was assumed on all edges of the plate having 29 fixed nodes. The number of unknown nodes was 35, and thus the overdetermined problem was dealt with. For the simulation, the (2,2) mode of vibration at 119 Hz and the (4,3) mode at 287 Hz were selected as shown in Fig. 2. The mode shapes of two vibration modes were estimated from a finite element analysis of the finely meshed model. The pressures at the field points were calculated by BEM and the measurement noise was artificially added, by means of a random number generator. All calculated results were the mean values of 100 simulations in which the incorporated noise was varied.

Figure 3 shows the condition number of partial transfer matrix $[G_u]$. The velocity-known nodes are sequentially selected, starting with node 1 in Fig. 1(a). Figure 3 shows that

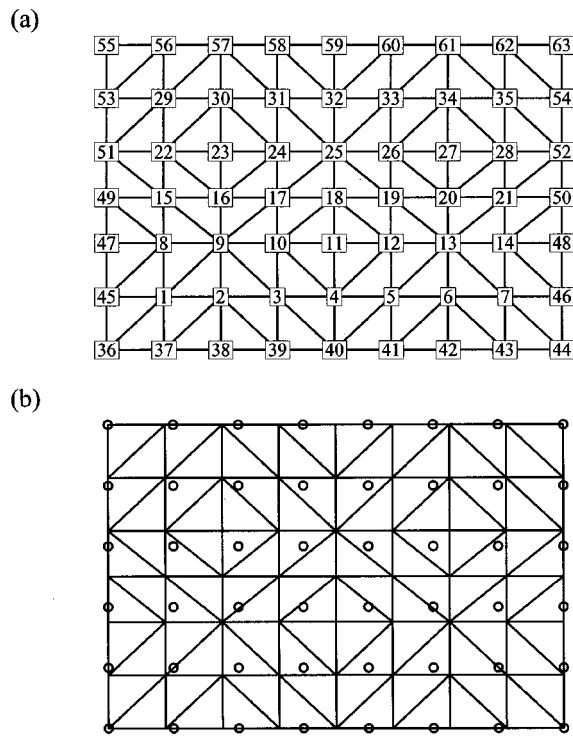


FIG. 1. (a) Boundary element model of a rectangular plate on an infinite baffle; (b) 48 evenly spaced field points on the measurement plane are separated from the source surface by 20 mm.

the singularity of the partial transfer matrix can be improved by increasing the number of nodes with known velocities. When the number of velocity-known nodes, N_k , is increased, the size of $[G_u]$ is reduced, the order of λ_{\min} is reduced, thus λ_{\min} is increased rapidly, and the condition number decreases.

Figure 4 shows a comparison of the reconstruction errors before and after applying the regularization and varying the number of known surface data for an S/N ratio=30. The iterative regularization method is applied in this simulation, as described above. The velocity reconstruction error is calculated as

$$\varepsilon = \frac{\|\{\nu\}_s - \{\hat{\nu}\}_s\|^2}{\|\{\nu\}_s\|^2} \times 100 (\%), \quad (12)$$

where $\{\hat{\nu}\}_s$ is the estimated velocity vector. The measure-

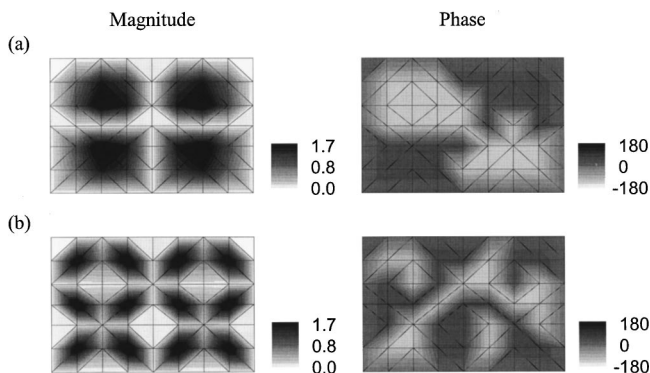


FIG. 2. Distribution of the actual measured surface velocity: (a) 119 Hz; (b) 287 Hz.

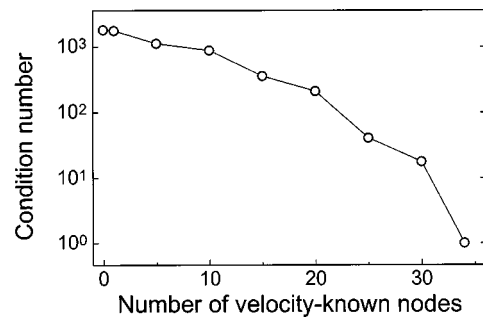


FIG. 3. The condition numbers for the partial transfer matrix for a 119 Hz mode with increasing number of velocity-known nodes.

ment noise is assumed to have an unbiased Gaussian random distribution. When the transfer matrix is not regularized, the reconstructed result is highly affected by the value of λ_{\min} . Therefore, a reduction in the condition number of the partial transfer matrix introduces a small reconstruction error. When all velocities at the surface nodes are unknown and the wave-vector filtering is not yet applied, the reconstruction error amounts to as much as 21 865% and 2942% for (2,2) or (4,3) mode, respectively. These large errors are reduced to 9.6% and 7.9%, respectively, by using the optimally designed wave-vector filter. The reconstruction error can be dramatically reduced with an increase in N_k when the regularization is not applied to the recovered source data. In contrast, a reduction in the reconstruction error is small with an increasing amount of *a priori* known source data, when the regularization is used. In particular, the reconstruction error is not improved when N_k is less than 15 or 10 for the (2,2) or (4,3) mode, respectively, although the reconstruction error for the nonregularized result is reduced. The reason for this is not clear, because the reconstructed results are complicated and depend on several factors, including the shape of the optimal wave-vector filter, the distribution of the surface velocity and field pressure, and the singularity of transfer matrix. One reason for this is that the reconstruction error has already been sufficiently reduced by regularization. However, the error will be suppressed by increasing the number nodes with *a priori* known velocities, when N_k becomes larger than 10 or 15. This trend is because the singularity of partial transfer matrix is lowered and the velocity data at some nodes are replaced with experimentally measured values. When N_k is

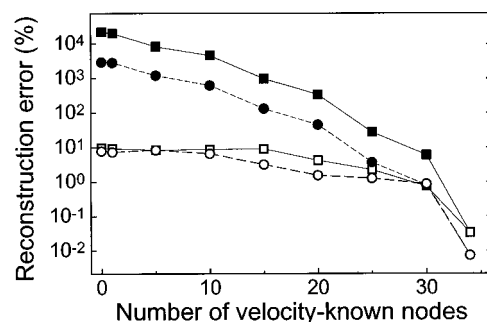


FIG. 4. Comparison of velocity reconstruction error before and after regularization (S/N ratio=30). For 119 Hz mode: —■—, before regularization; —□—, after regularization. For 287 Hz mode: ---●---, before regularization; ---○---, after regularization.

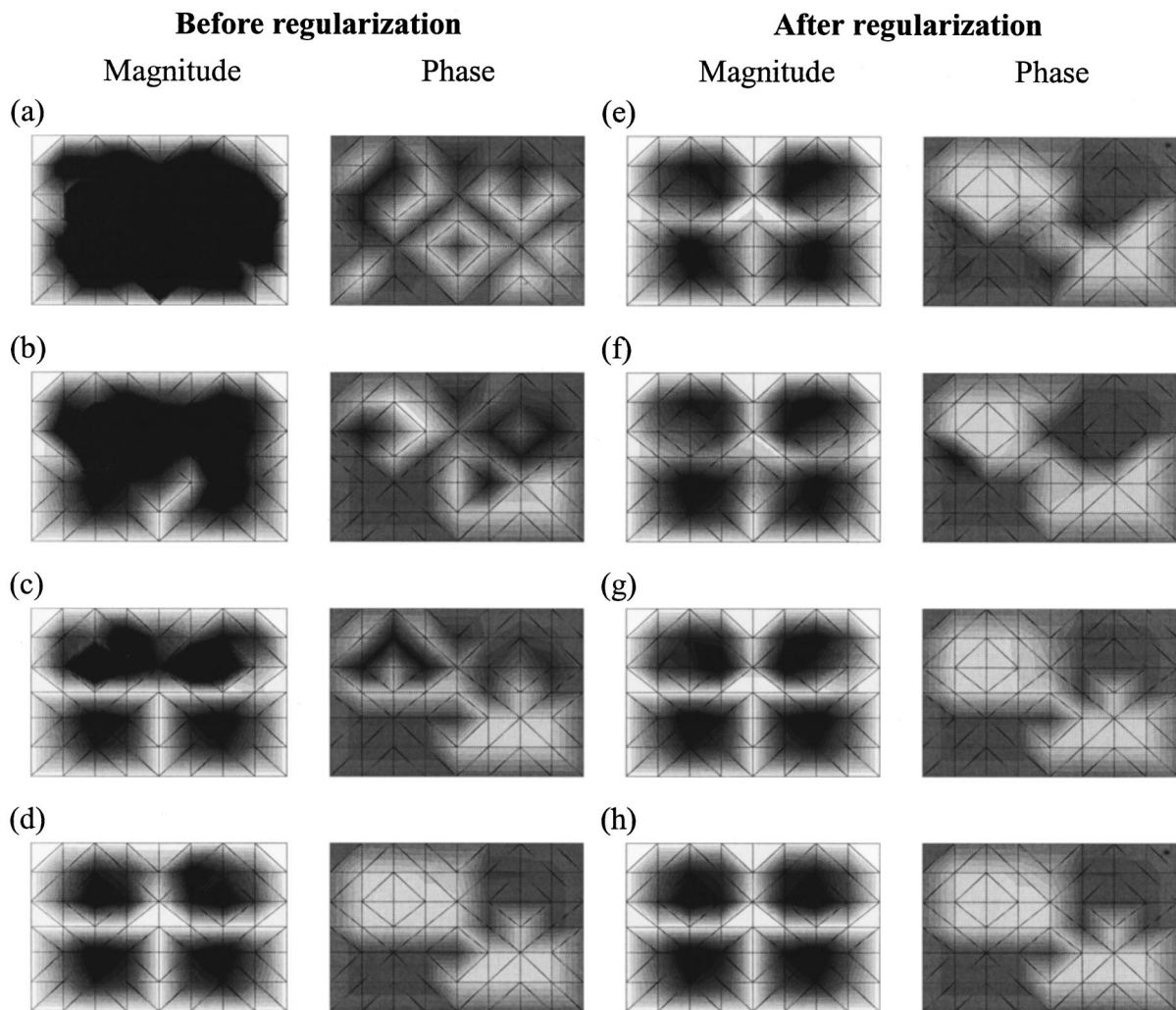


FIG. 5. Reconstructed surface velocity for a 119 Hz mode. Before regularization: (a) $N_k=0$ ($\varepsilon=21\,865.2\%$), (b) $N_k=10$ ($\varepsilon=4571.2\%$), (c) $N_k=20$ ($\varepsilon=336.7\%$), (d) $N_k=30$ ($\varepsilon=6.0\%$). After regularization: (e) $N_k=0$ ($\varepsilon=9.6\%$), (f) $N_k=10$ ($\varepsilon=8.9\%$), (g) $N_k=20$ ($\varepsilon=4.2\%$), (h) $N_k=30$ ($\varepsilon=0.7\%$).

30, the reconstruction error is reduced to about 10% of the result for $N_k=0$. In any event, one can reduce the reconstruction error by using the partially known source data in this case.

The reconstructed velocity distributions, which are illustrated in Figs. 5 and 6, compare the results before and after the regularization process when the source velocities are known *a priori* at 10, 20, or 30 nodes beginning from node number 1. Prior to regularization, the reconstructed results are highly distorted from the actual shapes, but they are amended dramatically by utilizing the optimally designed wave-vector filter. It can be seen that a far better result can be obtained by using some of the actually measured source data before the regularization is performed, and the recovered result is improved with increasing N_k . However, the level of improvement is not satisfactory at this point. The distribution of the final surface velocity can be further improved by conducting the regularization using optimal wave-vector filtering. This leads to only minor improvement with increasing N_k .

Figure 7 illustrates an example that compares the actual velocity field and reconstructed velocity distribution for $N_k=0$ with the amplitude of a reconstructed velocity field using

the first 15 nodes with known velocities where the S/N ratio is 30. The velocities that are reconstructed by utilizing the known velocity at the first 15 nodes are very close to the actual velocities for other nodes, especially for the case of nodes having large velocity amplitudes. This means that the overall improvement of the reconstructed result in the refinement process is caused, not only by the use of actual velocities at the velocity-known nodes, but by error reduction at other nodes as well.

If the velocity information is known for more nodes, the noise included in the measured pressure will remain in the reformed field pressure, which, in turn, will be used for estimating the velocity $\{v_u\}_s$ as indicated in Eq. (1). When the relative proportion of measurement noise $\{n\}_f$ becomes larger than $\{p_u\}_f$, the reconstructed result will diverge, eventually. If the nodes with unknown velocities were to be located on or near the nodal line, $\{p_u\}_f^2$ would be much smaller than $\{m\}_f^2$ and one would expect a large error in velocity reconstruction. In order to simulate this type of extreme case, the velocity reconstruction error is compared in Fig. 8 for the two sets of node arrangements with known velocities. Here, ‘‘set I’’ denotes a set of in-phase nodes chosen from the surface points having relatively small veloc-

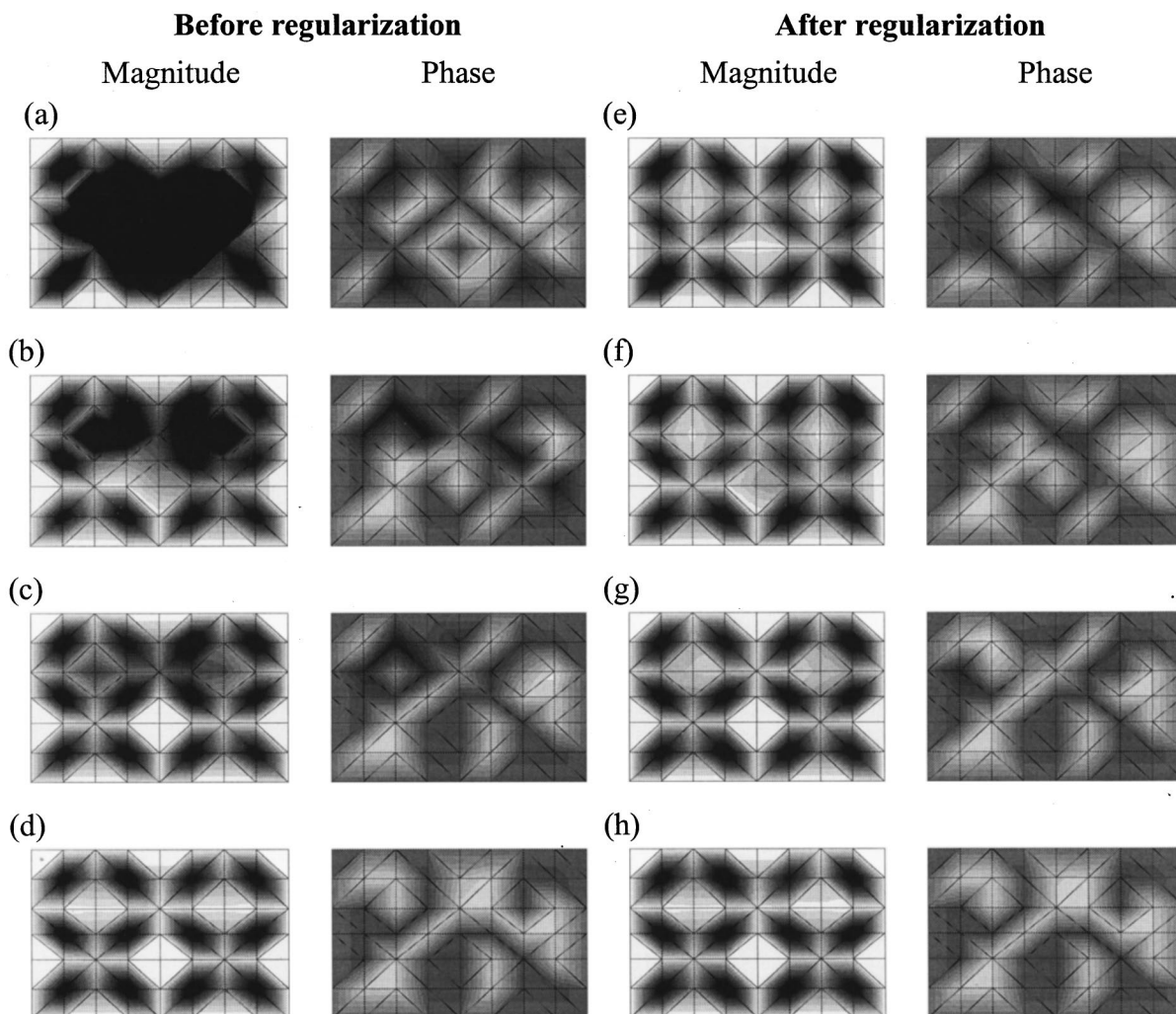


FIG. 6. Reconstructed surface velocity for a 287 Hz mode. Before regularization: (a) $N_k=0$ ($\varepsilon=2941.6\%$), (b) $N_k=10$ ($\varepsilon=624.4\%$), (c) $N_k=20$ ($\varepsilon=45.3\%$), (d) $N_k=30$ ($\varepsilon=0.8\%$). After regularization: (e) $N_k=0$ ($\varepsilon=7.9\%$), (f) $N_k=10$ ($\varepsilon=6.7\%$), (g) $N_k=20$ ($\varepsilon=1.6\%$), (h) $N_k=30$ ($\varepsilon=0.9\%$).

ity amplitudes, and “set II” consists of in-phase nodes with relatively large velocity amplitudes. For two cases, the overall reconstruction error is continuously decreased even for large N_k . This result for set II is due to the fact that the measurements are performed at nodes mostly having large amplitudes, the reconstruction error is determined by this actual and large $\|\{v\}_s\|^2$, and the recovering error for large N_k is not diverged.

Before drawing conclusions, the effects of S/N ratio

should be mentioned. The dependence of error percentage on the signal-to-noise ratio is investigated as depicted in Fig. 9 where curves are shown for two S/N ratios of 10 and 30. It can be seen that the reconstruction errors decrease with the increase of S/N ratio and N_k , as expected. This result conclusively demonstrates the effect of noise included in the measured field pressure. It is clear that a great care is needed in the measurement of field pressure in order to obtain a sensitive reconstruction result.

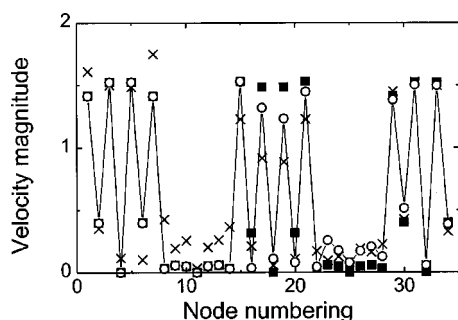


FIG. 7. Amplitude distribution of the reconstructed velocity of 287 Hz mode when the velocities at nodes 1–15 are known *a priori* (S/N ratio=30): ■, actual; ×, $N_k=0$; —○—, $N_k=15$.

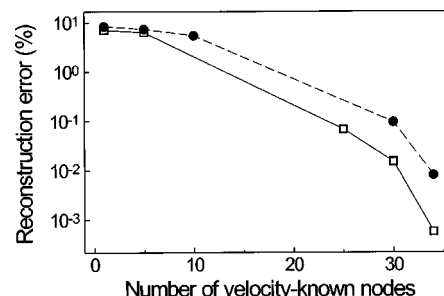


FIG. 8. Reconstruction error of the velocity for a 287 Hz mode with regularization (S/N ratio=30). —●—, set I; —□—, set II.

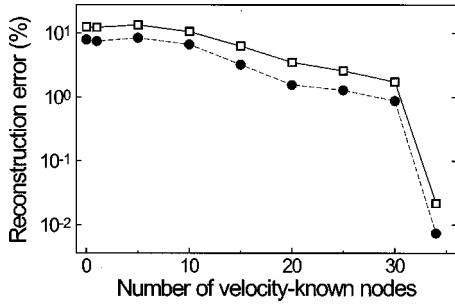


FIG. 9. Reconstruction error of the velocity at a 287 Hz mode; —□—, S/N ratio=10; ---●---, S/N ratio=30.

IV. CONCLUSIONS

The BEM-based NAH using partially known velocity data has been studied for the possible refinement of reconstruction of source velocity and reduction in problem size. For this purpose, the holography equation was modified by subtracting the pressure contributions of the velocity-known nodes to the measured pressure. By performing a numerical simulation, it was found that the singularity of partial transfer matrix of modified holography equation becomes considerably reduced as the number of nodes with known velocities is increased; the velocity reconstruction error follows this pattern as well. When the wave-vector filter is applied optimally in the regularization, the final reconstruction error can be dramatically improved. However, the velocity reconstruction error is not reduced substantially when the number of velocity-known nodes is less than 30%–40% of the total nodes. The shapes of wave-vector filter, the distributions of surface velocity and field pressure, and the singularity of transfer matrix can influence the reconstructed results in a very complicated manner. The effect of measurement noise remaining in the field pressure was also investigated. The measurement noise varies in accordance with the distribution of velocity-known nodes. When the velocity data are known *a priori* at the nodes with small vibration amplitudes, the reconstruction error for such nodes becomes very large. However, the overall reconstruction error for all the surface points will be greatly reduced compared to the reconstructed result for $N_k=0$. It should also be noted that the reconstruction error would be reduced with the measured field pressures having better S/N ratios.

APPENDIX: ALGORITHM FOR ITERATIVE REGULARIZATION

When field pressures are measured, surface velocities can be estimated from the inverse iterations as¹⁴

$$\{\hat{v}\}_s^{l+1} = \beta[G]^H\{\tilde{p}\}_f + ([I_n] - \beta[G]^H[G])\{\hat{p}\}_s^l, \quad (\text{A1})$$

where $\{\hat{v}\}_s^l$ denotes the estimated surface velocity at the l th iteration step, $[I_n]$ is the identity matrix with rank n , and β means the convergence parameter. The necessary and sufficient condition for the convergence of iteration as increasing the iteration number l is

$$|1 - \beta\lambda_i^2| < 1, \quad \text{for } i = 1, 2, \dots, n. \quad (\text{A2})$$

If $[G]$ is invertible, the limit value for $l \rightarrow \infty$ is equivalent to the pseudo-inverse solution as

$$\{\hat{v}\}_s^\infty = ([G]^H[G])^{-1}[G]^H\{\tilde{p}\}_f. \quad (\text{A3})$$

If measurement noises are taken into consideration, the reconstructed field does not converge to the true surface field, even if the number of iterations is infinitely increased. This is because the measurement noises are amplified by non-propagating wave components during the backward reconstruction. Therefore, the iteration process should be terminated at an appropriate stage of iterations, in order to achieve the minimum reconstruction errors.

The direct implementation of iterative method can be written as

$$\{\hat{v}\}_x^l = \beta \sum_{k=0}^l ([I] - \beta[G]^H[G])^k [G]^H\{\tilde{p}\}_f \quad (\text{A4a})$$

$$= \beta \sum_{k=0}^l [W][\Lambda] 1 - \beta[\Lambda]^2)^k [U]^H\{\tilde{p}\}_f \quad (\text{A4b})$$

$$= [W] \text{diag}[(1 - (1 - \beta\lambda_1^2)^{l+1}), \dots, (1 - (1 - \beta\lambda_n^2)^{l+1})] \times [\Lambda]^{-1} [U]^H\{\tilde{p}\}_f. \quad (\text{A4c})$$

By comparing Eq. (A4c) with Eq. (6), the wave-vector filter matrix $[F_l]$ can be defined as

$$\{\hat{v}\}_s^l = [W][F_l][\Lambda]^{-1} [U]^H\{\tilde{p}\}_f, \quad (\text{A5})$$

where

$$[F_l] = \text{diag}[(1 - (1 - \beta\lambda_1^2)^{l+1}), \dots, (1 - (1 - \beta\lambda_n^2)^{l+1})]. \quad (\text{A6})$$

All diagonal components of this matrix should be less than or equal to one.

¹E. G. Williams and J. D. Maynard, "Holographic imaging without the wavelength resolution limit," *Phys. Rev. Lett.* **45**, 554–557 (1980).

²J. D. Maynard, E. G. Williams, and Y. Lee, "Near-field acoustic holography: I. Theory of generalized holography and the development of NAH," *J. Acoust. Soc. Am.* **78**, 1395–1413 (1985).

³K. Gardner and R. J. Bernhard, "A noise source identification technique using an inverse helmholtz integral equation method," *Trans. ASME, J. Vib. Acoust. Stress, Reliab. Des.* **110**, 84–90 (1988).

⁴W. A. Veronesi and J. D. Maynard, "Digital holographic reconstruction of source with arbitrarily shaped surfaces," *J. Acoust. Soc. Am.* **85**, 588–598 (1989).

⁵M. R. Bai, "Application of BEM (boundary element method)-based acoustic holography to radiation analysis of sound sources with arbitrarily shaped geometries," *J. Acoust. Soc. Am.* **92**, 533–549 (1992).

⁶B.-K. Kim and J.-G. Ih, "On the reconstruction of the vibro-acoustic field over the surface enclosing an interior space using the boundary element method," *J. Acoust. Soc. Am.* **100**, 3003–3016 (1996).

⁷B.-K. Kim and J.-G. Ih, "On the use of the BEM-based NAH for the vibro-acoustic source imaging on the nonregular exterior surfaces," *Proc. NOISE-CON 98*, pp. 665–670, Ypsilanti (1998).

⁸B. K. Gardner, J. S. Bolton, and D. L. Hallman, "Nearfield acoustical holography in arbitrary geometries by the use of the boundary element method," *Proc. NOISE-CON 98*, pp. 671–676, Ypsilanti (1998).

⁹J.-G. Ih, S.-C. Kang, S.-J. Kim and K.-S. Kang, "Reconstruction of the vibro-acoustic field on the surface of the refrigerator compressor by using the BEM-based acoustic holography," *Proc. 1998 Int. Compressor Eng. Conference*, pp. 525–529, Purdue (1998).

¹⁰E. G. Williams, "An overview of the inverse problem for sound reconstruction in interior spaces," *J. Acoust. Soc. Am.* **105**, 968(A) (1999).

¹¹D. M. Photoiadis, "The relationship of singular value decomposition to

- wave-vector filtering in sound radiation problems,” J. Acoust. Soc. Am. **88**, 1152–1159 (1990).
- ¹²D. C. Kammer, “Effects of noise on sensor placement for on-orbit modal identification of large space structures,” Trans. ASME, J. Dyn. Syst. Meas. and Control **114**, 436–443 (1992).
- ¹³B.-K. Kim and J.-G. Ih, “Design of an optimal wave-vector filter for enhancing the resolution of reconstructed source field by NAH,” J. Acoust. Soc. Am. (to be published).
- ¹⁴J. Biemond, R. L. Lagendijk, and R. M. Mersereau, “Iterative methods for image deblurring,” Proc. IEEE **78**, 856–884 (1990).

Performance measurements on a thermoacoustic refrigerator driven at high amplitudes

Matthew E. Poesé and Steven L. Garrett^{a)}

Graduate Program in Acoustics, The Pennsylvania State University, P.O. Box 30, State College, Pennsylvania 16804

(Received 11 June 1998; revised 8 December 1999; accepted 20 December 1999)

Since the power density in a thermoacoustic device is proportional to the square of the acoustic Mach number, there is strong motivation to design thermoacoustic refrigerators to operate at larger pressure amplitudes. Measurements are reported of a modified version of the Space Thermo-Acoustic Refrigerator (STAR), driven at peak-to-mean pressure ratios up to 6%. This pressure ratio corresponds to 30 W of cooling power—five times as large as reported for STAR in 1993. The results of these measurements are compared to a DELTAE computer model of the low-amplitude (linear) performance that matches experimental conditions on a point-by-point basis. It is found that there is a small but measurable deviation in heat pumping power from the power predicted with a linear acoustic computer model at moderate amplitudes. This deviation in heat pumping power at 6% pressure ratio is about 23%. A large disagreement in the acoustic power needed to attain a specific pressure ratio is found between measured data and DELTAE results. An overview of the instrumentation, including a measurement of exhaust heat with an absolute accuracy of 65 mW, is also presented. © 2000 Acoustical Society of America. [S0001-4966(00)00504-X]

PACS numbers: 43.35.Ud [HEB]

INTRODUCTION

The proliferation in the last 50 years of very reliable and inexpensive refrigerators and air conditioners has led the population of many developed countries to regard cooling machines as a necessity rather than a luxury. Because a thermoacoustic refrigerator uses no environmentally harmful working fluids, has very few moving parts, and can utilize a proportional control scheme, the technology may become attractive for widespread commercial use.

The fact that the power density of a thermoacoustic device in the linear regime increases with the square of increasing pressure ratio provides strong motivation to design thermoacoustic refrigerators to operate at higher pressure ratios. This experiment investigated deviations from linear predictions of efficiency and heat pumping power at high pressure amplitudes and determined at what pressure ratio these deviations become significant. The predictions of linear system performance were made with a computer program called DELTAE^{1,2} that models the performance of thermoacoustic and other one-dimensional acoustic apparatus by numerically integrating a one-dimensional wave equation in the usual low-amplitude approximation.

This experiment was conducted on a small thermoacoustic refrigerator affectionately called “Frankenfridge” because the engine is made of parts³ from other refrigerators, namely the *Space ThermoAcoustic Refrigerator*⁴ (STAR) and the *Shipboard Electronics Thermoacoustic Cooler*⁵ (SETAC), both designed and built at the Naval Postgraduate School. Frankenfridge is the resonator section of STAR, which includes the stack and heat exchangers, coupled to a SETAC driver. The STAR driver produced about 10 W of

acoustic power while the SETAC driver is capable of producing up to 100 W of sound power. The SETAC driver creates large enough pressure amplitudes in the STAR resonator that nonlinear effects can be measured.

I. APPARATUS

A. The driver

The SETAC driver has an *in vacuo* mechanical resonance frequency of 316 Hz. It is instrumented with a miniature Endevco piezoresistive microphone (Model 8514-10) positioned very near the pusher cone and a miniature Entran piezoresistive accelerometer (Model EGA-125-1000D) mounted directly on the rear surface of the pusher cone. The signals from these two sensors, along with the effective cross-sectional area of the bellows^{6,7} (21 cm²), allow the measurement of acoustic input power as well as the stroke of the pusher cone and the acoustic impedance that the resonator presents to the driver.

The pertinent driver parameters are listed in Table I. The voice coil is attached to a reducing cone that ends in an aluminum piston face. This cone is attached to the driver housing with a two-convolution electroformed nickel bellows that provides a flexure seal for the resonator and eliminates the need for a sliding seal. The main function of the bellows is to seal the oscillating pressure in the resonator from the driver back volume. A capillary leak between the volume of the resonator and the driver back volume allows for equilibration of the 10 atm of static pressure.

The sinusoidal electrical signal supplied to the driver originates from an HP 3314A signal generator and is amplified by a Techron 7520 power amplifier. The drive frequency is adjusted to assure a 90-degree phase relationship between acceleration and pressure—a relationship that is monitored

^{a)}Electronic mail: garrett@sabine.acs.psu.edu

TABLE I. SETAC driver parameters.

Parameter	Value	Units	Relative uncertainty
Moving mass	36.4	g	0.8%
Stiffness	143	kN/m	0.5%
Mechanical resistance	2.10	kg/s	0.6%
B1	19.1	N/A	0.7%
dc electrical resistance	1.677	Ω	negligible
Effective bellows area	21	cm ²	3%

with a Lissajous pattern on an oscilloscope displaying the accelerometer and microphone signals. There are many reasons to stay on resonance, not the least of which is that heat pumping power is a strong function of stack position in the standing wave. This position is calculated and designed to be a constant parameter for the system at resonance and, if the frequency is not changed to accommodate the changing sound speed of the gas mixture, this stack position in the standing wave will vary with temperature. Another reason for resonant operation (the mechanical resonance frequency of the driver and the resonator are made coincident by adjusting the relative concentration of helium and argon in the resonator) is that the driver has the potential to deliver its maximum power when it operates close to its mechanical resonance frequency.⁸

B. Resonator

The resonator is instrumented with thermocouples to measure the external metal temperatures near the hot (exhaust) and cold (heat load) heat exchangers. There are no sensors mounted directly inside of the resonator. As shown in Fig. 1, the resonator is equipped with a small Minco electrical resistance heater just below the cold side heat exchanger. This electrical heater provides an easily controlled and measured amount of heat for the refrigerator to pump. Table II lists parameters that describe the working fluid, stack, and heat exchangers.

C. Exhaust heat sink flange

Even though the cold side of the refrigerator is insulated from the room using standard Corning Pink fiberglass insulation, heat leak to the cold side is expected when large temperature differences exist between the cold side duct and room air temperature. From the first law of thermodynamics and the measured power delivered by the driver, knowledge of the exhaust heat flux allows calculation of the heat leak and provides an accurate determination of the amount of heat that the refrigerator pumps. However, in the experiment presented here, the cold side was only a few degrees colder than the room making the heat leak too small to measure. For these measurements, the exhaust heat sink flange allowed the continuous verification that the measured amount of heat supplied by the resistance heater was moved up the stack from the cold side to the hot side. For other experiments with this refrigerator which included larger temperature differ-

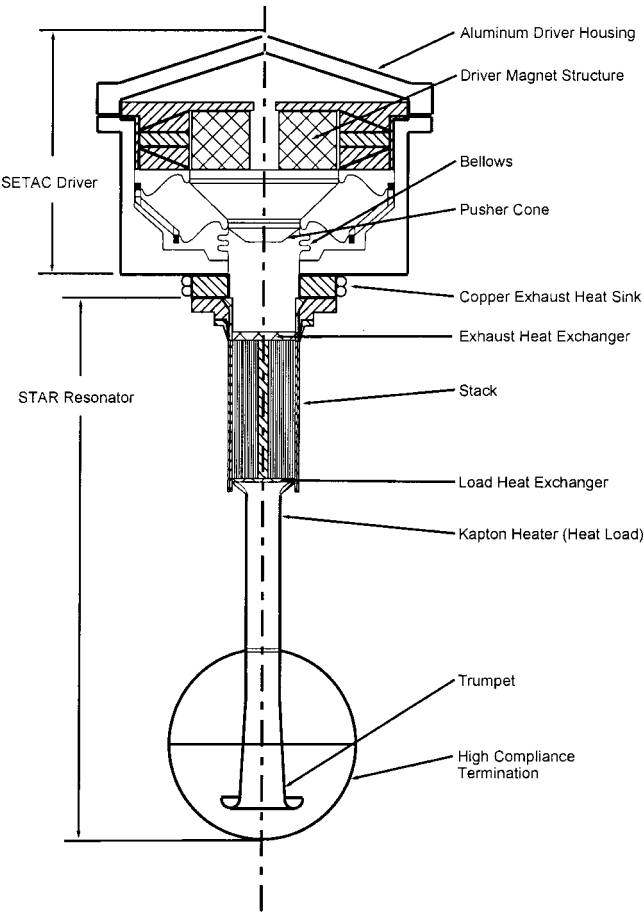


FIG. 1. Cross-sectional diagram of Frankenfridge. Approximate height is 56 cm and maximum width at the driver housing is 23 cm.

ences, the exhaust heat sink flange was relied upon to measure the significant heat leak from the room to the cold parts of the resonator.

To make this measurement of exhaust heat flux, a copper exhaust heat sink shown in Fig. 1 has been sandwiched between the SETAC driver and the STAR resonator. This $\frac{1}{2}$ -in.-thick flange has two loops of $\frac{1}{8}$ -in. copper refrigeration tubing wrapped and soldered around its perimeter to allow water to be circulated around the flange. The water that is pumped through this loop and around the flange absorbs the exhaust heat and therefore experiences an increase in temperature as it travels around the flange. The water circulation loop is instrumented with a ten junction thermopile (located directly across the inlet and discharge ports of the flange piping) and a Hedland flow meter (IR-OPFlow Model 502-101) located in-line with the TygonTM tubing that plumbs the pump, filter, and dissipating heat exchanger in the loop. The heat flux through the exhaust flange to the water flowing in the loop can be calculated from the heat capacity of the water, the temperature difference, and the fluid flow rate.

The driver housing is not insulated from the room. For this reason, thermal insulation between the exhaust heat sink and the aluminum driver housing is critical if a temperature increase is to be directly measured in the exhaust heat sink flange. The resonator, the new exhaust heat sink, and a $\frac{1}{8}$ -in.-thick DelrinTM insulating ring are bolted to the driver housing. The exhaust thermal measurement system is shown in

TABLE II. Physical parameters of the resonator and working fluid.

Parameter	Symbol	Value	Units
Mean pressure	p_m	1.07	MPa
Mean stack temperature	T_m	290	K
Gas mixture (helium/argon)		85.5% He	
Gas mixture atomic mass	M	9.214	kg/kmol
Gas mixture density	ρ	4.105	kg/m ³
Gas mixture sound speed	a	660.4	m/s
Gas mixture specific heat	c_p	2256	J/kg·K
Gas mixture polytropic coefficient	γ	1.667	
Gas mixture Prandtl number	σ	0.428	
Gas mixture kinematic viscosity	μ	2.10×10^{-5}	kg/s·m
Gas mixture thermal conductivity	K_g	0.111	W/m·K
Stack thermal conductivity	K_s	0.161	W/m·K
Stack specific heat	c_s	1101	J/kg·K
Stack material density	ρ_s	1348	kg/m ³
Stack plate thickness	$2l$	0.0762	mm
Stack plate separation	$2y_0$	0.191	mm
Stack length	Δx	78.5	mm
Center position of stack (ref. from driver)	x_s	106.5	mm
Stack radius	R	19.1	mm
Stack (spiral) perimeter	Π	4846	mm
Stack heat capacity correction	ϵ_s	0.067	
Cold exchanger length	Δx_C	6.35	mm
Hot exchanger length	Δx_H	2.54	mm
Exchanger fin thickness	$2l^{EX}$	0.254	mm
Exchanger fin separation	$2y_0^{EX}$	0.508	mm
Operating frequency	f	328	Hz
Gas thermal penetration depth	δ_κ	0.108	mm
Stack thermal penetration depth	δ_s	0.010	mm
Gas viscous penetration depth	δ_ν	0.070	mm

Fig. 2 (the insulating ring has been left out of the drawing for simplicity). The heat flow rate, \dot{Q}_H , is the exhaust heat flux of the thermoacoustic engine and \dot{Q}_d is the heat leak from the copper flange to the driver housing. The heat flux measured with the thermopile/flow-meter combination is the exhaust heat flux less the amount that escapes to the driver housing. The thermal resistance of the Delrin™ insulating ring was measured to be 1.0 K/W. Two Type-E thermocouples, one mounted on the exhaust heat sink flange and the other on the driver housing, allow measurement of \dot{Q}_d and correction for the measured value of exhaust heat flux. The temperature difference measured by these thermocouples was at the least experimentally zero (at lower input power)

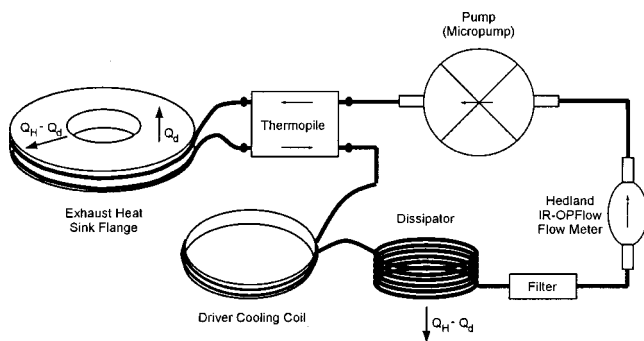


FIG. 2. Schematic of the exhaust heat sink flange and heat flux measurement instrumentation. The small arrows show the water flow direction and the large arrows indicate heat flux.

and at the most 7 °C (at the highest operating power).

The water that absorbs heat from the exhaust heat sink flange must also be cooled somewhere in the loop if a temperature difference is going to be maintained across the flange inlet and discharge. This is accomplished by submerging a coil of copper tubing in a large (50 gal) bucket of water. The dissipator is expected to dissipate this same amount of heat from the water loop that is introduced at the exhaust heat sink flange. The flow loop is diverted to loop twice around the driver housing after the discharge from the heat sink flange in order to reduce \dot{Q}_d by keeping the driver housing and the flange at the same temperature. The flow diversion also prevents the driver from overheating. No heat flux measurements are made across the driver cooling coil.

The accuracy of this measurement system is verified by running the closed water loop through two reservoirs of water without the inclusion of the refrigerator. One reservoir is the 50 gal bucket of water mentioned above. For calibration purposes, the closed flow loop is allowed thermal contact with another very small ($\frac{1}{4}$ gal) reservoir of water (in place of the refrigerator) that is heated with an electrical resistance heater. The part of the loop that is immersed in this small reservoir is a coil of copper tubing. After the system reached steady state (approximately 15 min) the standard deviation between the measured power to the electrical heater and the measured power dissipated by the flow loop is 0.052 W. This small variability ($0.052 \text{ W}/5.27 \text{ W}=1.0\%$) is within an acceptable limit for these measurements.

D. Heat exchanger performance model

In order to compare experimental data to the DELTAE model, knowledge of the stack end temperatures (hot and cold side) is necessary. Space qualification of the resonator in the STAR project required that no penetrations were made in the pressurized resonator for signal leads. Since the thermocouples are mounted outside the resonator and not directly on the fins, a method to infer the temperature of the fins (and hence the stack temperature) is needed. A simple mathematical calculation for the parallel thermal resistance of the 50 fins of the heat exchanger allows the temperature of the central fin region of the heat exchanger to be estimated. The hot side thermocouple is located on the copper flange where the hot heat exchanger is seated and the cold side thermocouple is located just below the cold side heat exchanger.⁹

The addition of a fin to a surface is usually motivated by a need for greater convective heat transfer from or to the surface. However, the addition of surface area (and therefore material) also carries with it some thermal resistance that depends partly on the fin geometry. Since the potential for heat transfer to or from a surface depends on the temperature of the surface and the temperature of the surrounding medium, the best performance from a fin would be realized if that fin were at the same temperature across the whole length of the fin. However, because a fin has nonzero thermal resistance, the fin sustains a temperature gradient and performs below this optimum. The fin efficiency is a multiplicative constant¹⁰ that expresses the degradation of heat transfer due

to the addition of the thermal resistance of the fin material and is defined as

$$\eta_f \equiv \frac{q_f}{q_{\max}} = \frac{q_f}{hA_f(T_b - T_\infty)}, \quad (1)$$

where q_f is the actual heat transferred by the fin, q_{\max} is the amount which would be transferred if the fin resistance were zero, h is the convective heat transfer coefficient, A_f is the surface area of the fin, T_b is the temperature at the base of the fin, and T_∞ is the temperature of the medium surrounding the fin. By symmetry we assure that the end of the fin is insulated from the medium, therefore q_f is

$$q_f = \sqrt{hP_c K_f A_c} (T_b - T_\infty) \tanh(mL), \quad (2)$$

where P_c and A_c are the cross-sectional perimeter and area of the fin, K_f is the thermal conductivity of the fin material, and the critical length $1/m$ is defined by $m^2 = hP_c / K_f A_c$. With a little manipulation, the fin efficiency [Eq. (1)] can be expressed as

$$\eta_f = \frac{\tanh(mL)}{mL}. \quad (3)$$

All of the parameters that make up the critical length, $1/m$, are geometric or tabulated except for the convective heat transfer coefficient, h . A common dimensionless parameter that includes the convective heat transfer coefficient is the Nusselt number which can be thought of as the ratio of total heat transfer to conductive heat transfer. This is expressed in Eq. (4) spatially averaged over a plate of length L ,

$$\text{Nu} \equiv \frac{\bar{h}L}{K}, \quad (4)$$

where K is the thermal conductivity of the fluid. Scaling arguments¹¹ (similitude) suggest that the Nusselt number over the length of a plate is correlated to the Reynolds number over the length and the Prandtl number (a nondimensional ratio of a fluid's viscous to thermal diffusivity). The standard correlation¹⁰ for steady, laminar flow that allows calculation of the heat transfer coefficient over the length of the plate, \bar{h} , in terms of these dimensionless parameters is

$$\text{Nu} = 0.664 \text{Re}_L^{1/2} \text{Pr}^{1/3}. \quad (5)$$

The above expression leads to a value of the convective heat transfer coefficient over the width of a fin in the heat exchanger. The value for \bar{h}_H at the hot side heat exchanger is $700 \text{ W/m}^2\text{K}$ for a flow velocity that corresponds to a pressure ratio of 3%.

However, in oscillatory flow the validity of this correlation is not clear. In this case, a time-averaged convective heat transfer coefficient, $\langle \bar{h} \rangle$, may have more relevance and the standard laminar correlation is modified as follows. Since the Nusselt number depends on the square root of the free-stream flow velocity (the Reynolds number is proportional to velocity) and this velocity is oscillating sinusoidally, a time average value over half a period (because the heat transfer is independent of the sign of the velocity) is used. This time average is¹²

TABLE III. Heat exchanger model summary.

Parameter	Hot exchanger	Cold exchanger
A_c	$1.62 \times 10^{-6} \text{ m}^2$	$6.45 \times 10^{-7} \text{ m}^2$
P_c	$1.32 \times 10^{-2} \text{ m}$	$5.59 \times 10^{-3} \text{ m}$
$\langle \bar{h} \rangle$	$530 \text{ W/m}^2\text{K}$	$1000 \text{ W/m}^2\text{K}$
m	104 m^{-1}	147 m^{-1}
R	0.17 K/W	0.30 K/W

$$\frac{\langle u_\infty^{1/2} \rangle}{u_\infty^{1/2}} = \frac{1}{\pi} \int_0^\pi (\sin x)^{1/2} dx = 0.763. \quad (6)$$

A more rigorous approach, which produces a similar numerical value, has been obtained by Mozurkewich.¹³

The fin efficiency, a value less than 1, increases the value for the effective thermal resistance of the fin,

$$R_{\text{fin}} = \frac{1}{\langle \bar{h} \rangle \eta_f A_f}, \quad (7)$$

where $\langle \bar{h} \rangle$ is the convective heat transfer coefficient averaged over the length of the fin and over time. The time average heat transfer coefficient corresponding to the hot side and cold side flow at a 3% pressure ratio is shown in Table III. With these values of $\langle \bar{h} \rangle$, the characteristic length, $1/m$, can be calculated using values for the fin cross-sectional area and cross-sectional perimeter, A_c and P_c , respectively. The results are shown in Table III. Since each fin has a different length, the total parallel thermal resistance, including the fin efficiency, is calculated with a computer. Since the resonator body is thin copper, it is assumed that there is very little thermal resistance from the inside of the resonator body to the external thermocouple.

E. Measurement procedure

The performance of Frankenfridge was measured in eight controlled experiments. The device was operated in the same manner for each of these experiments. First, the refrigerator was turned on with no heat load and the cold side was stabilized at 17.2°C by adjusting the acoustic pressure amplitude created by the driver. This modest temperature difference was chosen to minimize the heat leak from the room to the cold resonator parts. The frequency of the driver piston oscillation was adjusted so that an in-phase relationship between pressure and velocity at the face of the driver piston was maintained. Once the machine reached a steady state at the preestablished cold side temperature, the electric heat load was activated to dissipate $\frac{1}{2} \text{ W}$ at the cold side. The acoustic pressure was readjusted to keep the cold side at 17.2°C . Once this new state stabilized (taking about 6 to 8 mins), the heat load was again incremented by $\frac{1}{2} \text{ W}$ and the acoustic pressure adjusted accordingly to maintain a constant cold resonator temperature.

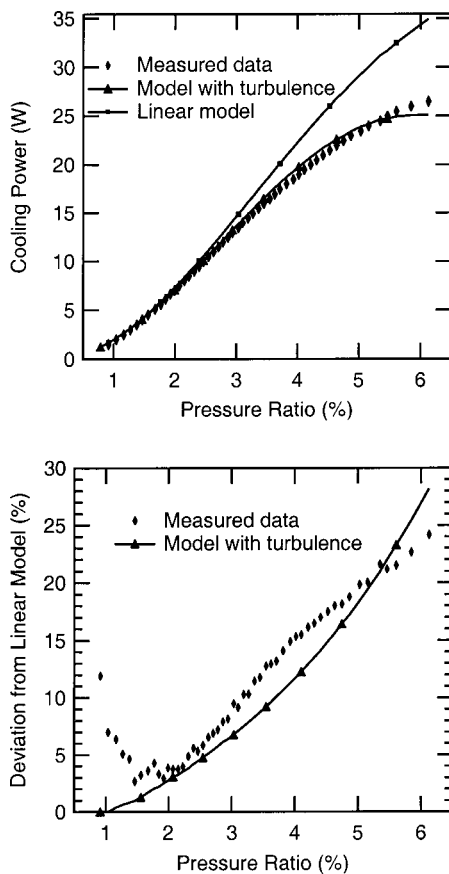


FIG. 3. Comparison of heat pumping power. The outside cold metal temperature is 17.2 °C.

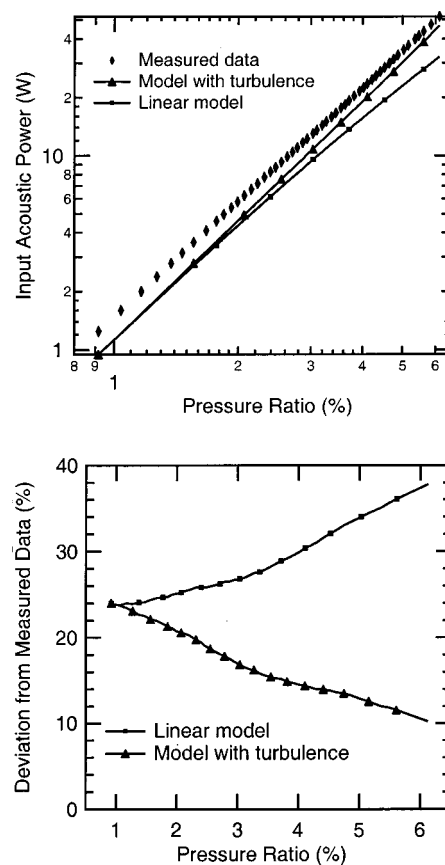


FIG. 4. Comparison of acoustic input power. The outside cold metal temperature is 17.2 °C.

II. RESULTS

A. Performance at low amplitudes ($P_A/p_m \leq 3\%$)

The performance of the refrigerator is illustrated in Figs. 3 and 4. Figure 3 shows the measured and modeled cooling power as a function of the ratio of peak-to-mean pressure in the resonator. Figure 4 is a plot of the input acoustic power required to attain a certain pressure ratio. In the presence of a stack and heat exchangers, this acoustic power is responsible for pumping heat from the side of the stack that gets cold to the side that gets hot.

The DELTAE model was generated by entering each experimental datum into the model, one at a time, and letting the model generate a solution for each point. For the comparisons in Figs. 3 and 4, the hot and cold side heat exchanger temperatures were targets for DELTAE and the program was allowed to solve for the input acoustic power and heat pumping power that correspond to the target temperatures. The acoustic pressure at the specific datum point was given to DELTAE as an invariant input parameter, allowing DELTAE to very accurately model an experimental condition.

The acoustic power required to attain a certain pressure ratio (as seen in Fig. 4) under the experimental condition is approximately 25% greater than DELTAE predicts it to be in the low-amplitude region. In Fig. 3 at low amplitudes, the graph of heat pumping power (or cooling power) shows agreement with DELTAE to within the accuracy of this experiment.

B. Performance at high amplitudes ($P_A/p_m \leq 6\%$)

The high-amplitude data (from 3% to 6% pressure ratio) is presented in Figs. 3 and 4 as well. The data in Fig. 3 show the nonlinear degradation in cooling power between the measured data and the (linear) DELTAE results above a 3% drive ratio. This nonlinear performance degradation is modest enough that the gain in power density associated with operation at higher pressure amplitudes may be worth the sacrifice in efficiency for some applications.

The discrepancy in the acoustic power required to attain a specific pressure ratio evident in Fig. 4 is greater in this high-amplitude region than is seen in the data below 3% pressure ratio. The cause of the discrepancy is undetermined at this time. In the following section we explore some possible loss mechanisms that may explain the disagreement between model predictions and measured data.

C. Resonator losses

In an attempt to explore the disagreement between model predictions and measured data, DELTAE models were created that include the effects of turbulent flow, a larger surface area for enhanced thermoviscous loss, and a different stack geometry that may more closely represent the rolled Mylar™ and fishing line stack used in this refrigerator.

Instead of using a STKS1ab segment in DELTAE (which models a parallel plate stack and does not account for spacers that support and separate the plates), a STKREct (which models the stack as rectangular pores) was substituted. It is

possible that this stack model more closely represents the Mylar™ film/fishing-string stack used in Frankenfridge. This modification decreases the acoustic power discrepancy by only a few percent and is not shown on a graph.

Because the discrepancy of DELTAE predictions and measured data at low amplitudes suggests an inaccurately modeled thermoviscous loss, comparison of the quality factor of the resonator to the quality factor predicted by the DELTAE model should reflect the discrepancy in input acoustic power. By driving the resonator with a negligibly small pressure oscillation, Q was measured to be 11.0. The Frankenfridge model in DELTAE can be programmed to produce a frequency “sweep” at very low pressure amplitudes to allow a determination of the model resonator’s Q . This sweep was executed at a constant volume velocity and at such very low pressure amplitudes so as not to cause DELTAE to form a temperature gradient across its stack. A half-power point fit for Q from this pressure response curve yields a value of 11.75.

Since the stack provides the dominant thermoviscous loss in the resonator (and it seems to be modeled well in DELTAE), the effects of smaller thermoviscous losses in the resonator could be magnified if the stack is removed. With the stack removed from the resonator of the refrigerator (and the heat exchangers left in place) the input power required to attain several values of pressure ratio was measured. The results of this experiment are compared to a DELTAE model that includes HX**** (finned heat exchanger) segments but no stack segment. The model uses the same values for acoustic pressure that were imparted into the physical resonator. The quality factor of the physical stackless resonator was measured to be 56. The model predicts Q to be 66 for the model that has no stack. The fact that the measured quality factor for the stackless refrigerator is smaller than the model predicts leads us to believe that there is some part of the resonator that is not understood, or not modeled in DELTAE correctly.

Along these same lines, the surface area of the cold reducer, cold duct, and trumpet was increased by 30% in DELTAE. This model, with a lower value for Q , causes the discrepancy to shrink by only 2% across the entire amplitude range.

The parameter that DELTAE uses to quantify the turbulence of the flow is the relative roughness. In a duct, this is defined¹¹ as the ratio of the surface roughness, e , to the diameter of the duct: e/D . A typical value¹¹ of surface roughness (e) for drawn tubing is 5×10^{-6} in. The result from the DELTAE model with turbulence that best matches the measured cooling power degradation corresponds to a relative roughness value of 0.1 as shown in Fig. 3. This value was assigned to the cold reducer, the cold duct, and the trumpet. A relative roughness of 10% is outrageously large for hardware of the size and type in this refrigerator but it is a way to get DELTAE to model a nonlinear loss that is on the order of that measured in Frankenfridge. In this only academically interesting limit, the discrepancy in acoustic power becomes as low 10% at the highest measured amplitude as compared to almost 40% in the linear model.

In summary, none of the above changes to the DELTAE

model correct the acoustic power discrepancy to within experimental error, nor do they destroy the agreement of cooling power for a given pressure ratio below pressure ratios of 3%.

III. CONCLUSIONS

There is a measurable deviation in heat pumping power of Frankenfridge from the power predicted with a linear acoustic computer model for the same pressure amplitude. This deviation in heat pumping power at 3% pressure ratio is about 8% and at 6% pressure ratio is about 24%. For some applications, the correspondingly poorer coefficient-of-performance may be an acceptable sacrifice in the face of a dramatically increased power density. This rolloff in heat pumping power could be due in part to the fact that the acoustic particle displacement is four to five times higher than the heat exchanger length—a high-amplitude design would likely have longer heat exchanger fins which might lessen the cooling power degradation.

The exhaust heat flux measurement system has allowed the device to trade the complicated vacuum insulation used in the STAR experiments for a simple fiberglass insulation shield and to increase the rate of data acquisition. The new driver can create stack power densities which are almost six times greater than could be obtained with the STAR driver.

While these experiments have shown excellent agreement with computer models of the stack, the resonator performance shows significant deviations from the DELTAE model. The computer model predicts the acoustic power requirements to be almost $\frac{2}{3}$ of the measured acoustic input power needed to reach a certain pressure ratio. This deviation is consistent in experiments with and without the stack. The discrepancy is greater in the stackless resonator model because the stack is the dominant loss mechanism and reduces the effect of the resonator model inaccuracy.

A preliminary analysis shows that the power discrepancy may be caused by minor losses in the stack, heat exchangers, and horn outlet of the machine. The term “minor loss” refers to nonlinear loss that occurs in the flow around obstacles, through orifices, and at abrupt changes in cross-sectional area or flow direction: this loss mechanism is not accounted for in the surface-roughness/turbulence module in DELTAE.

Not reported in this paper are measurements of the coefficient of performance for Frankenfridge. These measurements exhibit a disagreement with DELTAE predictions similar to that shown in this paper. Interestingly, a comparison of 1993 STAR coefficient of performance data with a DELTAE model of STAR shows the same disagreement.

ACKNOWLEDGMENTS

The authors gratefully thank the Office of Naval Research, NASA through the Pennsylvania Space Grant Consortium, and the Applied Research Lab through the Exploratory and Foundation Program for generous financial support. We thank Robert Smith for his effective technical assistance and Greg Swift and Bill Ward of Los Alamos National Laboratories for providing several upgrades to

DELTA-E. The BLKData and RPNTarget features were particularly useful in the analysis of this experimental data. Also, Anthony Atchley and Ray Wakeland provided insightful conversations and guidance during this project.

- ¹W. C. Ward and G. W. Swift, "Design environment for low-amplitude thermoacoustic engines," *J. Acoust. Soc. Am.* **95**, 3671–3672 (1994).
- ²W. C. Ward and G. W. Swift, *Design Environment for Low Amplitude ThermoAcoustic Engines, Version 3.04b* (software available from Energy Science and Technology Center, U.S. DOE, Oak Ridge, TN, <http://www.lanl.gov>).
- ³M. W. Shelley, *Frankenstein; or, The Modern Prometheus* (London, Lackington, 1818).
- ⁴S. L. Garrett, J. A. Adeff, and T. J. Hofler, "Thermoacoustic refrigerator for space applications," *J. Thermophys. Heat Transfer* **7**, 595–599 (1993).
- ⁵S. C. Ballister and D. J. McKelvey, "Shipboard Electronics Thermoacoustic Cooler," Master of Science Thesis, Physics Department, Naval Postgraduate School, 1995; DTIC Report No. ADA 300514; S. L. Garrett, U.S. Patent No. 5,647,216 (1997).

- ⁶Servometer Corporation, 501 Little Falls Road, Cedar Grove, NJ 07009.
- ⁷T. J. Hofler, "Accurate acoustic measurements with a high-intensity driver," *J. Acoust. Soc. Am.* **83**, 777–786 (1988).
- ⁸R. S. Wakeland, "Use of electrodynamic drivers in thermoacoustic refrigerators," *J. Acoust. Soc. Am.* **107**, 827–832 (2000).
- ⁹J. A. Adeff, "Measurement of the Space Thermoacoustic Refrigerator Performance," Master of Science Thesis, Physics Department, Naval Postgraduate School, 1990; DTIC Report No. AD A241-320.
- ¹⁰F. P. Incropera and D. P. DeWitt, *Fundamentals of Heat and Mass Transfer*, 3rd ed. (Wiley, New York, 1990), pp. 491–497.
- ¹¹R. W. Fox and A. T. McDonald, *Introduction to Fluid Mechanics*, 4th ed. (Wiley, New York, 1992), Chap. 7.
- ¹²S. L. Garrett, "Thermoacoustics life sciences refrigerator: heat exchanger design and performance prediction," June (1992); S. L. Garrett, D. K. Perkins, and A. Gopinath, "Thermoacoustic refrigerator heat exchangers: design, analysis and fabrication," *Heat Transfer 1994* (Proceedings 10th International Heat Transfer Conference) (1994), Vol. 4, pp. 375–380.
- ¹³G. Mozurkewich, "Time-average temperature distribution in a thermoacoustic stack," *J. Acoust. Soc. Am.* **103**, 380–388 (1998).

Thickness shearing vibration of the tangentially polarized piezoelectric ceramic thin circular ring

Lin Shuyu

*Applied Acoustics Institute, Shaanxi Teachers University, Xian, Shaanxi, 710062,
People's Republic of China*

(Received 11 May 1999; revised 4 January 2000; accepted 13 January 2000)

The thickness shearing vibration of the tangentially polarized piezoelectric ceramic thin circular ring is studied. The electrical and mechanical characteristics of the ring are analyzed in detail. The electro-mechanical equivalent circuit, the input electrical impedance, and the frequency equations for the resonance and anti-resonance frequencies are derived. A new concept that is characterized by the cross sectional shape and dimension is presented. An important conclusion is obtained that the thickness shearing vibration of the tangentially polarized piezoelectric ceramic ring is different from the traditional thickness extensional and radial extensional vibrations in that the electro-mechanical conversion coefficient, the electro-mechanical coupling coefficient, and the resonance frequency equation depend not only on the material parameters and the longitudinal dimension, but also on the lateral dimension, such as the cross sectional radius. It is shown that the measured resonance frequencies are in agreement with the theoretically calculated results, and the theoretical relationship between the resonance frequency and the cross sectional radius is also verified. © 2000 Acoustical Society of America. [S0001-4966(00)03904-7]

PACS numbers: 43.38.Fx [SLE]

INTRODUCTION

Piezoelectric ceramic transducers have been used for many years as sensors and transmitters of acoustic signals. In underwater acoustics and high power ultrasonics, low-frequency Langevin transducers that consist of two or more identical piezoelectric elements sandwiched between a back and front cylinder are widely used for finding fish, deep-sea seismology, sounding, ultrasonic cleaning, ultrasonic soldering, ultrasonic atomization, and other applications. The other principal applications of piezoelectric ceramic transducers are in medical ultrasound, nondestructive testing, and ultrasonic imaging, where they are used for ultrasonic diagnosis, ultrasonic therapy, material evaluation, and flaw detection.

Piezoelectric ceramic transducers or resonators have many vibrational modes. The widely used vibrational modes of transducers are the thickness extensional mode of a thin piezoelectric ceramic plate, the radial vibration of a thin ceramic circular disk or ring, the thickness shear vibration of a thin ceramic rectangular plate, and the longitudinal vibration of a slender piezoelectric ceramic cylinder. For these traditional vibrational modes, their analytical and design theory has been well established,¹⁻⁵ and widely used in the design of piezoelectric ceramic resonators that are employed in piezoelectric transformer, piezoelectric filter, ultrasonic emitter, and receiver.

Unfortunately, very little work has been carried out on the thickness shearing vibrational mode of piezoelectric ceramic circular elements. In recent years, ultrasonic motor and other ultrasonic applications such as ultrasonic soldering and material fatigue testing have been paid more and more attentions.⁶⁻⁹ In these applications, thickness shearing and longitudinal-torsional compound vibration are needed. Accordingly, the requirement for the insight into the thickness shearing vibration of piezoelectric ceramic elements be-

comes more and more pronounced. In our previous work, the torsional vibration of the slender piezoelectric ceramic cylinder and the piezoelectric ceramic stack polarized in the tangential direction was studied.¹⁰ Peter H. Ceperley analyzed the rotating waves in ring resonators and cylindrically symmetric resonators.¹¹⁻¹³ The aim of the present work is to perform an analytical modeling of the behavior of the thickness shearing vibration of the piezoelectric ceramic thin ring polarized in tangential direction. The electrical and mechanical characteristics of the ring in thickness shearing vibration are analyzed. An electro-mechanical equivalent circuit that is very useful in the analysis and design of the piezoelectric devices is derived. The effect of the cross sectional shape and dimension of the element on some important parameters of the piezoelectric ceramic ring polarized in tangential direction is studied. Finally, the measured results of the resonance frequencies of some piezoelectric ceramic circular rings in thickness shearing vibration are given.

I. THICKNESS SHEARING VIBRATION OF PIEZOELECTRIC CERAMIC RINGS POLARIZED IN TANGENTIAL DIRECTION

The tangentially polarized piezoelectric ceramic ring in thickness shearing vibration is shown in Fig. 1. The outer and inner radii are R_2 and R_1 , its thickness is l . The arrow in the circumferential direction also represents the polarization direction. The external exciting electric field is in the thickness direction of the ring. E_z is the external exciting electric field. For the thickness shearing vibration of the piezoelectric ceramic thin ring, the radial boundary can be regarded as clamped since the lateral dimension is much larger than its thickness. The displacement is in the tangential direction, while the propagation direction of thickness shearing vibration is in the thickness direction. On the other hand, since the

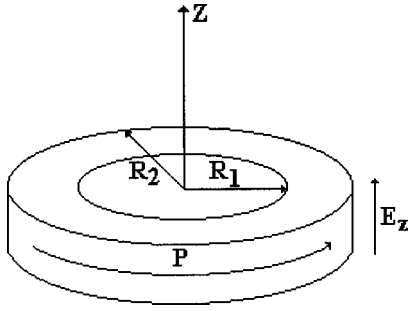


FIG. 1. A tangentially polarized piezoelectric ceramic ring.

piezoelectric ceramic material is isolated and the edge effect is neglected, we have $D_r = D_\theta = 0$, $D_z \neq 0$, $\partial D_z / \partial z = 0$. In this case, the piezoelectric constitutive equations can be expressed as

$$T_{\theta z} = c_{55}^D S_{\theta z} - h_{15} D_z, \quad (1)$$

$$E_z = -h_{15} S_{\theta z} + \beta_{11}^s D_z. \quad (2)$$

Here, $T_{\theta z}$, $S_{\theta z}$ are the tangential stress and strain; E_z , D_z are the axial electric field and electric displacement; c_{55}^D , h_{15} , and β_{11}^s are the elastic stiffness under the condition of constant electric displacement, the piezoelectric stiffness constant, and the clamped permittivity. If the thickness shearing angle displacement is φ , for the thickness shearing vibration of a thin piezoelectric ceramic ring, the relation between the thickness shearing angle and the tangential strain is

$$S_{\theta z} = r \partial \varphi / \partial z. \quad (3)$$

A. Mechanical characteristics of the thickness shearing thin ring

When the piezoelectric ceramic thin ring vibrates in thickness shearing vibration, the torsional moment can be expressed as

$$M = \int_s \int_s T_{\theta z} \cdot r \, ds. \quad (4)$$

Here $S = \pi(R_2^2 - R_1^2)$ is the cross sectional area of the piezoelectric ceramic thin ring. Substituting Eq. (1) into (4) yields

$$M = c_{55}^D I_p \partial \varphi / \partial z - h_{15} D_z W. \quad (5)$$

Here $I_p = \int_s \int_s r^2 \, ds = \pi(R_2^4 - R_1^4)/2$ is the moment of the cross section $W = \int_s \int_s r \, ds = 2\pi(R_2^3 - R_1^3)/3$. For the segment of dz in the ring, the equilibrium equation of the thickness shearing moment is

$$(\partial M / \partial z) dz = I (\partial^2 \varphi / \partial t^2). \quad (6)$$

Here $I = \int_s \int_s r^2 \, dm = \int_s \int_s r^2 \rho \, dz \, ds = \rho \, dz \, I_p$ is the moment of inertia of the segment dz . Substituting Eq. (5) into (6) yields

$$\partial^2 \varphi / \partial t^2 = c_t^2 (\partial^2 \varphi / \partial z^2). \quad (7)$$

Here $\partial D_z / \partial z = 0$ is used, and $c_t = (c_{55}^D / \rho)^{1/2}$ is the sound speed of thickness shearing vibration in the tangentially polarized piezoelectric ceramic thin ring. For harmonic vibration, substituting $\varphi = \Phi(z) \exp(j\omega t)$ into Eq. (7) yields

$$d^2 \Phi(z) / dz^2 + k_t^2 \Phi(z) = 0. \quad (8)$$

Here $k_t = \omega / c_t$ is the wave number of thickness shearing vibration in the tangentially polarized piezoelectric ceramic thin ring. From Eq. (8), the thickness shearing angle can be obtained:

$$\varphi = [A \sin(k_t z) + B \cos(k_t z)] \exp(j\omega t). \quad (9)$$

Here A and B are constants that can be determined by the boundary conditions. Let φ_1 and φ_2 be the thickness shearing angles at the two ends of the ring, i.e., $\varphi(z=0) = \varphi_1$, $\varphi(z=l) = -\varphi_2$. Substituting these two expressions into Eq. (9) and solving for A and B , we can get the thickness shearing angle displacement function,

$$\varphi = \frac{\varphi_1 \sin[k_t(l-z)] - \varphi_2 \sin(k_t z)}{\sin(k_t l)}. \quad (10)$$

Substituting Eq. (10) into (5) yields the thickness shearing moment function,

$$M = -c_{55}^D I_p k_t \times \frac{\varphi_1 \cos[k_t(l-z)] + \varphi_2 \cos(k_t z)}{\sin(k_t l)} - h_{15} D_z W. \quad (11)$$

Let the external thickness shearing moments exerted on the two ends of the ceramic ring be M_1 and M_2 , i.e., $M(z=0) = -M_1$, $M(z=l) = -M_2$. Substituting these two expressions into Eq. (11) yields

$$M_1 = c_{55}^D I_p k_t \times \frac{\varphi_1 \cos(k_t l) + \varphi_2}{\sin(k_t l)} + h_{15} D_z W, \quad (12)$$

$$M_2 = c_{55}^D I_p k_t \times \frac{\varphi_2 \cos(k_t l) + \varphi_1}{\sin(k_t l)} + h_{15} D_z W. \quad (13)$$

B. Electrical characteristics of the thickness shearing thin ring

Substituting Eq. (3) into (2) yields

$$E_z = -h_{15} r (\partial \varphi / \partial z) + \beta_{11}^s D_z. \quad (14)$$

Integrating Eq. (14) on the cross sectional area of the ring yields

$$E_z = -h_{15} W (\partial \varphi / \partial z) / S + \beta_{11}^s D_z. \quad (15)$$

Let the voltage across the ring be V , using Eq. (15) we have

$$V = \int_0^l E_z \, dz = -h_{15} W / S \int_0^l (\partial \varphi / \partial z) \, dz + \beta_{11}^s D_z l. \quad (16)$$

Using the above definitions, we have

$$\int_0^l (\partial \varphi / \partial z) \, dz = -\varphi_2 - \varphi_1. \quad (17)$$

Substituting Eq. (17) into (16) yields

$$V = h_{15} W / S (\varphi_2 + \varphi_1) + \beta_{11}^s D_z l. \quad (18)$$

Let the current into the ring be I , we have

$$I = dQ / dt = j\omega Q = j\omega D_z S. \quad (19)$$

Here Q is the electric charge on the end surface of the ring. Substituting Eq. (19) into (18) and solving for I , we have

$$I = j\omega C_0 V - n(\dot{\phi}_1 + \dot{\phi}_2). \quad (20)$$

Here $C_0 = S/(\beta_{11}^s l)$ is the clamped capacitance of the ring in thickness shearing vibration, $\dot{\phi}_1 = j\omega\phi_1$, $\dot{\phi}_2 = j\omega\phi_2$, $\dot{\phi}_1$ and $\dot{\phi}_2$ are the thickness shearing angular velocities at the two ends of the ring, and $n = h_{15}W/(\beta_{11}^s l)$ is the electro-mechanical conversion coefficient of the piezoelectric ceramic ring in thickness shearing vibration. It can be seen that it is different from the traditional thickness shear vibration of a piezoelectric ceramic rectangular thin plate in which the electro-mechanical conversion coefficient is expressed as

$$n = h_{15}A/(\beta_{11}^s l). \quad (21)$$

Here A is the area of the electrode surface of the rectangular thin plate in thickness shear vibration.

C. Electro-mechanical characteristics of the piezoelectric ceramic thin ring in thickness shearing vibration

1. Electro-mechanical equivalent circuit of the piezoelectric ceramic thin ring in thickness shearing vibration

From Eq. (18), we have

$$nV = nh_{15}W/S(\dot{\phi}_2 + \dot{\phi}_1) + n\beta_{11}^s D_z l. \quad (22)$$

Substituting the expression of the electro-mechanical conversion coefficient into Eq. (22) yields

$$nV = (\dot{\phi}_2 + \dot{\phi}_1)n^2/(j\omega C_0) + D_z h_{15}W. \quad (23)$$

Using Eq. (23), Eqs. (12) and (13) can be rewritten as

$$M_1 = C_{55}^D I_p k_t \times \frac{\phi_1 \cos(k_t l) + \phi_2}{\sin(k_t l)} - (\dot{\phi}_1 + \dot{\phi}_2)n^2/(j\omega C_0) + nV, \quad (24)$$

$$M_2 = C_{55}^D I_p k_t \times \frac{\phi_2 \cos(k_t l) + \phi_1}{\sin(k_t l)} - (\dot{\phi}_1 + \dot{\phi}_2)n^2/(j\omega C_0) + nV. \quad (25)$$

Let $Z_t = \rho I_p c_t$, Z_t is the characteristic impedance of the ring in thickness shearing vibration, and Eqs. (24) and (25) can be further rewritten as

$$M_1 = \left[\frac{Z_t}{j \sin(k_t l)} - \frac{n^2}{j\omega C_0} \right] (\dot{\phi}_1 + \dot{\phi}_2) + jZ_t \tan(k_t l/2) \dot{\phi}_1 + nV, \quad (26)$$

$$M_2 = \left[\frac{Z_t}{j \sin(k_t l)} - \frac{n^2}{j\omega C_0} \right] (\dot{\phi}_1 + \dot{\phi}_2) + jZ_t \tan(k_t l/2) \dot{\phi}_2 + nV. \quad (27)$$

Using Eqs. (20), (26), and (27), the electro-mechanical equivalent circuit of the piezoelectric ceramic thin ring in thickness shearing vibration can be obtained as shown in Fig. 2. In the figure, $Z_1 = jZ_t \tan(k_t l/2)$, $Z_2 = Z_t/[j \sin(k_t l)] - n^2/(j\omega C_0)$. It can be seen from Fig. 2 that it is similar to the electro-mechanical equivalent circuits of the longitudinal vi-

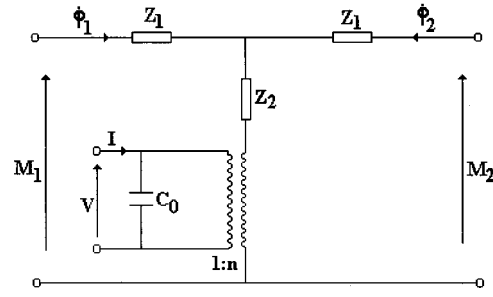


FIG. 2. Electro-mechanical equivalent circuit of the piezoelectric ceramic thin ring polarized in tangential direction in thickness shearing vibration.

bration of a slender piezoelectric ceramic rod polarized in the longitudinal direction, the thickness extensional vibration of a thin piezoelectric ceramic plate polarized in the thickness direction, and the thickness shear vibration of a thin piezoelectric ceramic rectangular plate polarized in the width direction. Therefore, the thickness shearing vibration in the thin piezoelectric ceramic circular ring is a kind of vibration of longitudinal effect, or a kind of stiffness vibration. On the other hand, since the propagation direction of the thickness shearing vibration in the ring is perpendicular to the particle vibrational displacement, it is a kind of transverse wave.

2. Frequency equation of the piezoelectric ceramic thin ring in thickness shearing vibration

When the piezoelectric ceramic thin ring is free of external torques at the two ends, i.e., $M_1 = M_2 = 0$. From Fig. 2, the input electrical impedance of the ring can be obtained:

$$Z_i = \frac{1}{j\omega C_0} \times \left[1 - (K_{15}^t)^2 \frac{\tan(k_t l/2)}{k_t l/2} \right]. \quad (28)$$

Here $K_{15}^t = h_{15}W/(\beta_{11}^s S I_p c_{55}^D)^{1/2}$ is the electro-mechanical coupling coefficient of the tangentially polarized piezoelectric ceramic thin ring in thickness shearing vibration. Let $T = (S I_p / W^2)^{1/2}$, T is defined as the cross sectional shape factor of the ring in thickness shearing vibration, and K_{15}^t can be rewritten as

$$K_{15}^t = K_{15}/T. \quad (29)$$

Here $K_{15} = h_{15}/(\beta_{11}^s c_{55}^D)^{1/2}$ is the electro-mechanical coupling coefficient of a rectangular piezoelectric ceramic thin plate in thickness shear vibration. It can be seen that for the thickness shearing vibration of a thin piezoelectric ceramic ring, its electro-mechanical coupling coefficient depends not only on the material parameters, but also on the cross sectional shape and dimension; while for the thickness shear vibration of a piezoelectric ceramic rectangular thin plate, its electro-mechanical coupling coefficient depends only on the material parameters. From the above analysis, it can be seen that the introduced parameter T is a function of the cross sectional shape and dimension of the ring. It can be further expressed as

$$T = \sqrt{\frac{9(R_1 + R_2)^2(R_1^2 + R_2^2)}{8(R_1^2 + R_2^2 + R_1 R_2)^2}}. \quad (30)$$

If the inner radius is much less than its outer radius, i.e., $R_1 \ll R_2$, the ring can be approximated to a circular disk. We have

$$T = (9/8)^{1/2}. \quad (31)$$

On the other hand, if the inner radius is approximate to its outer radius, i.e., $R_1 \approx R_2$, we have

$$T = 1. \quad (32)$$

In this case, the electro-mechanical coupling coefficient of the piezoelectric ceramic thin ring in thickness shearing vibration is reduced to that of a piezoelectric ceramic rectangular thin plate in thickness shear vibration. This is obvious. The reason is that when the inner radius is approximate to the outer radius, the ring can be regarded as a long rectangular strip approximately. In any other cases, the cross sectional shape factor of the ring can be given as

$$1 \leq T \leq (9/8)^{1/2}. \quad (33)$$

From Eq. (28), when the input electrical impedance is zero, the resonance frequency equation can be obtained:

$$1 - (K'_{15})^2 \frac{\tan(k_t l/2)}{k_t l/2} = 0. \quad (34)$$

When the input electrical impedance is infinity, the anti-resonance frequency equation can also be obtained:

$$\tan(k_t l/2) = \infty. \quad (35)$$

3. Effect of the cross sectional shape and dimension on the resonance frequency of the piezoelectric ceramic circular ring in thickness shearing vibration

For some traditional vibrational modes of piezoelectric ceramic elements, such as the longitudinal vibration of a slender piezoelectric ceramic rod polarized in longitudinal direction and the thickness shearing vibration of a piezoelectric ceramic thin rectangular plate polarized in the width direction, their resonance frequency is determined by the material parameters and the dimension in the propagation direction of wave. The dimension perpendicular to the propagation direction of wave has no effect on the resonance frequency when one-dimensional theory is assumed. For the thickness shearing vibration of a piezoelectric ceramic ring polarized in the tangential direction, however, it can be seen from Eq. (34) that the resonance frequency of the tangentially polarized piezoelectric ceramic thin ring in thickness shearing vibration depends not only on the material parameters and the thickness of the ring, but also on the cross sectional shape and dimension. From Eqs. (28) and (30), the input electrical impedance and the cross sectional shape factor of the ring can be rewritten as

$$\frac{Z_i}{Z_0} = \frac{k_{t0} l}{k_t l(1 - B^2)} \times \left[1 - \frac{K_{15}^2}{T^2} \times \frac{\tan(k_t l/2)}{k_t l/2} \right], \quad (36)$$

$$T = \sqrt{\frac{9(1+B)^2(1+B^2)}{8(1+B+B^2)^2}}. \quad (37)$$

Here, $B = R_1/R_2$, $k_{t0} l = \pi$, $k_{t0} = \omega_0/c_t$, $\omega_0 = 2\pi f_0$, $Z_0 = 1/(j\omega_0 C_s)$, $C_s = S_0/(\beta_{11}^s l)$, $S_0 = \pi R_2^2$, and f_0 is the first

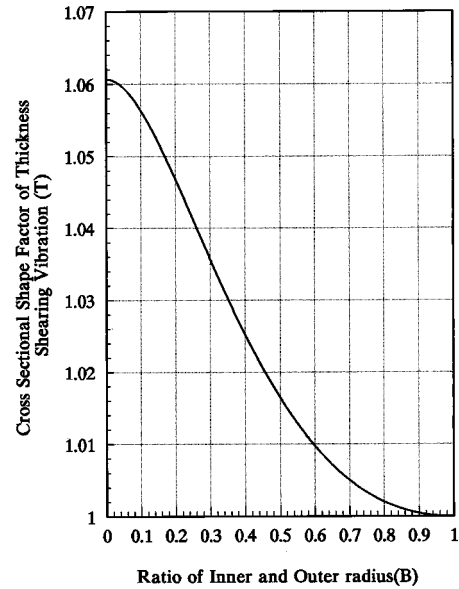


FIG. 3. Theoretical relationship between the cross sectional shape factor and the radius ratio of the ring.

anti-resonance frequency of the tangentially polarized piezoelectric ceramic thin ring in thickness shearing vibration. Figure 3 shows the relationship between the cross sectional shape factor T and the radius ratio B . It can be seen that the variation range of the cross sectional shape factor of the ring in thickness shearing vibration is consistent with Eq. (33).

From Eqs. (36) and (37), we can get the relationship between the normalized input electrical impedance Z_i/Z_0 and the nondimensional phase length $k_t l$ of the thickness shearing ring as shown in Fig. 4. In the figure, B is chosen as a parameter. The solid line, the dotted line, and the dashed and dotted line represent three different values of B of 0, 0.35, and 0.75, respectively.

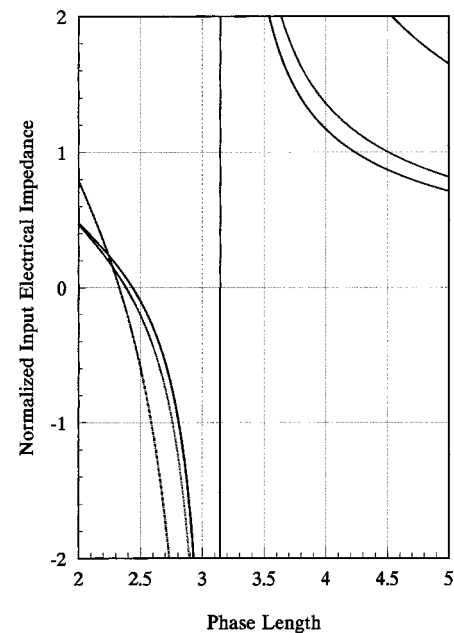


FIG. 4. Theoretical relationship between the normalized input electrical impedance and the nondimensional phase length (the solid line, the dotted line, and the dashed and dotted line represent three different values of B of 0, 0.35, and 0.75, respectively).

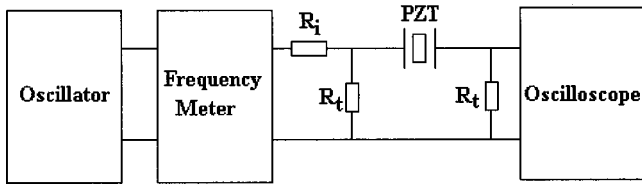


FIG. 5. Experimental setup for the measurement of the resonance frequency of the resonator.

0.35, and 0.75, respectively. From Fig. 4, it can be seen clearly that for different values of radius ratio, the resonance frequency of the ring changes. When the radius ratio is increased, the resonance frequency of the ring in thickness shearing vibration is decreased. This means that when the resonator changes from a circular disk to a ring, its resonance frequency is decreased. This is different from traditional vibrational modes of piezoelectric ceramic elements in which the resonance frequency is determined by the thickness.

II. EXPERIMENTS

Experimental investigations have been done to verify the theoretical analysis and to measure the resonance frequencies of the piezoelectric ceramic thin ring in thickness shearing vibration. The piezoelectric ceramic material is an equivalent of PZT-4, its standard material parameters are as follows: $\rho = 7.5 \times 10^3 \text{ (kg/m}^3\text{)}$, $K_{15} = 0.71$, $c_{55}^D = 5.18 \times 10^{10} \text{ (N/m}^2\text{)}$. The transmission line method is used for the measurement of the resonance frequency of the piezoelectric ceramic circular ring in thickness shearing vibration. The experimental setup is shown in Fig. 5. In Fig. 5, R_i and R_t are two noninductive electrical resistances, R_i is much larger than R_t , PZT represents the resonator in thickness shearing vibration to be measured. The resonance frequency is measured according to the following procedures: Change the frequency of the oscillator until the output of the oscilloscope has a maximum. The frequency corresponding to this maximal output is the maximal admittance frequency; this is the resonance frequency of the test resonator under first approximation. In general cases, if the loss in the resonator is low, this approximation is satisfactory. This method is widely used in the resonance frequency measurement of piezoelectric ceramic elements. The geometrical dimensions, the calculated frequencies and measured frequencies are listed in Table I. In the table, f is the calculated frequency, and f_m is the measured frequency, $\Delta = |f_m - f|/f_m$.

In the experiment, the piezoelectric ceramic circular ring polarized in the tangential direction is composed of eight small sectors of the same dimension. The manufacturing process of the tangentially polarized ceramic ring is as follows. First, cut a ceramic circular ring into eight small sectors with

the same dimensions, then polarize every small sector in its tangential direction. Second, erase the electrodes from the rectangular cross section of the sector and cement the eight sectors into a circular ring. Third, coat silver powder on the two end surfaces of the ring as the electrodes. Finally, connect the two end electrodes of the ceramic ring into the measuring circuit and the resonance frequency of the ring can be measured.

From Table I, it can be seen that the measured resonance frequencies of the resonator in thickness shearing vibration are basically in agreement with the calculated results. As for the frequency error, the following factors should be taken into account. First, the standard material parameters are different from those of the practical material used. Second, the tangentially polarized ring is composed of a number of small sectors; the cementing of these small sectors can cause some mechanical loss and/or additional mechanical compliance. Third, the polarization of the piezoelectric ceramic sectors is nonuniform; the polarization direction is not in ideal tangential direction. Finally, since the process to manufacture the piezoelectric ceramic ring is crude, the loss cannot be ignored, and the mechanical quality factor is not very high. Therefore, the maximal admittance frequency is different from the resonance frequency, and the approximation in the measurement of the resonance frequency can cause some error definitely.

From the measured results, the effect of the cross sectional shape and dimension on the resonance frequency of the ceramic ring in thickness shearing vibration can be explained. In Table I, the thickness of No. 1 and No. 2 rings are the same, since the values of B are different, their resonance frequencies are also different, and the variation law is consistent with the theoretical relationship in Fig. 4.

III. CONCLUSIONS

The thickness shearing vibration of tangentially polarized piezoelectric ceramic thin ring is studied theoretically and experimentally. The electrical and mechanical characteristics are analyzed, and the electro-mechanical equivalent circuit is obtained. A new concept defined as the cross sectional shape factor of the piezoelectric ceramic thin ring in thickness shearing vibration is introduced. Theoretical analyses show that the electro-mechanical coupling coefficient and the resonance frequency of the ceramic thin ring depend not only on the material parameters, but also on the transverse dimensions. This is different from the traditional vibrational modes such as the thickness shear vibration of a rectangular thin plate and the thickness extensional vibration. Experimental results show that the calculated resonance frequencies are in good agreement with the measured results, and the

TABLE I. Geometrical dimensions and the resonance frequencies of the thickness shearing rings.

No.	l (mm)	R_1 (mm)	R_2 (mm)	B	f (Hz)	f_m (Hz)	$\Delta(\%)$
1	2.8	3.0	5.0	0.60	350 087	342 148	2.32
2	2.8	4.0	20.0	0.20	360 412	351 765	2.46
3	3.5	6.0	10.0	0.60	280 069	273 968	2.23
4	3.5	6.0	20.0	0.30	286 076	280 152	2.11

effect of the cross sectional shape and dimensions on the resonance frequency is verified. The theory developed in this work can be used in the design and analysis of circular thickness shearing transducers and ultrasonic motors.

- ¹M. Brissaud, "Characterization of piezoceramics," IEEE Trans. Ultrason. Ferroelectr. Freq. Control **38**, 603–617 (1991).
- ²Philips, *Piezoelectric Ceramics: Properties and Applications* (N. V. Philips Gloeilampenfabrieken, 1991), pp. 42–45.
- ³W. P. Mason, *Physical Acoustics*, Vol. 1, Part A (Academic, New York, 1964).
- ⁴O. E. Mattiat, *Ultrasonic Transducer Materials* (Plenum, New York, 1971).
- ⁵IRE Standards on Piezoelectric crystals: Measurements of piezoelectric ceramics, Proc. IRE **49**, 1161 (1961).
- ⁶S. E. Stanzl-Tschegg, H. R. Mayer, and E. K. Tschegg, "High frequency method for torsional fatigue testing," Ultrasonics **31**, 275–280 (1993).
- ⁷Y. Tomoikawa, K. Adachi, and M. Aoyagi, "Some constructions and characteristics of rod-type piezoelectric ultrasonic motors using longitudinal and torsional vibrations," IEEE Trans. Ultrason. Ferroelectr. Freq. Control **39**, 600–608 (1992).
- ⁸K. Nakamura, M. Kurosawa, and S. Ueha, "Characteristics of a hybrid transducer-type ultrasonic motor," IEEE Trans. Ultrason. Ferroelectr. Freq. Control **38**, 188–193 (1991).
- ⁹K. Nakamura, M. Kurosawa, and S. Ueha, "Design of a hybrid transducer type ultrasonic motor," IEEE Trans. Ultrason. Ferroelectr. Freq. Control **40**, 395–401 (1993).
- ¹⁰Lin Shuyu, "Torsional vibration of coaxially segmented, tangentially polarized piezoelectric ceramic tubes," J. Acoust. Soc. Am. **99**, 3476–3480 (1996).
- ¹¹Peter H. Ceperley, "Acoustic spinal wave field," J. Acoust. Soc. Am. **93**, 2278 (1993).
- ¹²Peter H. Ceperley, "Observation of single-point excitation and shocking of rotating acoustic waves," J. Acoust. Soc. Am. **91**, 2331(A) (1992).
- ¹³Peter H. Ceperley, "Rotating waves," Am. J. Phys. **60**, 938–942 (1992).

Interaction between wave number pairs for a capped cylinder

J-F. Ille and J. H. Ginsberg

G. W. Woodruff School of Mechanical Engineering, Georgia Institute of Technology, Atlanta, Georgia 30332-0405

(Received 12 March 1999; accepted for publication 11 January 2000)

When acoustic-structure interaction is analyzed according to the wave number-based version of the surface variational principle (SVP), considerable computational effort is required to generate each coefficient in the quadratic sum forming the variational quantity. Examination of the wet-surface impedance, which represents the spectrum of pressure amplitudes generated by a specified spectrum of surface velocity amplitudes, reveals that many cross-impedance terms are very small. This suggests that some coefficients need not be computed. The article introduces *a priori* criteria for selecting the impedance terms to be omitted, based on the supersonic cutoff wave number. Each truncation scheme is assessed by comparing its predictions to the convergent SVP solution. It is shown that, for the nonsymmetric azimuthal harmonics, the field quantities and the radiated power are well predicted if subsonic waves are ignored. In contrast, for the axisymmetric component, substantial errors (6 dB or more) for radiated power arise, unless a broad spectrum of subsonic waves are included in the formulation. The complex power is mostly reactive, being associated with an evanescent field, even for the supersonic spectrum. The study shows that small interactions between subsonic waves in the axisymmetric case can result in the high wave number spectrum being associated with substantial radiation contributions. © 2000 Acoustical Society of America. [S0001-4966(00)02904-0]

PACS numbers: 43.40.Ey, 43.40.Rj [CBB]

INTRODUCTION

Aside from a few cases where the wetted surface suits separation of variables, acoustic radiation from elastic structures requires substantial computational resources. Boundary or finite element methods are widely used because of the relative simplicity of their implementation. Good entry points for the study of boundary elements are Schenck's paper on CHIEF (Schenck, 1968) and the paper by Ciskowski and Brebbia (1991), while Burnett (1994), and Harari and Hughes (1992) discuss finite elements issues. An alternative semi-analytical method, derived from the second surface Helmholtz equation, is offered by the surface variational principle (SVP) (Pierce, 1993). The wave number based implementation of SVP (Ginsberg and Wu, 1998) uses Ritz series to represent the surface response. This formulation provides more physical insight, and its convergence can be judged by examining the modal amplitudes of the higher basis functions relative to those of lower ones. Such advantages are offset by the fact that considerable computational effort is required to form the system equations.

A useful quantity that may be derived from any formulation of the fluid-structure interaction is the wet-surface impedance, which may be viewed as a Green's function from which the surface pressure corresponding to any specified surface velocity field may be obtained. Computation of this quantity is an inherent feature of SVP. When a body of revolution is analyzed by the SVP method, the meridional and azimuthal dependence of surface pressure and normal velocity are described by Fourier series. Because the series span the entire wetted surface, this description leads to an analysis in terms of interacting waves propagating along the surface.

The full implementation of SVP entails retention of all possible wave interactions.

The present work will examine the wet-surface impedance to show that many of these interactions are quite weak. This observation will lead to an exploration of the degree to which wave interactions may be ignored on an *a priori* basis in order to increase the computational efficiency of an SVP analysis. More importantly, an examination of the quantitative effect of omitting these interactions will shed light on the importance of supersonic and subsonic waves for acoustic radiation from slender capped cylinders. The data obtained by Wu and Ginsberg (1998) in their analysis of radiation from a slender capped cylinder are the starting point for the investigation.

I. BASIC PRINCIPLES

A cylindrical shell of diameter $2a$ and length L , closed at both ends by hemispherical shells of the same thickness, and fully submerged in an infinite fluid domain, provides the data that will be used to identify the essential features of the wet surface impedance. In Wu and Ginsberg's 1998 study of radiation from capped cylinders, the acoustic field on the wet surface is represented as a sum of waves propagating in the circumferential and meridional directions. Pressure, velocity, and time are nondimensionalized relative to ρc^2 , c , and a/c , respectively, where ρ and c are the ambient density and speed of sound in water. Harmonic steady-state response is represented by a factor $\exp(-ikat)$, which is suppressed. The pressure and normal velocity series are expressed as

$$\begin{Bmatrix} p(\alpha, \theta) \\ V_n(\alpha, \theta) \end{Bmatrix} = \frac{1}{4} \sum_{m=-\infty}^{\infty} \sum_{j=-N}^N \begin{Bmatrix} \rho c^2 \hat{P}_j^m \\ c \hat{V}_j^m \end{Bmatrix} e^{i(jk_0 \alpha + m \theta)}, \quad (1)$$

where θ is the azimuthal angle of a point on the wetted surface, α is the arclength along the shape generator formed by a meridian, nondimensionalized by the total arclength $s_0 = L + \pi a$, and $k_0 = \pi/s_0$. The coefficients \hat{P}_j^m and \hat{V}_j^m are nondimensional pressure and normal velocity amplitudes of a helical wave on the cylindrical surface. The amplitudes at various azimuthal wave numbers m are uncoupled by axisymmetry. Coupling of the meridional spectra j is described by the wet-surface impedance $[Z^m]$, which relates the pressure coefficients \hat{P}_j^m and velocity coefficients \hat{V}_j^m . Satisfaction of the surface variational principle, which pertains to the acoustic field resulting from a specified surface velocity, leads to a set of linear equations relating the two sets of coefficients, whose solution may be written as

$$\{P^m\} = [Z^m]\{V^m\}. \quad (2)$$

In addition to the acoustic relations, the structural dynamics laws must be satisfied. Wu and Ginsberg derived these relations by using the Ritz method. Love's assumptions, which accounts for both extensional and flexural deformations, were used to describe the shell. The displacement vector at a point on the surface is represented in terms of tangential components in the meridional and azimuthal directions and the normal component. The basis functions for the Ritz series are the *in vacuo* eigenmodes for a spherical shell, which are mapped onto the meridional arclength. Specifically, let s denote arclength measured from one apex, and let $\tilde{\Psi}_{jn}^m(\phi)$ denote the modal displacement (vector) of a spherical shell at polar angle ϕ , where j and m are the meridional and azimuthal wave numbers, and $n = 1, 2$, or 3 is the number for the branch number of the vibration mode. Then the basis function for the capped cylinder at arclength s is taken to be $\tilde{\Psi}_{jn}^m(\pi s/s_0)$. The Ritz series for each m extends over the same range of j as those for the surface pressure and normal velocity. The contribution of each basis function to each of the three shell displacement components is given by the value of the generalized coordinates q_{nj}^m . (The $j=3$ branch when $m=0$ is purely torsional, so it is not excited.)

The shell basis functions are not orthogonal relative to the capped cylinder's surface, so the structural generalized coordinates at each m are inertially and elastically coupled. The kinetic energy, the strain energy, and the virtual work done by the pressure field and excitation are expressed in terms of the displacement series. Lagrange's equations for these quantities, with each time derivative replaced by a $-ika$ factor, yields

$$[[K^m] - (ka)^2[M^m]]\{q^m\} = \{Q_f^m\} - \{Q_p^m\}, \quad (3)$$

where $\{q^m\}$ consists of the structural generalized coordinates for azimuthal harmonic m , $[K^m]$ is the stiffness matrix, and $[M^m]$ is the mass matrix. Also $\{Q_f^m\}$ and $\{Q_p^m\}$ represent, respectively, the generalized forces due to the excitation and the fluid pressure on the wetted surface.

The SVP and dynamics equations are coupled when continuity at the wetted surface is enforced. Matching the normal velocity yields

$$\{V^m\} = -ika[\Gamma^m]\{q^m\}. \quad (4)$$

The fluid and structure media also are coupled by the generalized forces due to fluid pressure on the wetted surface. Substitution of the Ritz series into the definition of the generalized forces leads to

$$\{Q_p^m\} = [\Lambda^m]\{P^m\}. \quad (5)$$

The elements of $[\Gamma^m]$ and $[\Lambda^m]$ are inner products of the basis functions for normal displacement and surface acoustic response. The difference between them is that $[\Lambda^m]$ contains the surface area metric in its definition.

The full set of equations is obtained by substituting Eq. (4) into Eq. (2), and Eq. (5) into Eq. (3), from which one obtains

$$[[K^m] - (ka)^2[M^m] - ika[\Delta^m]]\{q^m\} = \{Q_f^m\}, \quad (6)$$

where $[\Delta^m]$ is the structural impedance associated with fluid loading:

$$[\Delta^m] = [\Lambda^m][Z^m][\Gamma^m]. \quad (7)$$

Because $[\Lambda^m]$ and $[\Gamma^m]$ are real, it follows that the resistive (i.e., real) part of the wet surface impedance $[Z^m]$ is responsible for radiation damping, whereas the reactive (i.e., imaginary) part is responsible for the added mass effect.

Chen and Ginsberg (1995) showed that $[\Delta^m]$ should be symmetric as a consequence of physical laws, rather than its mathematical form. (This property will provide a focal point for the analysis that follows.) They showed that the complex power associated with each azimuthal harmonic m may be expressed as

$$\Pi = 0.5(ka)^2\{q^{m*}\}^T[\Delta^m]\{q^m\}, \quad (8)$$

where $()^*$ designates the complex conjugate of the argument. The reactive power, $\text{Im}(\Pi)$, is the part of the instantaneous power radiation that vanishes when averaged over a period. Its value relative to the radiated power, $\text{Re}(\Pi)$, provides information regarding the radiation efficiency.

Evaluation of the coefficients in the SVP quadratic sums forming the variational quantity, from which $[Z^m]$ is derived, require much computational effort. Each coefficient is obtained by a double integration along the meridian, in which the integrand contains products of the basis functions and θ -integrated free space Green's functions. The numerical mesh for the meridional integrations must be quite fine because the integrand contains an integrable line singularity that stems from the property of the Green's functions. Hence, being able to ignore the evaluation of some wet-surface impedance coefficients would result in significant reduction of the computational effort.

II. BEHAVIOR OF THE WET-SURFACE IMPEDANCE

Equation (2) indicates that waves having different values for the meridional wave number j , but belonging to the same azimuthal harmonic m , couple. The term Z_{nj}^m gives the pressure amplitude in the meridional spectrum $-N < n < N$ generated by a unit amplitude normal velocity wave at meridional wave number j . Diagonal terms Z_{jj}^m are "self-impedances," and off-diagonal terms are "cross-impedances." Because a hemi-capped cylindrical shell is symmetric about the transverse midplane, any response can

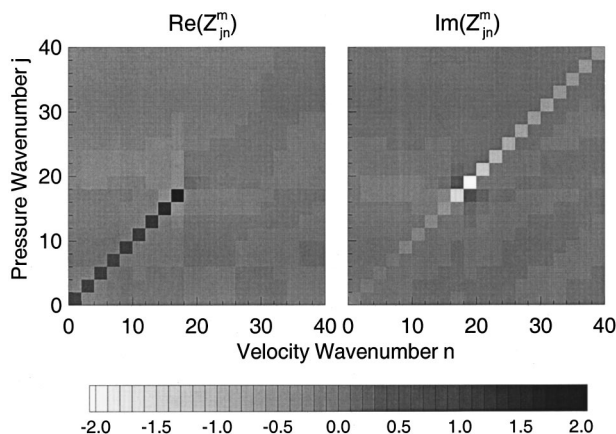


FIG. 1. Wet surface impedance contours, $L/2a=5$, $ka=4$, $m=0$.

be decomposed into parts that are symmetric and antisymmetric with respect to that plane. Such decoupling is manifested by the elements of $[Z^m]$ being zero if the row and column number have different parity. The like parity impedance values will be depicted as contour plots of $[Z^m]$ for the symmetric effects, which correspond to even row and column numbers for $m=0$ and odd numbers for $m \neq 0$. These contours will be drawn as two by two cells, whose lower left corner location corresponds to the wave numbers associated with the cell. (In essence, the lower left corner of the contour shows the 1,1 element of $[Z^m]$.)

The data set to be examined corresponds to a reasonably slender aspect ratio, $L/2a=5$, although similar behavior of the wet-surface impedance is obtained for $L/2a=2$. The excitation is a point force applied at the midplane. This system was analyzed in the mid-frequency range $1 < ka < 10$. Within this range, frequencies $ka=4$ and $ka=7$ illustrate the main trends across the entire frequency band.

Figure 1 depicts the real and imaginary parts of $[Z^m]$ for $ka=4$. Recall that the real part of the wet-surface impedance is responsible for the acoustic radiation damping, whereas the imaginary part is reactive. If the imaginary part is negative, it corresponds to an added mass effect, in which pressure is in-phase with acceleration, whereas a positive value is a springlike pressure loading. Inspection of the contours plot shows that, except for a region around the diagonal term j

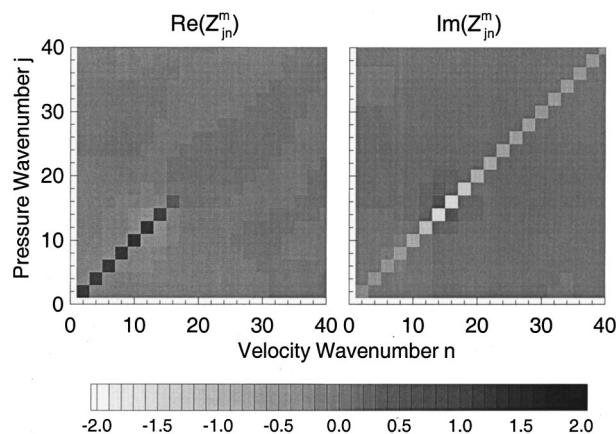


FIG. 3. Wet surface impedance contours, $L/2a=5$, $ka=4$, $m=3$.

$=n \approx 18$, which corresponds to the maximum impedance in each graph, the cross-impedance terms are very small.

Profiles showing the real and imaginary parts of the self-impedance terms are displayed in Fig. 2. The real part rises to a maximum at $j=16$, while the imaginary part, which is negative, falls to a minimum at $j=18$. The self-impedances $(Z_{int})_{jj}$ for waves on a infinite cylinder are also shown in this figure. (The cross-impedances for an infinite cylinder are identically zero.) It is clear from the infinite cylinder data that extrema correspond to a transition from supersonic to subsonic waves. Figures 3 and 4 show the impedance contours and self-impedance for $m=3$. Similar behavior to those depicted in Figs. 1 and 2 occur for other values of $L/2a$ and ka .

Analysis using cylindrical wave functions (Junger and Feit, 1986) shows the cutoff wave number between supersonic and subsonic waves, which corresponds to matching of the helical and acoustic wavelengths, to be

$$j_{\text{crit}} = \left(\frac{L}{\pi a} + 1 \right) \sqrt{(ka)^2 - m^2}. \quad (9)$$

For the values associated with Figs. 1 and 2 ($ka=4$, $m=0$, $L/2a=5$), $j_{\text{crit}} = 16.73 \approx 17$, which matches the observed transition range of the self-impedance terms. Similarly, in the case $ka=4$, $m=3$, $L/2a=5$, $j_{\text{crit}} = 11.06 \approx 11$, which matches the highlight in Figs. 3 and 4. A comparison of Figs.

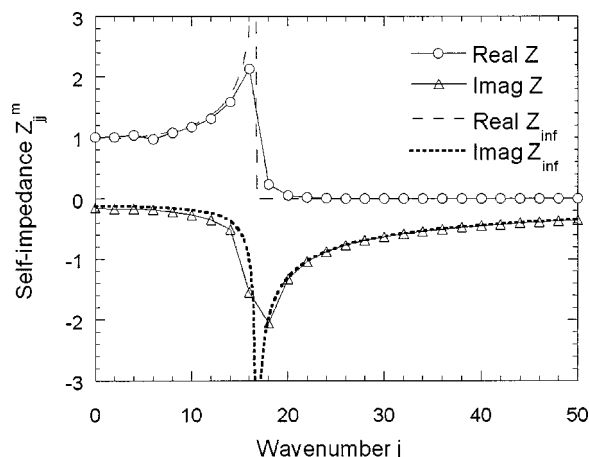


FIG. 2. Self-impedance, $L/2a=5$, $ka=4$, $m=0$.

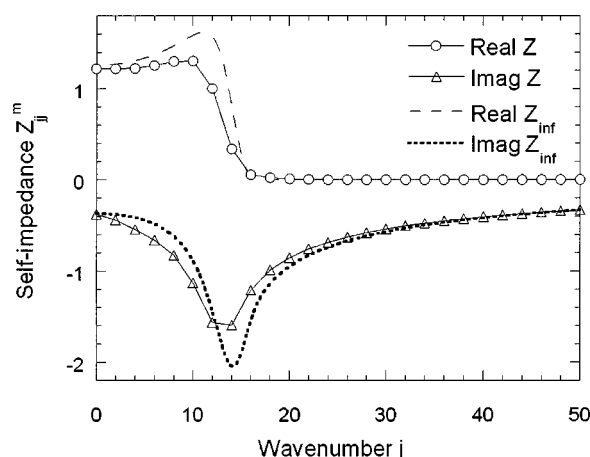


FIG. 4. Self-impedance, $L/2a=5$, $ka=4$, $m=3$.

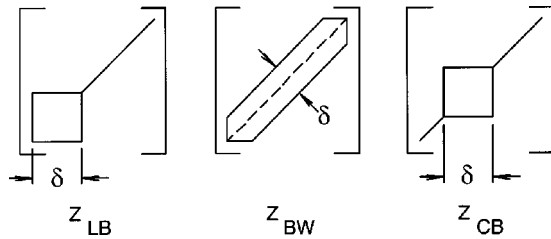


FIG. 5. Depiction of the truncation schemes for the wet-surface impedance.

1 and 3 shows that the patterns are essentially the same when m increases, aside from shifting the highlights from the diagonal at $j=17$ to the diagonal at $j=11$ as a result of the shift of j_{crit} . When $m \geq ka$, the entire meridional spectrum is subsonic. The supersonic waves correspond to waves that propagate away from the surface, whereas the subsonic waves correspond to evanescence of the radial field.

III. TRUNCATION OF THE IMPEDANCE MATRIX

Figures 1 and 3 exemplify a common attribute of $[Z^m]$. Most of the cross-impedances are quite small, except in the vicinity of the self-impedance at the cutoff wave number j_{crit} . When $ka=4$ with $m=0$ or $m=3$, 93% of the impedance terms are smaller than the largest impedance term by at least three orders of magnitude. When $ka=7$ with $m=0$ or $m=3$, this percentage is approximately 90%. It is reasonable to believe that there is no need to compute these small values. The validity of this hypothesis, which is a main objective of this article, will be examined as follows. Selected parts of $[Z^m]$ will be equated to zero, and the truncated $[Z^m]$ will be used to solve Eqs. (2)–(6). This solution represents an approximation of the structural and acoustical response. The quality of this approximation will be assessed by comparing it to the full SVP predictions of shell displacement, normal velocity, surface pressure, radiated power, and the acoustic field. For brevity, only azimuthal harmonics $m=0$ and $m=3$ at $ka=4$ and $ka=7$, which display the main trends for any case, will be discussed.

An examination of the wet-surface impedance in Figs. 1 and 3 suggests that the selection should keep at least the diagonal elements of $[Z^m]$, which are the self-impedances, and some cross-impedances in the vicinity of the transition wave number j_{crit} . Three truncation schemes, which differ in the selection of these off-diagonal terms, were considered in Ille's thesis (1998). They are the lower box method, the bandwidth method, and the centered box method, whose field quantities shall, respectively, be denoted with subscripts LB , BW , and CB . Each is characterized by a single truncation parameter δ . The lower box method keeps all self-impedances, plus the cross-impedances $Z_{n,j}^m$ satisfying $1 \leq n, j \leq \delta$. The bandwidth method keeps the self-impedances plus an equal number of sub-diagonals and super-diagonals, such that the bandwidth is δ . The centered box method keeps the self-impedances plus the terms $Z_{i,j}$ satisfying $j_{\text{crit}} - \delta/2 \leq i, j \leq j_{\text{crit}} + \delta/2$. Figure 5 displays the pattern of the impedance contours obtained by each of these methods.

If N denotes the size of the impedance matrix, the meth-

TABLE I. Number of wet-surface impedance terms retained in the lower box, centered box and bandwidth methods.

δ LB and CB method	8	10	12	14	16	18
# of terms to compute	117	151	193	243	301	367
δ BW method	3	5	7	9	11	12
# of terms to compute	181	299	415	529	641	751

ods require the evaluation of the following numbers of terms of $[Z^m]$:

Lower box method: $\delta^2 - \delta + N$

Bandwidth method: $\delta N - \frac{1}{4}(\delta^2 - 1)$

Centered box method: $\delta^2 - \delta + N$.

Table I presents the number of terms to be computed versus the value of the parameter for each method when $N=61$. It is evident that the bandwidth method for a given truncation number retains many more terms than the two others.

The center box method is similar to, but somewhat more restrictive than, the lower box method. In the center box method, one begins by including the off-diagonal impedances in the vicinity of the supersonic cutoff. Increasing the box size introduces more supersonic and subsonic waves, until it reaches the limit at which the lower left corner is at the lowest wave number. The lower box method begins with the lowest wave numbers, and can be increased without limit until all cross-impedances have been included. Ille (1998) reported that, for a given box size, the center box method never gives better results than the lower box method. Hence, only results obtained from the lower box and bandwidth method shall be discussed.

IV. ASSESSMENT

The shell displacement is defined by the values of the generalized coordinates $\{q^m\}$. A minimal approximation in any of the schemes retains only the diagonal elements of $[Z^m]$. When $m > 0$, the real and imaginary parts of $\{q^m\}$ have the same order of magnitude. In that case the largest error in $\{q^m\}$ obtained from the self-impedance terms relative to values obtained from the full $[Z^m]$ is 2%. In contrast, when $m=0$, the real part of $\{q^m\}$ is two orders of magnitude greater than the imaginary part. For the minimal approximation when $m=0$, the largest relative error in the real part is found to be 4%, while the largest relative error of the imaginary part reaches 30%, with an average error of 8%. However, such discrepancies in the imaginary part have little effect on the generalized coordinates magnitudes because of the dominance of the real part. Thus if one is solely interested in the structural displacements, any of the truncation schemes will suffice.

Small errors in $\{q^m\}$ obtained from the truncation schemes are magnified for acoustic variables, but the behavior is distinctly different for $m=0$ and $m > 0$. The velocity coefficients described by Fig. 6, for $m=0$, exhibit significant differences in the low wave number spectrum between either truncation and the full SVP analysis. The truncation numbers for this figure were selected such that the errors resulting from either scheme are comparable. Because this leads to the

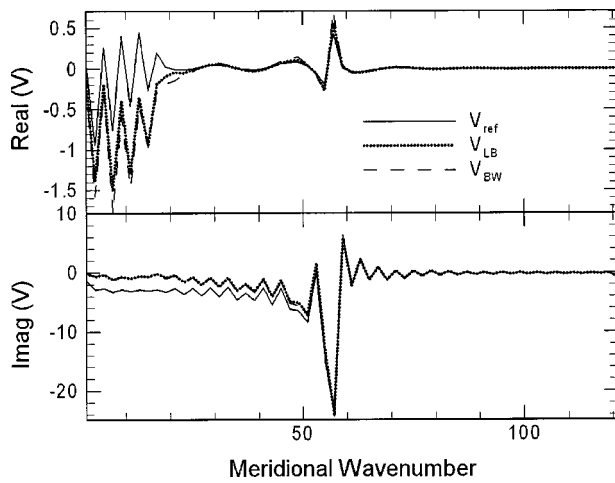


FIG. 6. Velocity coefficients, $L/2a=5$, $ka=4$, $m=0$, LB: $\delta=8$, 117 [Z^m] terms; BW: $\delta=3$, (181 [Z^m] terms).

retention of different number of terms, such selection of the truncation parameter sheds light on which scheme will be numerically more efficient. Notice that the ratio of the imaginary part of $\{V\}$ to the real part in the case of Fig. 6 is of the order of 10. The reactive power for $ka=4$, $m=0$, is 87 times larger than the radiated power. Figure 7 depicts the velocity coefficients for $m=3$. Contrary to the situation for axisymmetric terms, each approximate scheme for $m>0$ yields results for $\{V^m\}$ that are consistent with the full evaluation. When $m=3$, the reactive power is only a factor of 2.9 greater than the radiated power when $ka=4$, and 1.8 greater when $ka=7$.

An assessment of Fig. 7, as well as others not presented here (Ille, 1998), leads one to conclude that, for surface normal velocity when $m>0$, the bandwidth scheme is less approximate than the lower box method for a given truncation parameter δ . However, the number of terms it retains increases very rapidly with δ . For a given number of terms, the error in either scheme is lower than it is for $m=0$.

Truncation of $[Z^m]$ leads to pressure coefficients that show greater discrepancies than those displayed by the corresponding velocity coefficients. The explanation lays in the

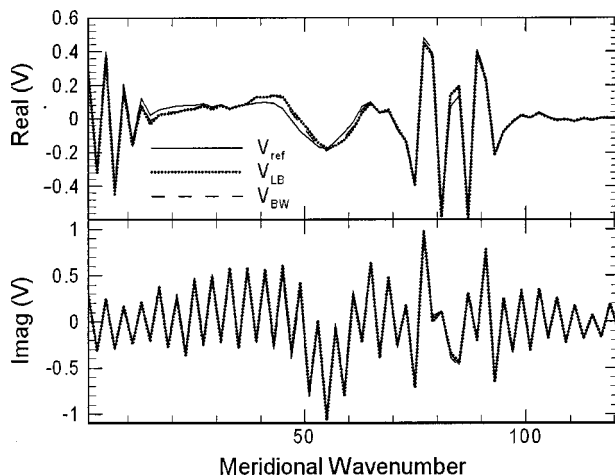


FIG. 7. Velocity coefficients, $L/2a=5$, $ka=4$, $m=3$, LB: $\delta=5$, (81 [Z^m] terms), BW: $\delta=3$ (181 [Z^m] terms).

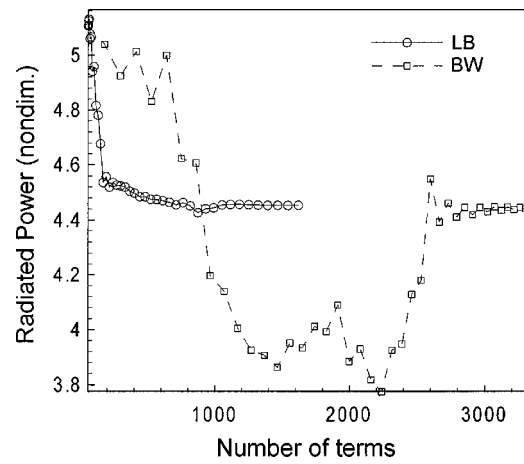


FIG. 8. Convergence of radiated power, $L/2a=5$, $ka=4$, $m=0$.

relations from which these quantities are computed. The velocity is derived directly from the generalized displacements $\{q^m\}$ according to $\{V^m\} = -ika[\Gamma^m]\{q^m\}$. The velocity coupling coefficients $[\Gamma^m]$ are not modified in any of the truncation methods, so good agreement for $\{q^m\}$ leads to good agreement for $\{V^m\}$. In contrast, the pressure is given by the relation:

$$\{P^m\} = -ika[Z^m][\Gamma^m][K^m] - (ka)^2[M^m] - ika[\Lambda^m][Z^m][\Gamma^m]^{-1}\{Q^m\}. \quad (10)$$

Hence, the error in pressure stems from truncation of $[Z^m]$ both directly and in the inverted term.

Acoustic power is another indicator of the ability of the approximate schemes to predict the acoustic response. Note that good results for the acoustic pressure and velocity amplitudes do not necessarily imply good agreement for the power, because the relative phase angle between these quantities is crucial. The radiated power for $m=0$ is very sensitive to modification of the impedance matrix, because the resistive portion of $[Z^m]$ is small compared to the added mass portion. For example, for $ka=7$, $m=0$, the largest error in the radiated power, obtained when $\delta=1$, is 14% for the lower box and bandwidth method, while for $ka=7$ and $m=3$, the results for $\delta=1$ are within a range of 2% for the lower box method and 5% for the bandwidth method. To understand the sensitivity of the power to the modification of $[Z^m]$ recall that the complex power is computed by $\Pi = 0.5(ka)^2\{q^{m*}\}^T[\Delta^m]\{q^m\}$. Changing $[Z^m]$ alters the power directly by modifying $[\Delta^m]$, and indirectly because $\{q^m\}$ depends on $[\Delta^m]$. According to Eq. (6), the latter dependence entails inversion of a matrix to which $[\Delta^m]$ is a contributor, and errors in a matrix tend to be magnified in an

TABLE II. Complex power, $ka=4$, $m=3$, $j_{\text{crit}}=6$, $\Pi_{\text{ref}}=2.39-6.95i$.

Box size	Π_{LB}	Bandwidth	Π_{BW}
2	2.32-7.24i	1	2.32-7.24i
4	2.37-7.08i	3	2.46-6.89i
6	2.4-7.05i	5	2.37-7.18i
8	2.4-7.02i	7	2.43-6.96i
10	2.39-6.98i	9	2.4-7.04i

TABLE III. Symmetry properties of the wet-surface impedance.

m	ka			
	4		7	
	0	3	0	3
$\max(\Delta)$	10.11	7.16	12.35	18.7898
$\max(\Delta - \Delta^T)$	0.6016	0.1255	0.1661	0.9424
$asym$	0.0595	0.0175	0.0134	0.05

inversion operation. Figure 8 shows that the power for $m=0$ obtained from the two truncation methods converges in a different manner when the truncation parameter is increased. The lower box method converges almost monotonically toward the solution, while convergence of the bandwidth method is oscillatory. As seen in Table II, the case $m>0$ converges very fast for any truncation schemes. It follows that the lower box method is preferable because it requires fewer terms than the bandwidth method.

V. SIGNIFICANCE OF RECIPROCITY

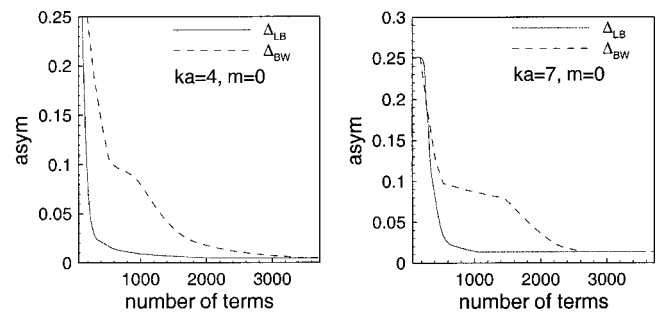
Chen and Ginsberg's proof that $[\Delta^m]$, defined in Eq. (6), is symmetric relies on physical principles. Any computed result for $[Z^m]$ will deviate from symmetry as a result of numerical errors, regardless of how it is obtained. Truncation of $[Z^m]$ will increase the deviation from symmetry. Thus the degree to which $[\Delta^m]$ is symmetric indicates the degree to which the truncation schemes are consistent with the laws of acoustics. A measure of the symmetry of the $[\Delta^m]$ matrix for each value of ka and m is a relative norm defined as

$$asym = \frac{\max(|\Delta - \Delta^T|)}{\max(|\Delta_{jn}|)} \quad (11)$$

Table III presents the $asym$ values obtained from SVP without truncation.

In order to ascertain whether the 1% to 6% deviation from symmetry for the untruncated $m=0$ data is significant, the acoustic field quantities and the radiated power were computed using a symmetricized $[\Delta^m]$. This entailed replacing $[\Delta^m]$ with $[\Delta^m]_s = 0.5([\Delta^m] + [\Delta^m]^T)$. The effect of this alteration was insignificant, with a maximum error of 4% in magnitude and 1° in phase for V_j^m . In addition, the basic SVP coefficients were recomputed for $m=0$, $ka=4$, and $ka=7$, which show the greatest deviation from symmetry. That computation used a refined numerical integration, in which the mesh size was decreased by a factor of 3. The results were an untruncated $[\Delta^m]$ for which $asym=0.0046$ when $ka=4$, which is 13 times smaller than the first set and $asym=0.0043$ when $ka=7$. Although the differences between response variables obtained with and without truncation were found to be insignificant, the refined SVP is now taken to be the standard. Doing so ensures that any conclusions regarding truncation error are not contaminated by errors in the reference solution.

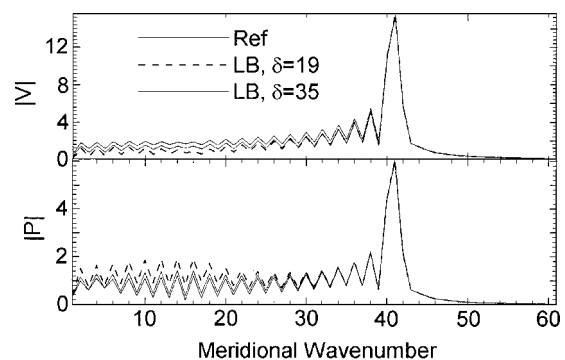
The next step is to examine the symmetry of $[\Delta^m]$ for each truncation method. For this, only the $m=0$ case needs to be examined, because the truncation schemes for $m>0$ were found in the previous section to give good results for

FIG. 9. Asymmetry parameter as a function of the number of terms retained in $[Z^m]$, $L/2a=5$, $ka=4$ and 7 , $m=0$.

small δ values. Figure 9 displays the convergence of $asym$ versus the truncation parameter for the lower box and bandwidth methods.

The cutoff wave number for $ka=4$, $m=0$ is $j_{crit}=17$, and $asym=0.0046$ when the full $[Z^m]$ matrix is used to compute $[\Delta^m]$. Figure 9 indicates that it is necessary to employ a truncation number greater than 26 (711 terms for LB method) to obtain a $[\Delta^m]$ matrix from the lower box method whose symmetry is comparable to that associated with the full impedance matrix. The bandwidth method requires 3000 terms to reach the same level of symmetry. The case $ka=7$, $m=0$ in Fig. 9 exhibits the same behavior. A lower box of size greater than 31 (991 terms) is necessary in this case to get a $[\Delta^m]$ matrix whose deviation is comparable to that computed with the full impedance matrix, while the same quality in the bandwidth method is attained when 2500 terms are retained ($\delta=50$).

Evidence that the symmetry quality of $[\Delta^m]$ indicates the correctness of the solution is presented in Fig. 10, which compares the pressure and normal velocities amplitudes obtained from $\delta=19$ (403 terms) and $\delta=35$ (1251 terms) when $ka=7$. Only the lower box method is depicted in these graphs because the bandwidth method retains many more elements of $[Z^m]$ for a given level of asymmetry. The cutoff wave number for $ka=7$, $m=0$ is $j_{crit}=29$. According to Fig. 10, if an insufficient number of terms are retained in $[Z^m]$, the acoustic field quantities at the higher wave number spectrum j are acceptable, but the values in the lower spectrum are not. A similar graph for $ka=4$ would show that the lower box scheme gives an $asym$ value that is comparable to that obtained from the full $[Z^m]$ matrix if $\delta \geq 27$. A response

FIG. 10. Magnitude of $\{V^m\}$ and $\{P^m\}$, $L/2a=5$, $ka=7$, $m=0$, LB: $\delta=19$ (403 $[Z^m]$ terms), BW: $\delta=35$ (1251 $[Z^m]$ terms).

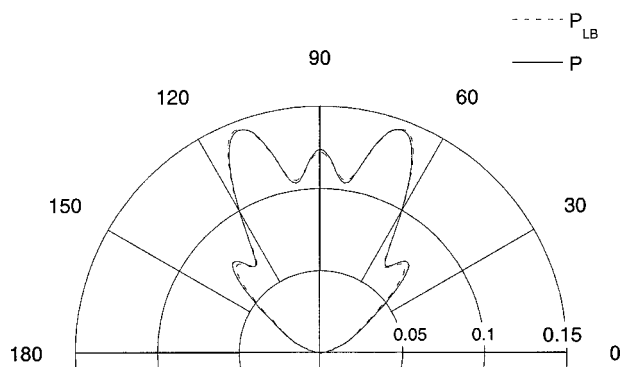


FIG. 11. Directivity of pressure, $ka=4$, $m=0$, $r=45a$, $\delta=29$, Lower box method.

calculation using $\delta=28$ for $ka=4$, $m=0$ yields results that agree very well with those obtained from SVP. For example, the computed $\text{Re}(\Pi)=4.372$ is extremely close to the reference value of 4.377.

Thus the symmetry of the fluid loading matrix obtained from a truncation method for $m=0$ is a direct indicator of the accuracy of the acoustical response obtained from that method. It is interesting to consider the behavior of $[\Delta^m]$ for $m>0$ from this perspective. When $m=3$, the *asym* value in either truncation method quickly converges to the value associated with the full impedance matrix. These trends are consistent with our earlier observations.

Setting some cross-impedance terms to zero is a useful simplification only if it leads to a reasonably accurate description of the radiated pressure field. Once the surface pressure and normal velocity have been evaluated, the pressure at any exterior point can be determined from the Kirchhoff–Helmholtz integral theorem. The following graphs display the pressure directivity in the azimuthal plane $\theta=\theta_0$ at a radial distance r/a and a polar angle Φ relatively to the symmetry axis of the shell. (The Fraunhofer far-field approximation was not used to compute the Kirchhoff–Helmholtz integral. Rather, the far-field location was identified by comparing the quantity $r^*|P(r,\Phi)|$ for increasing values of r and constant Φ , until a nearly constant value was attained.)

In order to obtain good agreement with the full evaluation of the far-field directivity, a lower box size with $\delta=29$ is needed in the case $ka=4$, $m=0$. This result is displayed in Fig. 11. The result of the full $[Z^m]$ is reproduced with $\delta=6$ when $ka=4$, $m=3$, as shown in Fig. 12. These values are comparable to the truncation parameters leading to good agreement of the normal velocity and surface pressure coefficients.

VI. CONCLUSIONS

The wave number version of the surface-variational principal describes the acoustic field on the surface in terms of a set of waves that propagate in the meridional and circumferential directions. The different azimuthal harmonics uncouple due to axisymmetry of the shape. Examination of the wet-surface impedance matrix, whose columns indicate the meridional spectrum of pressure waves that are generated by a velocity wave at a specific meridional wave number,

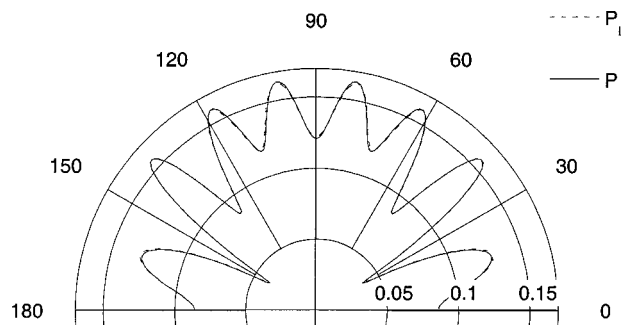


FIG. 12. Directivity of pressure, $ka=4$, $m=3$, $r=45a$, $\delta=6$, Lower box method.

shows that many cross-impedances are small compared to the self-impedances. The present analysis examined the influence of these cross-impedances for predictions of structural displacement, surface normal velocity, surface pressure, radiated power, and far-field directivity. The study was initially motivated by a desire to increase the computational efficiency of SVP by omitting evaluation of terms in the wet-surface impedance that are unimportant. Concurrently, an assessment of the alteration in response variables resulting from artificially removing the cross-impedance terms provides information regarding the physical importance of the associated wave interactions.

This study confirmed that it is only necessary to retain the self-impedance terms and selected cross-impedance terms. An important aspect of the selection of cross-impedance terms to be retained is the critical wave number j_{crit} at which the waves shed from the surface of an infinite cylinder transition from propagative to evanescent character. Three truncation schemes were tried. In each scheme, a single parameter defines the fraction of the full matrix to be computed. The “lower box” scheme, which retains interactions between all meridional wave numbers up to a certain maximum, was found to give the best result in terms of accuracy for a specified number of impedance terms.

Evaluation of the quality of the result relative to the truncation parameter showed that the axisymmetric azimuthal harmonic, $m=0$, behaves differently from the $m\neq 0$ harmonics. In the latter case, the radiated acoustic field quantities and power converge quickly as the truncation parameter is increased. In fact, reasonably good results for $m>0$ can be obtained by ignoring all cross-impedance effects. In contrast, when the azimuthal harmonic number is zero, the complex power is mostly instantaneous, and trapped in the evanescent field. All axisymmetric waves propagate at a zero helical angle. Consequently, coupling between waves having different meridional wavelength is strong, and the radiated power is sensitive to small errors. It is necessary to go well beyond the cutoff wave number to capture all the interactions between waves. An empirical guideline that fits the available data sets, including those not reported here, suggests that retention of all self-impedance terms Z_{jj}^m in the range $0\leq j\leq 2j_{\text{crit}}$ plus cross-impedance terms Z_{jn}^m in the range $0\leq j,n\leq j_{\text{crit}}+12$ should be adequate.

An important aspect of the study was demonstration that symmetry of the fluid-loading matrix $[\Delta^m]$ enables one to

verify *a priori* the adequacy of a truncation. This matrix represents the portion of the structure's dynamic stiffness that is attributable to fluid loading. A symmetry parameter measuring the norm of $[\Delta^m] - [\Delta^m]^T$ may be computed in any truncation method, without knowledge of the reference value associated with an analysis in which all wet-surface impedance terms are computed. It was examined as a function of the truncation parameter. When this parameter attains a 2% value, the results for the surface response, radiated power, and far-field directivity are consistent with the untruncated evaluations. Thus the truncation methods identified here, guided by the symmetry parameter, provide an economical way to obtain a correct structural and acoustical analysis, both on and outside the surface. The computational savings for SVP are significant, in that less than 35% of the system coefficients are required to obtain good results for $ka=7$ when $m=0$, while less than 10% of the coefficients are required for $ka=7$ when $m=3$. (These reductions pertain to the longer cylinder, $L/2a=5$, but the savings for $L/2a=2$ are comparable. Also, given the close agreement between the self-impedances for $L/2a=5$ and those for an infinite cylinder, it is reasonable to conjecture that these conclusions are equally valid for longer cylinders.) It is interesting to contemplate using symmetry of the fluid-loading impedance as an accuracy check for standard computational methods using BEM or FEM.

The data presented here also sheds light on an important physical aspect of the interaction between surface motion and pressure on slender, capped cylinders in the mid-frequency range. Specifically, the high meridional wave number spectrum of $[Z^m]$ at $m=0$ is required to obtain accurate values for the low wave number spectrum of the response, and hence for the radiated field. A hypothesis is that this behavior arises from constructive interference between short waves, which generate long waves at the difference of

the wave numbers of interfering pairs. The absence of this phenomenon when $m \neq 0$ may be explained from the perspective of helical waves. The waves for $m=0$ propagate solely in the meridional direction, so they may interact over long spatial scales. An accurate analysis must account for a broad spectrum of subsonic (short wavelength) $m=0$ surface waves, even though such waves themselves do not radiate. In contrast, the surface waves for $m \neq 0$ propagate over the surface in different helical directions. Their interactions are weak, so their self-impedances represent the dominant physical mechanism. A corollary is that any of the standard computational procedures should be zoned to capture the high wave number axisymmetric waves.

- Burnett, D. S. (1994). "A three-dimensional acoustic infinite element based on a prolate spheroidal multiple expansion," *J. Acoust. Soc. Am.* **96**, 2798–2816.
- Chen, P.-T., and Ginsberg, J. H. (1995). "Complex power, reciprocity, and radiation modes for submerged bodies," *J. Acoust. Soc. Am.* **98**, 3343–3349.
- Ciskowski, R. D., and Brebbia, C. A. (1991). *Boundary Element Methods in Acoustics* (Computational Mechanics Publications, Elsevier).
- Ginsberg, J. H., and Wu, K. (1998). "Nonaxisymmetric acoustic radiation and scattering from rigid bodies of revolution using the surface variational principle," *ASME J. Vibr. and Acoust.* **120**, 95–103.
- Harari, I., and Hughes, T. J. (1992). "A cost comparison of boundary element and finite element methods for problems of time-harmonic acoustics," *Comput. Methods Appl. Mech. Eng.* **97**, 77–102.
- Ille, J.-F. (1998). "Interaction of spatial scales in acoustic radiation from hemicapped cylinders," Master's thesis, Georgia Institute of Technology.
- Junger, M. C., and Feit, D. (1986). *Sound, Structures, and Their Interaction* (MIT Press, Cambridge).
- Pierce, A. D. (1993). "Variational formulations in acoustic radiation and scattering," *Phys. Acoust.* **22**, 195–371.
- Schenck, H. (1968). "Improved integral formulation for acoustic radiation," *J. Acoust. Soc. Am.* **44**, 41–58.
- Wu, K., and Ginsberg, J. H. (1998). "Mid-frequency range acoustic radiation from slender elastic bodies using the surface variational principle," *J. Vib. Acoust.* **120**, 392–400.

Acoustic radiation from an elastic baffled rectangular plate covered by a decoupling coating and immersed in a heavy acoustic fluid

Olivier Foin and Alain Berry

*Groupe d'Acoustique de l'Université de Sherbrooke, Département de Génie Mécanique,
Université de Sherbrooke, Sherbrooke, Québec J1K 2R1, Canada*

Jeffrey Szabo

Defence Research Establishment Atlantic, P.O. Box 1012, Dartmouth, Nova Scotia B2Y 3Z7, Canada

(Received 12 March 1997; revised 19 April 1999; accepted 16 February 2000)

The vibroacoustic behavior of an elastic, simply supported rectangular plate covered by a locally reacting decoupling layer supporting thickness deformation is presented. The model simulates the vibration and acoustic response of the system immersed in water and subjected to a point force disturbance. A simplified version of the theory is derived in the limiting case of a large decoupling (low mechanical impedance of the layer/high frequency). An appropriate vibratory indicator, representative of the acoustic attenuation provided by the decoupling treatment, and independent of the structure dimensions, is also investigated from the perspective of small-scale laboratory characterization © 2000 Acoustical Society of America. [S0001-4966(00)00405-7]

PACS numbers: 43.40.Rj, 43.30.Jx, 43.40.Dx, 43.40.Tm [CBB]

INTRODUCTION

The addition of a decoupling (or compliant) coating on a resonant structure is a common way to reduce the sound radiated by a structure immersed in water. The aim of this treatment is to isolate the fluid from the vibrations of the resonant structure. The decoupling material should have a low stiffness in comparison to the elastic material forming the base structure to support thickness deformation induced by the action of the dynamic pressure in the fluid. This deformation leads to the decoupling of the motions of the base structure and of the fluid in contact with the decoupling layer. The decoupling effect is small if the structure is immersed in a light fluid because the acoustic pressure exerted by a light fluid on the decoupling layer is negligible.

The objective of this article is to propose a theoretical model to describe the vibroacoustic behavior of a rectangular, simply supported elastic plate covered by a decoupling layer and immersed in water in order to quantify the sound attenuation provided by such a system. The rectangular plate is inserted in an infinite rigid baffle that separates vacuum on one side from water on the other side. The vibration of the base plate, the vibration of the outer surface of the decoupling layer, and the sound radiated in the fluid half-space are investigated when the excitation is a point force applied on the base structure.

In order to accurately model the dynamics of the decoupling layer, the three-dimensional theory of elasticity is usually required. This theory ensures that thickness deformation as well as bending, shear, and extensional deformations of the decoupling layer are properly accounted for (Noor *et al.*, 1989; Garrison *et al.*, 1994; Laulagnet *et al.*, 1994). Such models devoted to finite structures lead to very complicated formulations that usually prevent parametric study and high-frequency investigations. In order to obtain simpler formulation, Ko (1997) and Keltie (1998) considered the three-

dimensional theory of elasticity but with infinite structures. A finite structure is under consideration in this paper, so that the three-dimensional elasticity theory is not used to model the decoupling layer.

Another category of models to describe the behavior of compliant coating is the locally reacting model, which assumes that the decoupling material behaves as evenly distributed massless springs on the plate. This model is able to account for the transverse strain of the decoupling layer. It is called “locally reacting” because the thickness deformation of the layer at a given position depends on the surface acoustic pressure at this position only. It also neglects the mass of the decoupling layer.

Numerous works have used the locally reacting model to describe the vibroacoustics of a base structure covered by a compliant material and immersed in water. Crighton (1979) investigated the sound reflected from an infinite elastic panel covered by a decoupling material under fluid loading. He used the locally reacting model and examined the plane wave reflection coefficient when the excitation is a plane acoustic wave propagating in the fluid under grazing incidence. Maidanik *et al.* (1968, 1974, 1984) studied the acoustic behavior of multilayered infinite panels immersed in fluid media. They defined two types of layers: the compliant layer, that can support transverse strains, modeled by the locally reacting model, and the blanket that has dynamical properties that are fluidlike, so that the propagation within the blanket can be specified by a density and a speed of sound. Sylwan (1987) investigated the *in vacuo* vibrational behavior of an infinite sandwich beam when the core is modeled by evenly distributed springs in the transverse direction (to take into account transverse strain) and in the longitudinal direction (to take into account shear strain). He compared both effects as a function of the frequency and obtained their relative contribution in the global loss factor of the structure. House (1991)

studied the transmission loss of an infinite panel covered by a compliant coating immersed in water. He found good agreement between experimental results in a 1/10th scale section of a ship hull and a simple theory where the hull and the coating are assumed to behave as a rigid wall with a given impedance. Laulagnet *et al.* (1991, 1994) used the locally reacting model to calculate the sound radiated from a thin, simply supported cylindrical shell covered by a compliant coating. This method was compared to a rigorous model that involves the three-dimensional theory of elasticity for the compliant layer motion, and it was shown that the locally reacting model appropriately describes the vibroacoustic behavior of thin shells covered by a decoupling layer. Recently Sandman *et al.* (1995) used the locally reacting model when the base structure is a rigid piston. In order to obtain a very simple model, only limiting cases were considered by these authors for the calculation of the radiation impedance of the rigid piston.

As seen in the literature review, the sound radiation under mechanical excitation of a finite elastic plate, covered by a decoupling coating, is not yet fully covered. It is the purpose of this article to propose a model where the base plate is a rectangular elastic plate and the coating is assumed to conform to the locally reacting model. In particular, an objective of the proposed model is to derive a global vibratory indicator appropriate to represent the acoustic efficiency of a decoupling treatment, independently of the substrate (base plate) dimensions. Such an indicator would be useful to predict the efficiency of a decoupling treatment on a large structure from experiment on small-scale test specimens.

The acoustic behavior of the elastic plate-compliant coating system inserted in an infinite baffle involves the evaluation of the surface acoustic pressure. The modal basis, used for the displacement of the simply supported plate, cannot be used for the surface pressure because on the baffle/plate boundary the displacement is zero, but not the acoustic pressure. Several authors (Ginsberg *et al.*, 1995, 1991) used the surface variational principle (SVP, Pierce, 1987) to evaluate the surface acoustic pressure. These authors define a distance r —from the boundary of the plate—beyond which the surface pressure is assumed to vanish. The surface acoustic pressure is then defined by a set of trigonometric functions that vanish at the distance r . In this paper, the surface acoustic pressure is expanded on a complete Fourier series that allows the nonzero condition of the surface acoustic pressure on the baffle to be satisfied. Furthermore, it avoids the need to define the distance r .

The theoretical formulation of the proposed model is developed in the next section. The locally reacting model and the basic equations that describe the vibroacoustic behavior of the plate-coating-water system are presented. The governing equations are solved to obtain global structural and acoustic indicators of the problem. A simple theory is derived from the proposed model to treat the case of a large decoupling (low stiffness decoupling material/high frequency). The results obtained by the proposed model are analyzed in the numerical results section where the vibration and acoustic isolation effect provided by a compliant layer are presented.

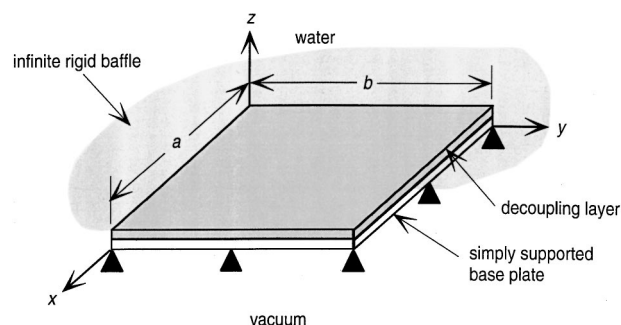


FIG. 1. Schematic representation of the baffled plate covered with a decoupling layer and immersed in water.

I. THEORETICAL FORMULATION

A. Statement of the problem

The structure under study is assumed to be a rectangular, baffled plate, simply supported on the four edges. The plate-baffle system separates a heavy fluid ($z > 0$) from vacuum ($z < 0$) (Fig. 1). The structure itself consists of a base flexural plate covered by a compliant layer allowing for thickness deformation. The position of the rigid baffle coincides with the outer surface of the decoupling layer. The base plate is assumed to conform to the Love–Kirchhoff theory that considers only bending deformation, while the coating layer is described with the locally reacting model. The purpose of the following model is to calculate the transverse vibration of the base plate, the transverse vibration at the outer surface of the coating, and the resulting sound radiation in the semi-infinite heavy fluid when the base plate is excited by an harmonic point force.

The output parameters are the mean-square velocity of the plate and of the outer surface of the decoupling layer, the radiation efficiency, and the radiated sound power in the heavy fluid.

B. Equation of motion of the structure

The thickness deformation of the decoupling layer implies that the base plate and the outer surface of the decoupling layer (the surface in contact with the fluid) have distinct transverse displacements (Fig. 2). Thus \tilde{w}_1 is defined to be the transverse displacement of the base plate and \tilde{w}_2 the transverse displacement of the outer surface of the decoupling layer.

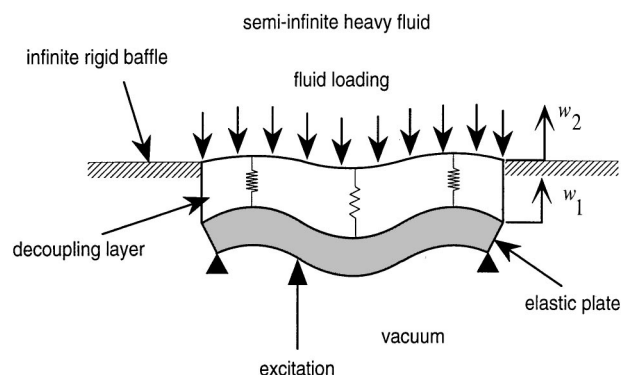


FIG. 2. Schematic representation of the locally reacting model.

pling layer. Throughout this document, tildes represent complex values expressed in the frequency domain.

The equation of motion of the base plate, expressed in the frequency domain, is

$$\tilde{D}\nabla^4\tilde{w}_1(Q) - \rho h \omega^2 \tilde{w}_1(Q) = \tilde{f}(Q) - \tilde{\sigma}(Q), \quad (1)$$

where ω is the angular frequency, Q is a point on the plate, \tilde{D} is the complex bending stiffness of the plate: $\tilde{D} = \tilde{E}h^3/12(1-\nu^2)$, ρ , h and ν are, respectively, the density, the thickness, and the Poisson's ratio of the plate, \tilde{E} is the complex Young's modulus, defined by $\tilde{E} = E(1+j\eta)$, E is the real part of the Young's modulus, and η the structural loss factor, ∇^4 is a partial differential operator: $\nabla^4 = (\partial^4/\partial x^4) + 2(\partial^4/\partial x^2\partial y^2) + \partial^4/\partial y^4$, $\tilde{f}(Q)$ is the external transverse force per unit area applied to the base plate and $\tilde{\sigma}(Q)$ is the normal stress in the z -direction applied by the coating on the base plate.

Assuming that the coating behaves like a massless stiffness, the surface acoustic pressure $\tilde{P}(Q)$ exerted by the fluid on the coating is the same as the normal stress $\tilde{\sigma}(Q)$ exerted by the coating on the structure. Thus, $\tilde{\sigma}(Q) = \tilde{P}(Q)$ in Eq. (1). In the following, the unknowns of this equation are the transverse displacement of the plate \tilde{w}_1 and the acoustic pressure \tilde{P} . According to previous work (Laulagnet *et al.*, 1994), the zero mass assumption for the coating is acceptable in situations where the mass per unit area of the coating is smaller than the mass per unit area of the substrate.

The simply supported boundary conditions assumes that the transverse displacement and the bending moment vanish on the plate boundaries,

$$\tilde{w}_1(x,y)=0 \quad \text{and} \quad \frac{\partial^2 \tilde{w}_1(x,y)}{\partial x^2}=0 \quad \text{for} \quad x=0 \quad \text{and} \quad x=a, \quad (2)$$

$$\tilde{w}_1(x,y)=0 \quad \text{and} \quad \frac{\partial^2 \tilde{w}_1(x,y)}{\partial y^2}=0 \quad \text{for} \quad y=0 \quad \text{and} \quad y=b, \quad (3)$$

where a and b are, respectively, the x -dimension and the y -dimension of the plate.

Considering a locally reacting decoupling material allows the following relation to be written for the motion of the decoupling layer under the action of the surface acoustic pressure $\tilde{P}(Q)$:

$$\tilde{P}(Q) = Z_c(\tilde{w}_2(Q) - \tilde{w}_1(Q)). \quad (4)$$

Equation (4) assumes that the thickness deformation of the decoupling layer at point Q depends only on the surface acoustic pressure at point Q . This equation ignores extended reaction of the layer, and thus conforms to a locally reacting assumption. Z_c is the impedance of the decoupling material, which can be written in terms of the elastic properties of the materials,

$$Z_c = \frac{\tilde{B}_c}{h_c}. \quad (5)$$

In Eq. (5), \tilde{B}_c is the complex modulus of the decoupling layer: $\tilde{B}_c = B_c(1+j\eta_c)$, and h_c is the thickness of this mate-

rial. Recent experiments (Hamm, 1996) have suggested that the *bulk* modulus is the best parameter to characterize the decoupling material when the locally reacting model is used, although a more thorough experimental validation using solid polymers and foams with known elastic properties still needs to be carried out. The decoupling material is thus characterized only by its bulk modulus and its thickness.

The acoustic pressure in the fluid is governed by the Helmholtz equation,

$$\Delta \tilde{P}(M) + k_0^2 \tilde{P}(M) = 0, \quad (6)$$

where M is a point in the fluid space, k_0 is the acoustic wave number defined by $k_0 = \omega/c_0$, and c_0 is the velocity of sound in the fluid.

The equation of continuity of the structural and acoustic normal accelerations on the outer surface of the decoupling material is

$$\frac{\partial \tilde{P}(Q)}{\partial z} = \rho_0 \omega^2 \tilde{w}_2(Q), \quad (7)$$

where ρ_0 is the density of the fluid.

The acoustic pressure on the surface of the coating is evaluated with the Rayleigh integral,

$$\tilde{P}(Q) = -\omega^2 \rho_0 \int \int_s \tilde{w}_2(M) G(M, Q) ds_M, \quad (8)$$

where s is the surface of the decoupling layer, points Q and M are on the surface of the decoupling layer, and the acoustic Green's function G is defined in the Appendix. The system formed by Eqs. (1), (4), and (8) has three unknowns: $\tilde{w}_1(Q)$, $\tilde{w}_2(Q)$, and $\tilde{P}(Q)$. Equation (4) allows the displacement $\tilde{w}_2(Q)$ to be expressed in terms of $\tilde{w}_1(Q)$ and $\tilde{P}(Q)$, so that the two remaining unknowns are $\tilde{w}_1(Q)$ and $\tilde{P}(Q)$. Using Eq. (4), Eq. (8) becomes

$$\tilde{P}(Q) = -\omega^2 \rho_0 \int \int_s \left(\tilde{w}_1(M) + \frac{\tilde{P}(M)}{Z_c} \right) G(m, Q) ds_M. \quad (9)$$

The displacement $\tilde{w}_1(Q)$ is expanded over the orthogonal *in vacuo* modes of the plate, while the acoustic pressure $\tilde{P}(Q)$ is expanded over a complete orthogonal Fourier series,

$$\tilde{w}_1(Q) = \sum_{m=1}^{\infty} \sum_{n=1}^{\infty} \tilde{a}_{mn} w_{mn}(Q), \quad (10a)$$

$$\tilde{P}(Q) = \sum_{p=-\infty}^{\infty} \sum_{q=-\infty}^{\infty} \tilde{b}_{pq} \psi_{pq}(Q), \quad (10b)$$

where $w_{mn}(Q) = \sin(m\pi x/a) \sin(n\pi y/b)$, and $\psi_{pq}(Q) = e^{j2\pi(px/a + qy/b)}$.

Using Eq. (10), Eqs. (1) and (9) lead to two linear systems of coupled equations written in terms of the unknowns of the problem, i.e., the modal displacements \tilde{a}_{mn} and the Fourier coefficients \tilde{b}_{pq} for the surface pressure,

$$\rho h \frac{ab}{4} \tilde{a}_{mn} (\omega_{mn}^2 (1 + j\eta) - \omega^2) = \tilde{f}_{mn} - \sum_{i=-\infty}^{\infty} \sum_{j=-\infty}^{\infty} b_{ij} S_{mnij}, \quad (11a)$$

$$ab \cdot \tilde{b}_{pq} = -\omega^2 \rho_0 \sum_{k=1}^{\infty} \sum_{l=1}^{\infty} \tilde{a}_{kl} H_{klpq} - \omega^2 \rho_0 \sum_{i=-\infty}^{\infty} \sum_{j=-\infty}^{\infty} \frac{\tilde{b}_{ij} L_{ijpq}}{Z_c}, \quad (11b)$$

where ω_{mn} are the *in vacuo* natural angular frequencies of the plate,

$$\omega_{mn} = \sqrt{\frac{D}{\rho h} \left[\left(\frac{m\pi}{a} \right)^2 + \left(\frac{n\pi}{b} \right)^2 \right]}, \quad (12)$$

and D is the real part of \tilde{D} .

The vector of the generalized forces \tilde{f}_{mn} in the case of a point force is

$$\tilde{f}_{mn} = F_0 \sin\left(\frac{m\pi x_0}{a}\right) \sin\left(\frac{n\pi y_0}{b}\right), \quad (13)$$

where F_0 is the magnitude of the force and x_0 and y_0 are the coordinates of the excitation point. In Eqs. (11a) and (11b), the three terms S_{mnij} , H_{klpq} , and L_{ijpq} are defined by the following integrals:

$$S_{mnij} = \int_s \int_s w_{mn}(Q) \psi_{ij}(Q) ds, \\ H_{klpq} = \int_s \int_s \int_s \int_s w_{kl}(Q) G(Q, M) \psi_{pq}^*(M) ds_Q ds_M, \quad (14)$$

$$L_{ijpq} = \int_s \int_s \int_s \int_s \psi_{ij}(Q) G(Q, M) \psi_{pq}^*(M) ds_Q ds_M.$$

The calculation of S_{mnij} is trivial. The integrals H_{klpq} and L_{ijpq} need to be calculated numerically; this calculation is done following the idea of Sandman (1975) and Nelisse (1996) that allows the quadruple integral to be analytically reduced into a double integral using an appropriate change of integration variables. The resulting double integral is then calculated with standard Gauss integration (see Appendix). The method to avoid the singularity problem of these integrals is also described in the Appendix.

The equation (11) can be written in matrix form,

$$\begin{bmatrix} [A] \\ [C] \end{bmatrix} \begin{bmatrix} [\tilde{a}_{mn}] \\ [\tilde{b}_{pq}] \end{bmatrix} = \begin{bmatrix} \tilde{f}_{mn} \\ 0 \end{bmatrix}, \quad (15)$$

where

$$[A] = \begin{bmatrix} \rho \frac{ab}{4} (\omega_{mn}^2 (1 + j\eta) - \omega^2) \end{bmatrix} \quad (\text{diagonal matrix}), \\ [B] = [S_{mnij}] \quad (\text{full matrix}), \\ [C] = [\omega^2 \rho_0 H_{klpq}] \quad (\text{full matrix}), \quad (16)$$

$$[D] = \left[\omega^2 \frac{\rho_0}{Z_c} L_{ijpq} \right] + [\dots ab \dots] \quad (\text{full matrix}).$$

The infinite series on m, n, p, q, i, j, k, l are now truncated such that m, n, k, l vary between 1 and N for the displacement and p, q, i, j vary between $-N_p$ and N_p for the acoustic pressure. The system can be written in the form

$$\{\tilde{a}_{mn}\} = [A]^{-1} (\{\tilde{f}_{mn}\} - [B] \{\tilde{b}_{pq}\}), \quad (17)$$

$$([D] - [C][A]^{-1}[B]) \{\tilde{b}_{pq}\} = [C][A]^{-1} \{\tilde{f}_{mn}\}. \quad (18)$$

To solve the problem, Eq. (18) has to be first solved for the magnitude of the surface acoustic pressures \tilde{b}_{pq} and then Eq. (17) can be used to calculate the modal displacements \tilde{a}_{mn} . Due to the diagonal structure of $[A]$, $[A]^{-1}$ is trivial.

C. Vibroacoustic indicators

Several vibroacoustic indicators can be deduced from the modal displacements \tilde{a}_{mn} and the magnitude of the surface pressure \tilde{b}_{pq} . The indicators used in this paper are the mean-square transverse velocity of the base plate, the mean-square transverse velocity of the outer surface of the decoupling layer, the radiated sound power into the fluid, and the radiation efficiency of the system.

The mean-square velocity of the base plate is defined by

$$\langle V \rangle^2 = \frac{\omega^2}{2s} \int_s \int_s \tilde{w}_1(Q) \tilde{w}_1^*(Q) ds_Q, \quad (19)$$

where $*$ designates the complex conjugate. Using the expansion of \tilde{w}_1 over the orthogonal *in vacuo* modes of the plate, one gets

$$\langle V_1 \rangle^2 = \frac{\omega^2}{8} \sum_{m=1}^N \sum_{n=1}^N |\tilde{a}_{mn}|^2. \quad (20)$$

Similarly, the mean-square velocity of the outer surface of the decoupling layer is

$$\langle V_2 \rangle^2 = \frac{\omega^2}{2s} \int_s \int_s \tilde{w}_2(Q) \tilde{w}_2^*(Q) ds_Q \\ = \frac{\omega^2}{8} \sum_{m=1}^N \sum_{n=1}^N |\tilde{a}_{mn}|^2 + \frac{\omega^2}{2|Z_c|^2} \sum_{p=-N_p}^{N_p} \sum_{q=-N_p}^{N_p} |\tilde{b}_{pq}|^2 \\ + \frac{\omega^2}{ab} \text{Re} \left[\sum_{k=1}^N \sum_{l=1}^N \sum_{i=-N_p}^{N_p} \sum_{j=-N_p}^{N_p} \frac{a_{kl}^* b_{ij} S_{kloj}}{Z_c} \right], \quad (21)$$

where use has been made of the orthogonal properties of the $\psi_{pq}(Q)$, and Re denotes the real part.

The radiated sound power in the fluid is the integration of the active acoustic intensity normal to the decoupling layer,

$$W = \frac{1}{2} \text{Re} \left[\int_s \int_s \tilde{P}(Q) (-j\omega) \tilde{w}_2^*(Q) ds_Q \right], \quad (22)$$

which gives

$$W = \frac{1}{2} \text{Re} \left[(-j\omega) \left[ab \sum_{p=-N_p}^{N_p} \sum_{q=-N_p}^{N_p} \frac{|b_{pq}|^2}{Z_c^*} + \sum_{m=1}^N \sum_{n=1}^N \sum_{p=-N_p}^{N_p} \sum_{q=-N_p}^{N_p} a_{mn}^* S_{mnpq} b_{pq} \right] \right]. \quad (23)$$

Finally, the radiation efficiency in the heavy fluid is defined by

$$\sigma = \frac{W}{\rho_0 c_0 s \langle V_2 \rangle^2}. \quad (24)$$

In the following, a simplified version of the above formulation is derived for the case of a low stiffness decoupling material.

D. Case of large decoupling

The motion of the base plate and the decoupling layer is largely decouple when the decoupling layer has a low stiffness (Z_c small). In this case the equation of the problem [Eqs. (11a), (11b)] can be rewritten in a simplified form. Two assumptions may be done in the case of high decoupling.

First, if the motion of the base plate and the decoupling layer are uncoupled, the fluid loading exerted on the base plate can be neglected. Thus, the equation of motion of the base plate is

$$\rho h \frac{ab}{4} \tilde{a}_{mn} (\omega_{mn}^2 (1 + j\eta) - \omega^2) = \tilde{f}_{mn}. \quad (25)$$

It is the *in vacuo* equation of motion of the base plate which can be readily solved for the modal displacement \tilde{a}_{mn} ,

$$\tilde{a}_{mn} = \frac{4\tilde{f}_{mn}}{\rho h ab (\omega_{mn}^2 (1 + j\eta) - \omega^2)}. \quad (26)$$

Let us now consider the transverse displacement of the outer surface of the decoupling layer $\tilde{w}_2(Q)$, and expand this displacement over a complete Fourier series,

$$\tilde{w}_2(Q) = \sum_{p=-\infty}^{\infty} \sum_{q=-\infty}^{\infty} \tilde{c}_{pq} \psi_{pq}(Q), \quad (27)$$

where ψ_{pq} is defined in Eq. (10b). Using Eq. (4), the Rayleigh integral [Eq. (8)] can be written

$$w_2(Q) - w_1(Q) = -\omega^2 \frac{\rho_0}{Z_c} \int \int_s \tilde{w}_2(M) G(M, Q) ds_M. \quad (28)$$

The assumption of a large decoupling implies that the motion of the outer surface of the decoupling layer can be neglected in comparison to the motion of the base plate,

$$w_2(Q) \ll w_1(Q). \quad (29)$$

Thus, Eq. (28) simplifies into

$$w_1(Q) = \omega^2 \frac{\rho_0}{Z_c} \int \int_s \tilde{w}_2(M) G(M, Q) ds_M. \quad (30)$$

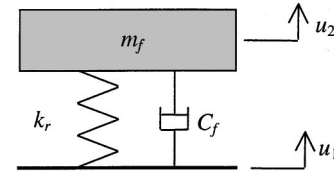


FIG. 3. Schematic representation of the lumped parameter system.

Integrating this equation over the surface of the plate and using orthogonal properties in *in vacuo* modes of the plate, leads to

$$\frac{ab}{4} a_{mn} = \frac{\omega^2 \rho_0}{Z_c} \sum_{r=-\infty}^{\infty} \sum_{s=-\infty}^{\infty} \tilde{c}_{rs} H_{mnrs}. \quad (31)$$

Equation (30) allows the unknown coefficients of the transverse displacement of the outer surface of the decoupling layer \tilde{c}_{rs} to be calculated,

$$\{\tilde{c}_{rs}\} = \frac{ab Z_c}{4 \omega^2 \rho_0} [H_{mnrs}]^{-1} \{\tilde{a}_{mn}\}. \quad (32)$$

This approximate equation relates the motion of the base plate and the motion of the surface of the decoupling layer, and hence reflects the amount of decoupling provided by the layer. The mean-square velocity of the outer surface of the decoupling layer may then be expressed as

$$\langle V_2 \rangle^2 = \frac{\omega^2}{2} \sum_{r=-N_p}^{N_p} \sum_{s=-N_p}^{N_p} |\tilde{c}_{rs}|^2. \quad (33)$$

Equation (32) shows that the Fourier coefficients for the displacement of the outer surface of the decoupling layer are proportional to the impedance of the decoupling material Z_c . This amount of the decoupling provided by the coating is thus simply related to the coating stiffness. This result is similar to classical vibration isolation using compliant mount. On the other hand, the frequency dependence of the decoupling cannot be simply drawn from Eq. (32) since the matrix H_{mnrs} has a complex dependence on the frequency.

E. Lumped parameter system

This section presents an alternative and simplified model of a base structure covered by a decoupling layer, by instead considering a lumped parameter system (Fig. 3). This approach is closely related to a previous work by Sandman *et al.* (1995). In this case, the substrate plate reduces to a rigid piston, and the decoupling coating still behaves as a massless stiffness. For such a system the displacement u_2 of the outer surface of the coating is expressed as a function of the displacement of the base u_1 by the following equation:

$$u_2 = \frac{k_r}{k_r - \omega^2 (m_f + jC_f)} u_1. \quad (34)$$

In this analogy, u_1 represents the displacement of the base plate, u_2 the displacement of the outer surface of the decoupling layer, k_r the stiffness of the decoupling layer, m_f represents the added mass due to the fluid, and C_f the radiation damping due to the fluid. In high frequency, when $\omega^2 m_f \gg k_r$, the displacement u_2 is proportional to k_r , as in Eq.

TABLE I. Characteristics of the steel base plate used in the numerical simulations.

x-dimension (m) a	y-dimension (m) b	Thickness (mm) h	Density (kg/m ³) ρ	Young's modulus (Pa) E	Loss factor η	Poisson's ratio ν
0.6	0.6	9	7850	2.1×10^{11}	0.005	0.3

(32). In the case of a rigid circular piston, the fluid loading term $m_f + jC_f$ is given by

$$m_f + jC_f = \pi(r_p)^2 \rho_0 c_0 \frac{Z_f}{j\omega}, \quad (35)$$

where r_p is the piston radius, the radiation impedance Z_f is given by

$$Z_f = R + jX,$$

$$\text{with } R = 1 - \frac{2J_1(2k_0 r_p)}{2k_0 r_p} \text{ and } X = \frac{2H_1(2k_0 r_p)}{2k_0 r_p}. \quad (36)$$

J_1 is the first-order Bessel function defined by

$$J_1(x) = \frac{(x/2)}{(1!)^2} - 2 \frac{(x/2)^3}{(2!)^2} + 3 \frac{(x/2)^5}{(3!)^2} - \dots \quad (37)$$

H_1 is the first-order Struve function,

$$H_1(x) = \frac{2}{\pi} \left(\frac{x^2}{1^2 \cdot 3} - \frac{x^4}{1^2 \cdot 3^2 \cdot 5} + \frac{x^6}{1^2 \cdot 3^2 \cdot 5} - \dots \right). \quad (38)$$

When $2k_0 r_p \gg 1$ Pierce (1981, p. 234, Eqs. 5-4.12a and 5-4.12b) uses an asymptotic expression,

$$J_1(x) \approx \sqrt{\frac{2}{\pi x}} \cos\left(x - \frac{3\pi}{4}\right), \quad (39)$$

$$H_1(x) \approx \frac{2}{\pi} + \sqrt{\frac{2}{\pi x}} \sin\left(x - \frac{3\pi}{4}\right). \quad (40)$$

Results given by this simple one degree of freedom system are compared to the proposed model of this paper in Sec. II.

II. NUMERICAL RESULTS

This section presents numerical results of the vibration reduction or sound attenuation that can be achieved by adding a compliant layer or an elastic base plate immersed in water. The vibroacoustic indicators calculated are the mean-square velocity of the base plate, the mean-square velocity of the outer surface of the decoupling layer, the radiation efficiency of the structure, and the radiated sound power in water. In order to characterize the decoupling efficiency, the ratio of the mean-square velocity of the base plate to the mean-square velocity of the outer surface of the coating is also presented. One underlying motivation of this numerical study is to derive a vibratory indicator representative of the acoustic performance of the decoupling treatment, and independent of the base plate dimensions.

The characteristics of the steel base plate used for all the numerical simulations are described in Table I. The excitation is a transverse point force applied at $x = y = 0.06$ m from a corner of the plate. In the locally reacting model the de-

coupling material is characterized by its impedance $Z_c = B_c(1 + j\eta_c)/h_c$; thus, increasing the modulus B_c is equivalent to decreasing the thickness h_c . In this study the thickness of the decoupling layer is fixed to 1 cm, whereas its modulus varies from 10^8 to 10^5 Pa. The loss factor of the decoupling layer is set to zero in order to study the decoupling effect without damping. In all cases the structure is immersed in water (density $\rho_0 = 1000$ kg/m³, sound speed $c_0 = 1460$ m/s). The expansion order used for this displacement and the acoustic pressure is $N = N_p = 15$, and the number of Gauss points used for the numerical calculation of H_{klpq} and L_{ijpq} is $N_g = 20$.

A. Vibration isolation results

Figures 4 and 5 show the mean-square velocity of the base plate $\langle V_1 \rangle^2$ [Eq. (20)] and the mean-square velocity of the outer surface of the decoupling layer $\langle V_2 \rangle^2$ [Eq. (21)] when $B_c = 10^8$ and $B_c = 10^6$ Pa. When $B_c = 10^8$ Pa, the velocity of the layer outer surface is not significantly reduced as compared to the velocity of the base plate; the decoupling provided by the layer is low in this case. When $B_c = 10^6$ Pa, the decoupling is much larger. It is seen from Figs. 4 and 5 that for a given value of B_c the frequency behavior of $\langle V_1 \rangle^2$ and $\langle V_2 \rangle^2$ shows the same peaks, corresponding to the resonance frequencies of the fluid-loaded base plate with the decoupling layer system. Moreover, when B_c is decreased, these resonance frequencies are shifted to higher frequencies, corresponding in the limiting case of a very compliant coating to the natural frequencies of the *in vacuo* base plate. In such a case, the base plate is isolated from the fluid loading by the compliant coating, and thus vibrates like an *in vacuo* plate. The approximate equation (33) would also predict that

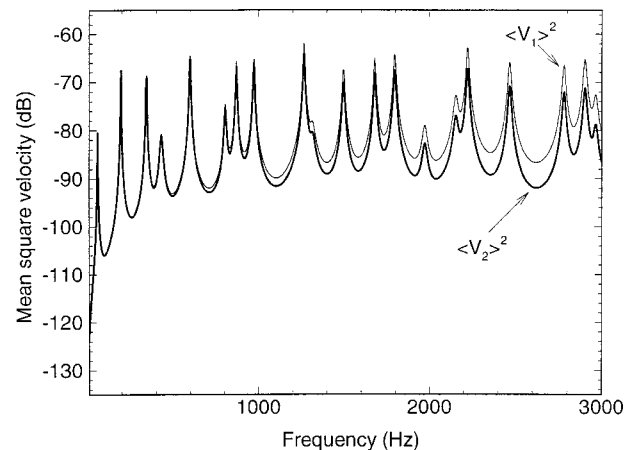


FIG. 4. Mean-square velocity of the base plate $\langle V_1 \rangle^2$ and mean-square velocity of the surface of the decoupling layer $\langle V_2 \rangle^2$ when $B_c = 10^8$ Pa, $\eta_c = 0$, $h_c = 1$ cm.

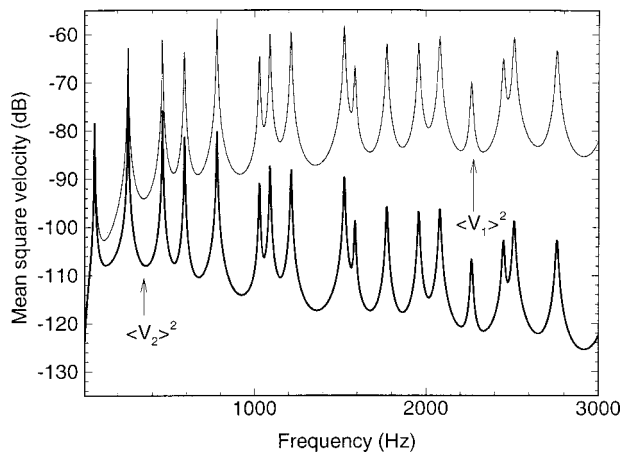


FIG. 5. Mean-square velocity of the base plate $\langle V_1 \rangle^2$ and mean-square velocity of the surface of the decoupling layer $\langle V_2 \rangle^2$ when $B_c = 10^6$ Pa, $\eta_c = 0$, $h_c = 1$ cm.

decreasing the layer modulus by a factor 10 in this case still decreases $\langle V_2 \rangle^2$ by an amount of 20 dB [$(Z_c)^2$ dependence of $\langle V_2 \rangle^2$, Eq. (33)].

B. Sound radiation results

Figure 6 shows the radiation efficiency for the base plate without coating and with coating ($B_c = 10^6$ Pa). This result shows that the radiation efficiency of the structure is globally unchanged by the addition of a compliant layer. Figure 7 shows the corresponding sound power radiated into water from the base plate without coating, and from the plate with coating ($B_c = 10^6$ Pa). The acoustic attenuation provided by the coating is commensurate with the vibration attenuation provided by the coating, since the radiation efficiency is essentially unchanged by the coating.

Figure 7 shows that the “large decoupling” approximation (Sec. ID) satisfactorily represents the radiated sound power of the structure. This validates the *in vacuo* approximation used to calculate the motion of the base plate as well as the assumption that neglects the motion of the outer surface of the coating in comparison to the motion of the base plate.

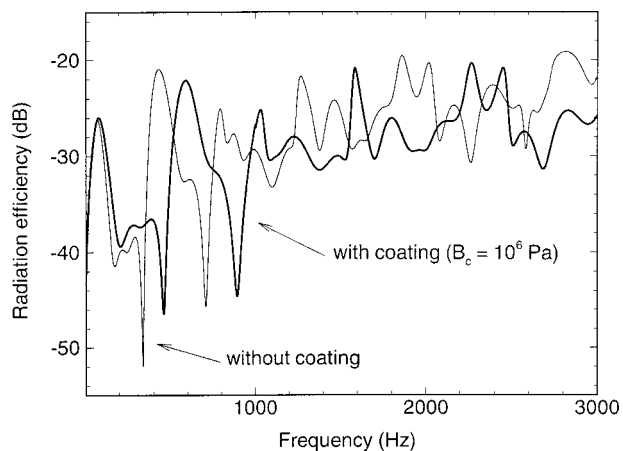


FIG. 6. Radiation efficiency of the structure without and with a decoupling layer when $B_c = 10^6$ Pa, $\eta_c = 0$, $h_c = 1$ cm.

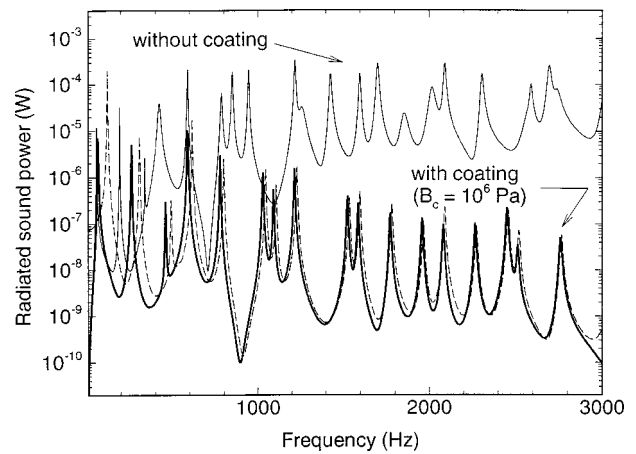


FIG. 7. Radiated sound power in water without and with a decoupling layer ($B_c = 10^6$ Pa, $\eta_c = 0$, $h_c = 1$ cm. Comparison between the proposed model (—) and the high decoupling approximation (---).

C. Lumped parameter system

Results given by the lumped parameter system presented in Sec. IE are compared to the proposed model of this paper in this section. The coating material has the following properties: $B_c = 10^6$ Pa, $h_c = 10$ mm, $\eta_c = 0$. The equivalent spring stiffness is given by: $k_r = \pi(r_p)^2 B_c / h_c$, the piston area is chosen to be equal to the late area: $ab = \pi(r_p)^2$. In the numerical results, eight terms were used in expansion of the Bessel function [Eq. (37)] and of the Struve function [Eq. (38)]. This allows an accurate calculation of this piston radiation impedance up to 2 kHz in this case; above this frequency more terms should be used in the series expansions, or the high-frequency approximation given by Pierce should be used. The velocity ratio is the indicator chosen for the comparison between the models. In Fig. 8, the indicator $10 \log_{10}(|u_1/u_2|^2)$ defined for the lumped parameter system is compared to the mean-square velocity ratio $10 \log_{10}[\langle V_1 \rangle^2 / \langle V_2 \rangle^2]$ defined for the plate-coating system. The two curves show the same tendencies; however, the lumped parameter system overestimates the elastic model by around 10 dB at low frequencies. This deviation is attributed to the difference in the type of motion (rigid body versus elastic) in both models. Above 2 kHz the lumped model

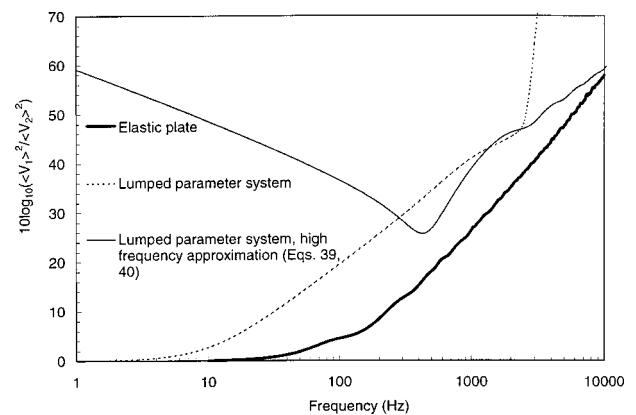


FIG. 8. Comparison between the elastic plate + locally reacting model proposed in this paper and a simple one degree of freedom system where the radiation impedance is modeled using a circular rigid piston.

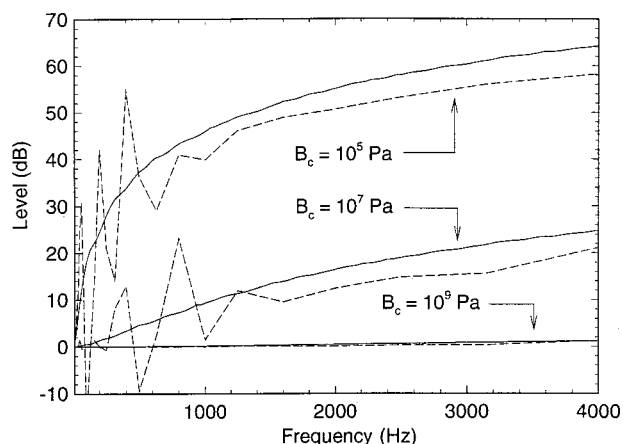


FIG. 9. Comparison between the ratio of the mean-square velocity of the base plate to the mean-square velocity of surface of the decoupling layer (—) and the acoustic insertion loss (---), where $B_c = 10^9$, $B_c = 10^7$, $B_c = 10^5$ Pa ($\eta_c = 0$, $h_c = 1$ cm in all cases).

diverges because of the insufficient number of terms used to evaluate the piston radiation impedance. The high-frequency approximation of the lumped model (using a high-frequency asymptotic expression of the piston radiation impedance) was plotted on the same graph; it satisfactorily approximates the elastic plate model at high frequencies but diverges at low frequencies. This example shows that a very simple system is able to approximate tendencies in the behavior of the plate-coating system, but it does not allow accurate prediction of the vibration or noise attenuation of an elastic plate covered by a compliant coating and immersed in water.

D. Decoupling indicator

In Fig. 9, the ratio of the mean-square velocity of the base plate to the mean-square velocity of the decoupling layer, $10 \log_{10}[\langle V_1 \rangle^2 / \langle V_2 \rangle^2]$, is compared to the acoustic power insertion loss, $10 \log_{10}(W_0/W)$, where W_0 is the radiated acoustic power of the structure without coating and W is the radiated acoustic power with coating. These results were obtained by solving Eqs. (17) and (18), without large decoupling approximation. From Fig. 9, it is seen that the mean-square velocity ratio is a nonmodal parameter that shows a smooth increase of the decoupling efficiency as a function of the frequency.

The acoustic power insertion loss characterizes the noise attenuation provided by the treatment. As seen in Fig. 7 there is a shift of the resonance frequencies between the response of the base plate alone and the response of the plate covered by a coating. This shift yields many peaks and dips when the acoustic power is plotted in the narrow-frequency band; the acoustic power insertion loss was thus plotted in 1/3-octave bands in Fig. 9. The differences at low frequencies between the mean-square velocity ratio and the insertion loss are due to the aforementioned frequency shifts. Both indicators show approximately the same frequency dependence but the velocity ratio is about 3 to 5 dB lower than the insertion loss. If the radiation efficiency is assumed to be unchanged when the coating is added (see Fig. 6), one can explain this 3 to 5 dB as the difference between the velocity of the base plate with and without coating.

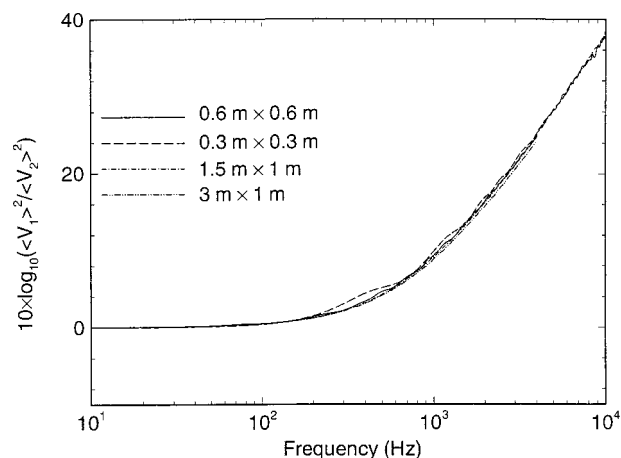


FIG. 10. Ratio of the mean-square velocity of the base plate to the mean-square velocity of the surface of the decoupling layer when the plate dimensions are 0.6×0.6 , 0.3×0.3 , 1.5×1 , 3×1 m ($B_c = 10^7$ Pa, $\eta_c = 0$, $h_c = 1$ cm).

E. Variation of the base structure dimensions

The effect of changing the base plate dimensions on the vibration isolation is investigated in this section. This point addresses the question of characterizing the vibroacoustic behavior of a decoupling layer on a test structure of small size, and transferring the result on a full scale, large structure. Figure 10 compares the mean-square velocity ratio for four different dimensions of the base plate (0.6×0.6 , 0.3×0.3 , 1.5×1 , and 3×1 m). For all cases the coating characteristics are: $B_c = 10^7$ Pa, $h_c = 1$ cm, $\eta_c = 0$. Clearly, this ratio is independent of the substrate size; similar results were obtained for all plate dimensions.

III. CONCLUSION

The vibroacoustic behavior of an elastic simply supported, rectangular baffled plate covered by a locally reacting decoupling layer supporting thickness deformation was presented in this paper. The model simulates the vibration and acoustic response of the system immersed in water, and subjected to a point force disturbance. A simplified version of the theory was derived in the limiting case of a large decoupling (low impedance of the layer/high frequency). On the other hand, the proposed model shows that the radiation efficiency of the structure is marginally affected by the addition of a locally reacting decoupling layer; the vibration attenuation provided by the decoupling layer is thus representative of the acoustic attenuation. Finally, the ratio of the mean-square velocity of the substrate to the mean-square velocity of the layer's surface is a smooth, nonmodal indicator approximately independent of the system x - and y -dimensions, and representative of the global acoustic performance of the decoupling treatment.

ACKNOWLEDGMENT

The work at the University of Sherbrooke was financed by the Defense Research Establishment Atlantic, Contract No. 95-00760-(107).

APPENDIX: CALCULATION OF MATRICES H_{klpq} and L_{ijpq}

Matrices H_{klpq} and L_{ijpq} are written,

$$H_{klpq} = \int_0^a \int_0^b \int_0^a \int_0^b w_k(x) w_l(y) G(x, y, x_0, y_0) \times \psi_p^*(x_0) \psi_q'(y_0) dx dy dx_0 dy_0, \quad (\text{A1})$$

$$L_{ijpq} = \int_0^a \int_0^b \int_0^a \int_0^b \psi_i(x) \psi_j(y) G(x, y, x_0, y_0) \times \psi_p^*(x_0) \psi_q'(y_0) dx dy dx_0 dy_0, \quad (\text{A2})$$

where $*$ denotes the complex conjugate, $\psi_p(x) = e^{j2\pi(px/a)}$, $\psi_q(y) = e^{j2\pi(qy/b)}$, $w_k(x) = \sin(k\pi x/a)$, $w_l(y) = \sin(l\pi y/b)$, $G = e^{-jk_0 R/2\pi R}$ and $R = \sqrt{(x-x_0)^2 + (y-y_0)^2}$.

The following change of variables is used:

$$\begin{cases} u = \frac{x-x_0}{a} \\ v = \frac{y-y_0}{b} \end{cases} \text{ and } \begin{cases} u' = \frac{x-x_0}{a} \\ v' = \frac{y-y_0}{b} \end{cases} \quad (\text{A3})$$

Using this change of variables, H_{klpq} and L_{ijpq} become

$$H_{klpq} = a^2 b^2 \int_{u'} \int_{v'} \left[\int_{-1}^0 \int_{-u}^1 w_k(u, v) w_l(u', v') \times G(u, u') \psi_p^*(v) \psi_q^*(v') dv du + \int_0^1 \int_0^{1-u} w_k(u, v) w_l(u', v') G(u, u') \times \psi_p^*(v) \psi_q^*(v') dv du \right] du' dv', \quad (\text{A4})$$

$$L_{ijpq} = a^2 b^2 \int_{u'} \int_{v'} \left[\int_{-1}^0 \int_{-u}^1 \psi_i(u, v) \psi_j(u', v') \times G(u, u') \psi_p^*(v) \psi_q^*(v') dv du + \int_0^1 \int_0^{1-u} \psi_i(u, v) \psi_j(u', v') G(u, u') \times \psi_p^*(v) \psi_q^*(v') dv du \right] du' dv', \quad (\text{A5})$$

In Eqs. (A4) and (A5), the limits of the integration of u' and v' are similar to the limits of integration on u and v , respectively.

The change of variable $u = -\tilde{u}$ is used in the integrals over u in (A4) and (A5). The first term of (A4) then leads to

$$\int_0^1 \int_{-u}^1 w_k(-u, v) w_l(u', v') G(u, u') \psi_p^*(v) \psi_q^*(v') dv du, \quad (\text{A6})$$

and the first term of (A5) leads to

$$\int_0^1 \int_{-u}^1 \psi_i(-u, v) \psi_j(u', v') G(u, u') \psi_p^*(v) \psi_q^*(v') dv du, \quad (\text{A7})$$

One can then analytically calculate the integrals over v in Eqs. (A4) and (A5). Repeating the same process for u' and v' leads to

$$H_{klpq} = a^2 b^2 \int_0^1 \int_0^1 F_{kp}^B(u) F_{lq}^B(u') G(u, u') du du', \quad (\text{A8})$$

$$L_{ijpq} = a^2 b^2 \int_0^1 \int_0^1 F_{ip}^C(u) F_{jq}^C(u') G(u, u') du du', \quad (\text{A9})$$

where $G(u, u') = (e^{-jk_0 R/2\pi R})$ and $R = \sqrt{a^2 u^2 + b^2 (u')^2}$.

The function $F_{kp}^B(u)$ is defined by

$$F_{kp}^B(u) = \int_{-u}^1 w_k(-u, v) \psi_p^*(v) dv + \int_0^{1-u} w_k(u, v) \psi_p^*(v) dv. \quad (\text{A10})$$

If k is even,

$$F_{kp}^B(u) = \frac{2j}{\pi(k^2 - 2p^2)} [2p \sin(\pi k u) - k \sin(2\pi p u)]. \quad (\text{A11})$$

If k is odd,

$$F_{kp}^B(u) = \frac{2k}{\pi(k^2 - 4p^2)} [\cos(2\pi p u) + \cos(\pi k u)]. \quad (\text{A12})$$

If $k = 2p$,

$$F_{kp}^B(u) = j(u-1) \cos(\pi k u) - \frac{j}{\pi(k+2p)} [\sin(\pi k u) + \sin(2\pi p u)]. \quad (\text{A13})$$

If $k = -2p$,

$$F_{kp}^B(u) = j(u-1) \cos(\pi k u) + \frac{j}{\pi(2p-k)} [\sin(\pi p u) - \sin(\pi k u)]. \quad (\text{A14})$$

The function $F_{ip}^C(u)$ is defined by

$$F_{ip}^C(u) = \int_{-u}^1 \psi_i(-u, v) \psi_p^*(v) dv + \int_0^{1-u} \psi_i(u, v) \psi_p^*(v) dv. \quad (\text{A15})$$

If $i \neq p$,

$$F_{ip}^C(u) = -\frac{2}{\pi(i-p)} \cos(\pi(i+p)u) \sin(\pi(i-p)u). \quad (\text{A16})$$

If $i = p$,

$$F_{ip}^C(u) = 2(1-u) \cos(2\pi i u). \quad (\text{A17})$$

The function $F_{ip}^C(u)$ satisfies the symmetry relation $F_{ip}^C(u) = F_{pi}^C(u)$, such that: $L_{ijpq} = L_{pjqi} = L_{iqpj} = L_{pqij}$. The integrals (A8) and (A9) are finally solved using Gaussian quadrature formulas,

$$H_{klpq} = a^2 b^2 \sum_{r=0}^{N_g} \sum_{s=0}^{N_g} p_r p_s F_{kp}^B(u_r) F_{lq}^B(u'_s) G(u_r, u'_s) \quad (A18)$$

$$\forall (u_r, u'_s) \in]0, 1[,$$

$$L_{ijpq} = a^2 b^2 \sum_{r=0}^{N_g} \sum_{s=0}^{N_g} p_r p_s F_{ip}^C(u_r) F_{jq}^C(u'_s) G(u_r, u'_s) \quad (A19)$$

$$\forall (u_r, u'_s) \in]0, 1[,$$

where p_r and p_s are the weights associated with the Gauss points u_r and u'_s . N_g is the number of Gauss points for the integration. The singularity $u_r = u'_s = 0$ is automatically removed because the Gauss points are in the interval $]0, 1[$.

- Crighton, D. G. (1979). "Aspects of the reflexion and free wave properties of a composite panel under fluid loading," *J. Sound Vib.* **64**, 467–474.
- Garrison, M. R., Miles, R. N., Sun, J. Q., and Bao, W. (1994). "Random response of a plate partially covered by a constrained layer damper," *J. Sound Vib.* **172**, 231–245.
- Ginsberg, J. H., and McDaniel, J. G. (1991). "An acoustic variational principle and component modes synthesis applied to the analysis of acoustic radiation from concentrically stiffened plate," *Trans. ASME, J. Vib. Acoust.* **113**, 401–408.
- Ginsberg, J. H., Cunefare, K., and Pham, H. (1995). "A spectral description of inertial effects in fluid-loaded plates," *Trans. ASME, J. Vib. Acoust.* **117**, 206–212.
- Hamm, C. A. (1996). "Evaluation of noise reduction coatings: Experimental design," DREA report CR/96/417.
- House, J. R. C. (1991). "Considerations for underwater decoupling treatment," *Proc. I.O.A.* **13**, 166–173.

- Keltie, R. F. (1997). "Signal response of elastically coated plates," *J. Acoust. Soc. Am.* **103**, 1855–1863.
- Ko, S. H. (1997). "Reduction of structure-borne noise using an air-voided elastomer," *J. Acoust. Soc. Am.* **101**, 3306–3312.
- Laulagnet, B., and Guyader, J. L. (1991). "Sound radiation from finite cylindrical shell covered with a compliant layer," *Trans. ASME, J. Vib. Acoust.* **113**, 267–272.
- Laulagnet, B., and Guyader, J. L. (1994). "Sound radiation from finite cylindrical coated shells, by means of asymptotic expansion of three-dimensional equations for coating," *J. Acoust. Soc. Am.* **96**, 277–286.
- Maidanik, G., and Reader, W. T. (1968). "Filtering action of a blanket dome," *J. Acoust. Soc. Am.* **44**, 497–502.
- Maidanik, G., and Tucker, J. (1974). "Acoustic properties of coated panels immersed in fluid media," *J. Sound Vib.* **34**, 519–550.
- Maidanik, G. (1984). "Modification caused by compliant layers and blankets in the pressure field induced on a boundary," *Trans. ASME J. Vib., Acoust., Stress, Reliab. Des.* **106**, 369–375.
- Nelisse, H., Beslin, O., and Nicholas, J. (1996). "Fluid-structure coupling for an unbaffled elastic panel immersed in a diffuse field," *J. Sound Vib.* **198**, 485–506.
- Noor, A. K., and Burton, W. S. (1989). "Assessment of shear deformation theories for multilayered composite plates," *Appl. Mech. Rev.* **42**, 1–13.
- Pierce, A. D. (1981). *Acoustics: An Introduction to its Physical Principles and Applications* (McGraw-Hill, New York).
- Pierce, A. D. (1987). "Stationary variational expressions for the radiated and scattered acoustic power and related quantities," *IEEE J. Ocean Eng.* **12**, 404–411.
- Sandman, B. E. (1975). "Motion of a three-layered elastic-viscoelastic plate under fluid loading," *J. Acoust. Soc. Am.* **57**, 1097–1107.
- Sandman, B. E., and Boisvert, J. E. (1995). "Simplified structural acoustic characterization of external compliant coatings on submerged surfaces," NUWC Division Newport Technical Digest, pp. 65–71.
- Sylwan, O. (1987). "Shear and compression damping effects of constrained layered beams," *J. Sound Vib.* **118**, 35–45.

On reconstruction of acoustic pressure fields using the Helmholtz equation least squares method

Sean F. Wu^{a)}

Department of Mechanical Engineering, Wayne State University, Detroit, Michigan 48202

(Received 7 June 1999; accepted for publication 11 January 2000)

This paper presents analyses and implementation of the reconstruction of acoustic pressure fields radiated from a general, three-dimensional complex vibrating structure using the Helmholtz equation least-squares (HELs) method. The structure under consideration emulates a full-size four-cylinder engine. To simulate sound radiation from a vibrating structure, harmonic excitations are assumed to act on arbitrarily selected surfaces. The resulting vibration responses are solved by the commercial FEM (finite element method) software I-DEAS. Once the normal component of the surface velocity distribution is determined, the surface acoustic pressures are calculated using standard boundary element method (BEM) codes. The radiated acoustic pressures over several planar surfaces at certain distances from the source are calculated by the Helmholtz integral formulation. These field pressures are taken as the input to the HELs formulation to reconstruct acoustic pressures on the entire source surface, as well as in the field. The reconstructed acoustic pressures thus obtained are then compared with benchmark values. Numerical results demonstrate that good agreements can be obtained with relatively few expansion functions. The HELs method is shown to be very effective in the low-to-mid frequency regime, and can potentially become a powerful noise diagnostic tool. © 2000 Acoustical Society of America.
[S0001-4966(00)03604-3]

PACS numbers: 43.40.Rj, 43.20.Rz, 43.50.Yw [CBB]

INTRODUCTION

It has been shown¹⁻³ that the radiated acoustic pressure fields in both exterior and interior regions can be reconstructed by using the Helmholtz equation least-squares (HELs) method. This method utilizes an expansion of the spheroidal functions that satisfy the Helmholtz equation to represent the radiated acoustic pressure field. Such an expansion is uniformly convergent because the basis functions consist of a uniformly convergent series of Legendre functions. The coefficients associated with these basis functions are determined by requiring the assumed-form solution to satisfy the pressure boundary condition at the measurement points. The errors incurred in this process are minimized by the least-squares method. The solutions thus obtained are unique. Further, the number of measurements is determined by the number of expansion terms, which is small when an appropriate coordinate system is selected for the particular source geometry under consideration. Consequently, the numerical computation efficiency is high.

The HELs method has recently been validated experimentally on reconstruction of the radiated acoustic pressure fields from a bowling ball and simplified, full-size vehicle front-end buck subject to random excitations.^{4,5} In both cases, the acoustic pressure distributions and spectra on the source surfaces are reconstructed and compared with those measured at the same locations. Satisfactory agreements are obtained in all cases.

However, the source geometry in the previous studies¹⁻⁵ are simple and the surfaces are smooth, containing no sharp

edges and corners. These simple sources are of great importance in the initial stage of development, because analytic solutions or accurate measurements can be obtained so that the validity of the HELs method can be checked rigorously. Once this is done, it would be necessary to extend the HELs method to general three-dimensional structures, since in engineering applications most vibrating structures are of arbitrary and complex geometry.

The objective of the present investigation is to examine the effectiveness and robustness of the HELs method on reconstructing the radiated acoustic pressure fields from a general three-dimensional complex vibrating structure. In particular, the structure under consideration contains sharp edges, corners, and abrupt changes in surface contour. The acoustic near field around these irregularities can be extremely complex because of local acoustic diffraction, scattering, and reflection effects.

The test object and environment simulate those of an automotive engine running in a free field. In this paper numerical simulations are employed, namely, harmonic excitations on the engine block are assumed and the vibration responses of the engine block are calculated using standard FEM codes. Once this is done, the surface acoustic pressures are determined by the BEM codes. The field acoustic pressures are calculated by the Helmholtz integral formulation and taken as the input to the HELs formulation to reconstruct the acoustic pressure distributions over the engine block surface as well as in the field. The reconstructed acoustic pressures thus obtained are then compared with the benchmark results.

Section I of this paper gives the fundamental HELs formulations. The advantage and limitation of the HELs

^{a)}Electronic mail: swu@mel.eng.wayne.edu

method and its implementation are discussed in Secs. II and III, respectively. Sections IV and V describe the test object, setup, and reconstruction scheme, respectively. Comparisons of the reconstructed acoustic pressure fields with benchmark results are demonstrated in Sec. VI. Conclusions are drawn in Sec. VII.

I. THE HELS METHOD

The HELS method assumes that the radiated acoustic pressures can be represented by an expansion of the spheroidal functions that satisfy the Helmholtz equation

$$\hat{p}(\mathbf{x}) = \rho c \sum_{n=0}^{\infty} \sum_{l=-n}^n \tilde{C}_{n,l} \tilde{\psi}_{n,l}(\mathbf{x}), \quad (1)$$

where \hat{p} represents the complex amplitude of the acoustic pressure at any point \mathbf{x} in the field, ρ and c are the density and speed of sound of the fluid medium, respectively, and $\tilde{\psi}_{n,l}(\mathbf{x})$ are the basis functions that can be generated by using Gram–Schmidt orthonormalization⁶ with respect to the particular solutions to the Helmholtz equation $\psi_{n,l}$

$$\langle \psi_{n,l}(\mathbf{x}_B), \tilde{\psi}_{n,l}(\mathbf{x}_B) \rangle = \int_{\partial B} \psi_{n,l}(\mathbf{x}_B) \tilde{\psi}_{n,l}(\mathbf{x}_B) dS. \quad (2)$$

The symbol $\langle \rangle$ on the left side of Eq. (2) indicates an inner product with respect to the boundary surface ∂B , and the particular solutions $\psi_{n,l}$ are expressible in terms of the spheroidal functions.⁷ For example, using the spherical coordinates we can write $\psi_{n,l}$ as

$$\psi_{n,l}(r, \theta, \phi) = h_n(kr) P_{n,l}(\cos \theta) \begin{cases} \cos l\phi \\ \sin l\phi \end{cases}, \quad (3)$$

where $h_n(kr)$ and $P_{n,l}(\cos \theta)$ denote the spherical Hankel functions and Legendre functions, respectively. The basis functions $\tilde{\psi}_{n,l}$ thus obtained are mutually orthonormal on ∂B and uniformly convergent, because $\psi_{n,l}$ consist of a uniformly convergent series of Legendre functions.⁸ In this way, the radiated acoustic pressure can be written as

$$\begin{aligned} \hat{p}(\mathbf{x}) = & \rho c \tilde{C}_{0,0} \tilde{\psi}_{0,0}(r, \theta, \phi) + \rho c \tilde{C}_{1,0} \tilde{\psi}_{1,0}(r, \theta, \phi) \\ & + \rho c \tilde{C}_{1,-1} \tilde{\psi}_{1,-1}(r, \theta, \sin \phi) \\ & + \rho c \tilde{C}_{1,1} \tilde{\psi}_{1,1}(r, \theta, \cos \phi) + \tilde{C}_{2,0} \tilde{\psi}_{2,0}(r, \theta, \phi) \\ & + \rho c \tilde{C}_{2,-1} \tilde{\psi}_{2,-1}(r, \theta, \sin \phi) + \tilde{C}_{2,1} \tilde{\psi}_{2,1}(r, \theta, \cos \phi) \\ & + \rho c \tilde{C}_{2,-2} \tilde{\psi}_{2,-2}(r, \theta, \sin 2\phi) \\ & + \tilde{C}_{2,2} \tilde{\psi}_{2,2}(r, \theta, \cos 2\phi) + \dots \end{aligned} \quad (4)$$

However, the double summation in Eq. (1) is cumbersome. To facilitate the derivations, we reset the basis functions in the following manner $[\Psi_0(\mathbf{x}) = \tilde{\psi}_{0,0}(r, \theta, \phi)$, $C_0 = \tilde{C}_{0,0}$, $[\Psi_1(\mathbf{x}) = \tilde{\psi}_{1,0}(r, \theta, \phi)$, $C_1 = \tilde{C}_{1,0}$, $[\Psi_2(\mathbf{x}) = \tilde{\psi}_{1,1}(r, \theta, \sin \phi)$, $C_2 = \tilde{C}_{1,-1}$, $[\Psi_3(\mathbf{x}) = \tilde{\psi}_{1,1}(r, \theta, \cos \phi)$, $C_3 = \tilde{C}_{1,1}$, etc. Accordingly, Eq. (4) can be rewritten as

$$\begin{aligned} \hat{p}(\mathbf{x}) = & \rho c [C_0 \Psi_0(\mathbf{x}) + C_1 \Psi_1(\mathbf{x}) + C_2 \Psi_2(\mathbf{x}) \\ & + C_3 \Psi_3(\mathbf{x}) + \dots] \\ = & \rho c \sum_{j=0}^{\infty} C_j \Psi_j(\mathbf{x}), \end{aligned} \quad (5)$$

where the coefficients C_j can be determined by requiring the assumed-form solution to satisfy the boundary condition at the measurement point \mathbf{x}_m

$$\rho c \sum_{j=0}^{\infty} C_j \Psi_{m,j}(\mathbf{x}_m) = \hat{p}_m(\mathbf{x}_m). \quad (6)$$

Note that there is no restriction on the locations of measurements so long as they do not overlap. If a J -term expansion in Eq. (6) is used, then M measurements in the field must be taken, where $M \geq J$. Theoretically, if the measured acoustic pressures $\hat{p}_m(\mathbf{x}_m)$ are exact, then the assumed-form solution (6) converges to the true value as $J \rightarrow \infty$.⁹ However, this never happens in reality because the measured quantities $\hat{p}_m(\mathbf{x}_m)$ always contain errors due either to measurement uncertainties or to rapid decay of the near-field effects. While the former can be reduced by taking more averages in the measurements, the latter is irreversible. Moreover, the number of measurements M is always finite. Hence the reconstructed acoustic pressure $\hat{p}_m(\mathbf{x}_m)$ will never converge to the true value.

To enhance the accuracy of reconstruction, the least-squares method is used to eliminate the first-order error imbedded in this process. Accordingly, Eq. (6) can be rewritten as

$$[T]_{J \times J} \{C\}_{J \times 1} = \{D\}_{J \times 1}, \quad (7)$$

where $[T]_{J \times J}$ represents the transformation matrix that correlates the measured data to the reconstructed acoustic pressures, and $\{D\}_{J \times 1}$ contains the measured information, whose elements are given, respectively, by

$$T_{i,j} = \rho c \sum_{m=1}^M \Psi_{m,i}(\mathbf{x}) \Psi_{m,j}(\mathbf{x}), \quad (8a)$$

$$D_i = \sum_{m=1}^M \hat{p}_m(\mathbf{x}_m) \Psi_{m,i}(\mathbf{x}). \quad (8b)$$

Note that the transformation matrix $[T]_{J \times J}$ is nonsingular. Therefore the coefficients $\{C\}$ can be solved by inverting the matrix $[T]_{J \times J}$

$$\{C\}_{J \times 1} = (\rho c)^{-1} [E]_{J \times M} \{\hat{p}\}_{M \times 1}, \quad (9)$$

where $[E]_{J \times M}$ is called the pseudoinverse matrix defined as

$$[E]_{J \times M} = ([\Psi]_{M \times J}^T [\Psi]_{M \times J})^{-1} [\Psi]_{M \times J}^T, \quad (10)$$

where $[\Psi]_{M \times J}$ represents an $M \times J$ matrix that consists of the expansion functions $\Psi_{i,j}(\mathbf{x})$. It can be shown that the condition number of the pseudoinverse matrix $[E]_{J \times M}$ is much smaller than that of a direct inverse.¹⁰ Hence the accuracy of numerical computations is high. Once $\{C\}$ is determined, the acoustic pressure anywhere can be reconstructed by using Eq. (5).

For convenience sake, the basis functions in this paper are expressed in terms of the spherical coordinates. This is

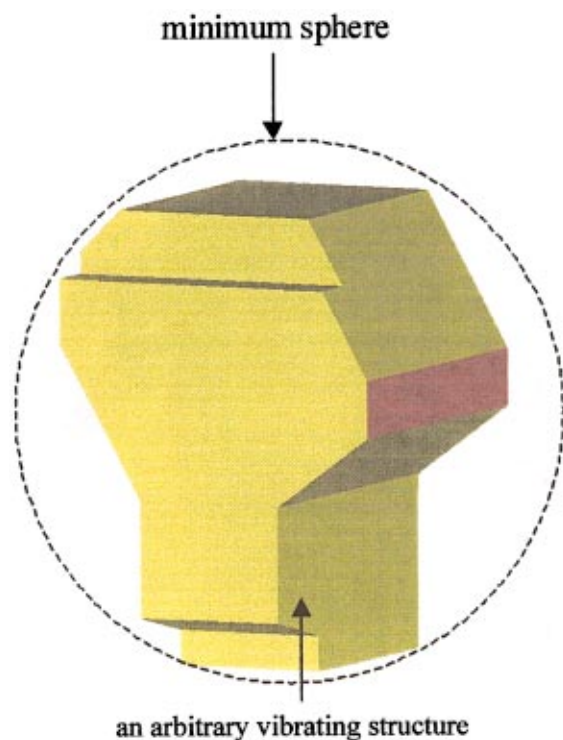


FIG. 1. Acoustic radiation from a non-spheroidal vibrating object.

because the spherical Hankel functions and Legendre functions are readily available in many mathematical libraries, such as IMSL (International Mathematical and Statistical Library) subroutines. Use of these functions can be extremely effective for spherical sources or chunky objects whose aspect ratios are close to unity, $x:y:z \rightarrow 1$. For an elongated, flat, or both elongated and flat object, the prolate, oblate, or elliptic coordinates may be used, respectively, to provide faster convergence in numerical computations. However, the spheroidal functions in these coordinates are not available in many mathematical libraries, and direct numerical computations of these functions can be time consuming.

II. ADVANTAGE AND LIMITATION OF THE HELS METHOD

The main advantage of the HELS method is its simplicity in expressing the acoustic pressure in terms of an expansion of certain basis functions. One can interpret these basis functions as a superposition of multipole expansion, with its coefficients being determined by matching the assumed-form solution with the pressure boundary condition at measurement points.

Like any methodology, the HELS method has limitations. Theoretically, the HELS method may yield an exact reconstruction of the radiated acoustic pressure field from a spheroid such as a spherical, oblate, prolate, elliptic spheroid, etc. Among these spheroids, the spherical surface is the most convenient because the basis functions are readily available and numerical computations can be extremely efficient. For a nonspheroidal surface, the acoustic pressure field reconstructed by the HELS method may be valid outside a minimum spheroid^{11,12} that encloses the source (see Fig. 1), but is approximate at best inside this control surface.

Note that one can use the multipole expansion theory to reconstruct the acoustic pressure field with its coefficients determined by the orthogonality properties of expansion functions. Under this condition, the reconstructed acoustic pressures would be exact outside the minimum sphere that encloses the source, but invalid inside this control surface.¹² The HELS method uses a superposition of multipole expansion with its coefficients determined by forcing the assumed-form solution to satisfy the pressure boundary condition at measurement locations. This renders the resulting reconstruction approximate in the entire exterior region for an arbitrary surface, with relatively high accuracy outside the minimum sphere but relatively poor accuracy inside. Of course it would be silly to apply the HELS method to a highly irregular surface. Experimental results⁵ have demonstrated that satisfactory reconstruction of the radiated acoustic pressure fields can be obtained for a nonspherical but smooth surface with relatively few measurements.

Needless to say, the accuracy of reconstruction depends on the accuracy of the input data, which is the case for all inverse problems. If the problem under consideration is to reconstruct the far-field acoustic pressure, for example in predicting vehicle pass-by noise, then the accuracy of reconstruction is determined primarily by the measurement uncertainties. The errors due to the loss of near-field acoustic pressures in the input will have a negligible impact on reconstruction, because this near-field effect decays exponentially with distance. On the other hand, if acoustic pressures on the source surface are sought, as in the case of noise diagnostics, the accuracy of reconstruction will depend critically on that of the measured near-field acoustic pressures. If this information is incomplete, then there will be no way of reconstructing accurately the surface acoustic pressure which is predominated by the near-field effect. Under these circumstances, increasing the number of expansion terms J will not help; in fact, it may even worsen the accuracy of reconstruction. This is because the problem is already ill-posed, hence adding more terms may simply introduce more errors and eventually cause the solution to diverge beyond a certain value of J .

It is emphasized that in engineering applications the acoustic near field cannot be captured completely. Hence an exact reconstruction of the radiated acoustic pressure distribution over the surface of a complex vibrating structure is not possible. Taking more measurements at closer distances can improve the accuracy of the near-field acoustic pressure measurements, thus leading to a more accurate reconstruction. However, this may prolong the diagnostic process and even make it impractical.

III. IMPLEMENTATION OF THE HELS METHOD

In this investigation we attempt to develop a simple yet effective methodology to reconstruct the radiated acoustic pressure fields. To achieve this goal we must devise a procedure that can yield the desired accuracy with relatively few measurements. The questions that must be addressed include: (1) Where and how the measurements should be taken? (2) How many measurements are necessary to achieve

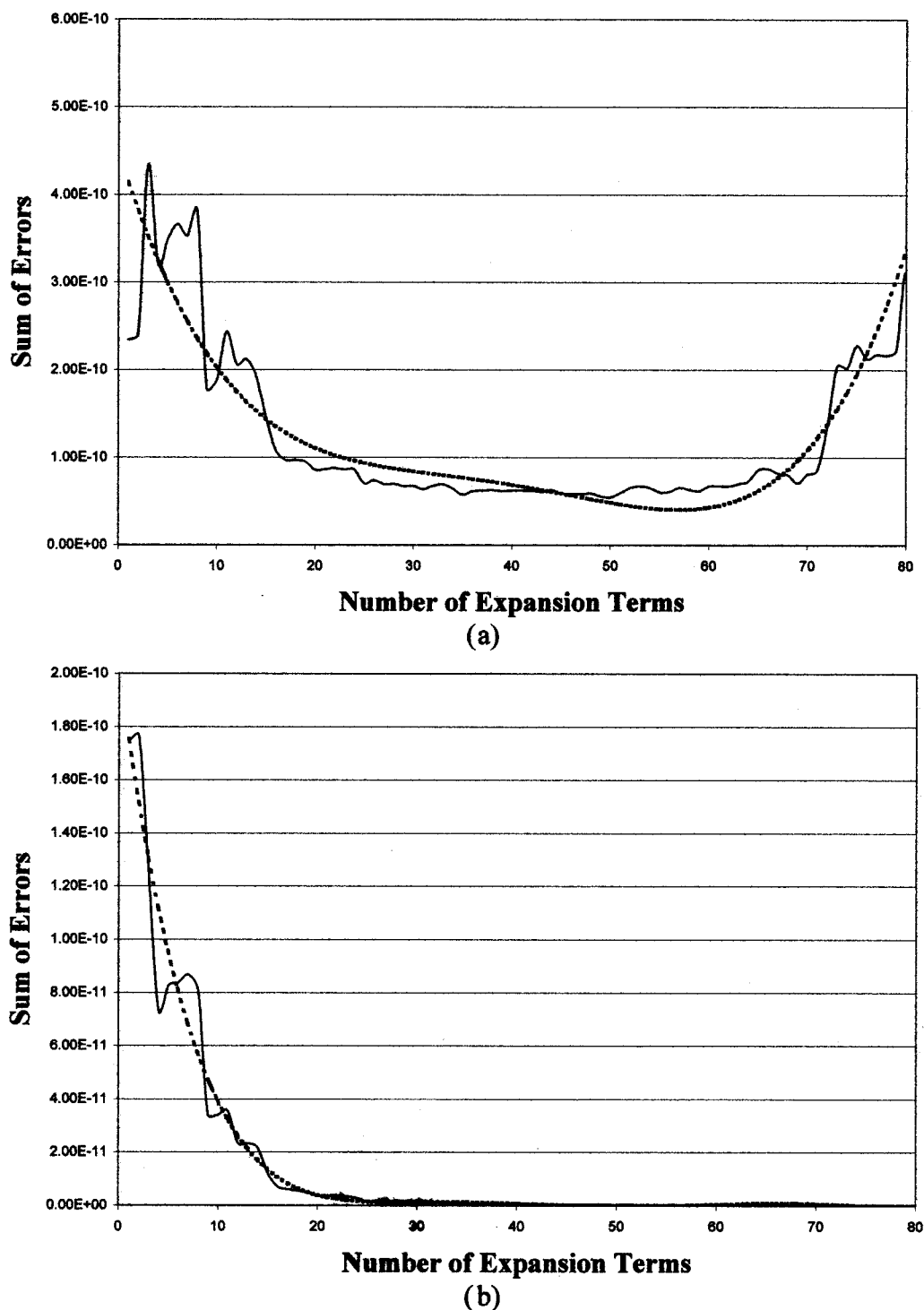


FIG. 2. Typical $\|L\|^2$ -norm errors in reconstruction versus the number of expansion terms. Solid line: Original errors; dashed line: Curve-fitted errors. (a) On source surface; (b) on measurement surface.

the desired resolution? and (3) Given the number of measurements, what is the optimal number of expansion terms?

Unfortunately, there are no definite answers to all these questions because of the uncertainties involved in an inverse problem. What we can do is to develop some guidelines with which satisfactory reconstruction of the radiated acoustic pressure fields can be obtained.

Experiments have demonstrated^{4,5} that in general the closer the measurement to a source surface, the more accu-

rate the reconstructed acoustic pressures. Therefore we must try to take measurements in the near field. Suppose that the measurement distance is d , the characteristic dimension of a source a , and the maximum frequency of interest f_{\max} . Then in order to stay in the near field, the value of d should satisfy the following three conditions simultaneously: (1) $d \ll a$; (2) $d \ll \lambda_{\min} = c/f_{\max}$; and (3) $d \ll k_{\max} a^2/2$, where $k_{\max} = 2\pi/\lambda_{\min}$. Also, it would be better to take measurements over a conformal surface that completely encloses the

source. Experimental data have shown⁴ that a conformal measurement surface always leads to a more accurate reconstruction than a planar measurement surface does. This is because a planar surface often extends beyond the near-field region, thus making measurements less accurate. However, measurements over a conformal surface are not easy in practice, whereas those over a planar surface can be facilitated by an x - y translator. Hence a preferred way is to take measurements over several small planar surfaces enclosing a source at the closest possible distance.

The spatial resolution of reconstruction depends on the frequency of interest. From signal analysis, we have learned that the sample rate must be greater than twice the highest frequency of interest in order to avoid aliasing.¹³ In practice, the sample rate is often selected to be 2.5 times the highest frequency of interest to make the process cost-effective. If this analogy were adopted, one would need at least 2.5 samples over the shortest wavelength of interest, namely, the interval between measurements must be $\Delta \leq \lambda_{\min}/2.5$. Thus, reconstruction of acoustic pressures radiated from a source of characteristic dimension of 1 m up to 1000 Hz can be done by taking measurements over planar surfaces at $d=0.1$ m away from the source at an interval of $\Delta=0.1$ m. This measurement grid would satisfy both the distance and spatial resolution requirements. However, the total number of measurements over six sides would reach almost 600, which may be excessive.

One unique feature of the HELS method is that reconstruction of the radiated acoustic pressure is not based on spatial sampling, but on synthesis of spheroidal functions. For example, if the source radiates a pure dipole sound, then theoretically reconstruction can be done exactly with no more than four expansion terms or equivalently, four measurements regardless the frequency. For an arbitrarily vibrating structure, the radiated acoustic pressure may be quite complex. Nevertheless, this acoustic pressure field can be expressed as a multipole expansion. Moreover, at low-to-mid frequencies the major contributions are from the first few expansion functions, which represent the effects of one monopole, three dipoles pointing in the (x,y,z) -axes directions, nine quadrupoles (three longitudinal plus six lateral quadrupoles), etc. The higher-order terms represent the small-scale effects and can be neglected. Accordingly, the number of measurements necessary to determine the corresponding coefficients in Eq. (6) may be significantly reduced.

Since an exact solution to the surface acoustic pressure distribution over a complex vibrating structure is unknown, it is difficult to judge the accuracy of reconstruction. The only information available would be the measured acoustic pressures. Experimental results^{4,5} have demonstrated that the accuracy of reconstruction on the measurement surface increases with the number of expansion terms. However, the accuracy of reconstruction on the source surface increases first with the number of expansion terms up to certain value, then decreases monotonically thereafter. Thus, it is possible to devise a systematic approach that enables one to select an optimal number of basis functions (see Fig. 2) to reconstruct the radiated acoustic pressure.

It is emphasized that this optimal number is not caused by the use of Gram–Schmidt orthonormalization, but inherent in this inverse acoustic problem. In a recent paper,¹⁴ we skipped this orthonormalization and least-squares processes and solved a set of simultaneous equations by singular value decomposition (SVD) followed by regularization as suggested by an anonymous reviewer. Still, we observed an optimum number of expansion terms, only occurring at a different value. Such a phenomenon was also reported by Bobrovnikii¹⁵ in reconstructing full vibration field using an expansion of normal modes of surface vibration.

The presence of an optimum number of expansion terms seems to be reasonable. Physically the higher-order terms represent the small-scale or near-field effect, which may have been lost in the measured data and cannot be recovered anyway. On the other hand, the lower-order terms depict the large-scale or propagating wave effect, which may be captured in the measured data and, therefore, can be reconstructed. Consequently, by using an optimal number of basis functions it is possible to obtain a satisfactory reconstruction of the radiated acoustic pressure fields in both near and far fields in a cost-effective manner.

IV. TEST SETUP

Figure 3 shows the test setup for a full-size four-cylinder engine and its accessories designed for a passenger vehicle. Obviously, modeling the dynamic responses of the whole system would be extremely complicated, and in this case unnecessary. For the purpose of examining the effectiveness of the HELS method on reconstructing the radiated acoustic pressures from a general three-dimensional structure, it suffices to consider a simplified yet arbitrarily shaped object.

Without loss of generality, we approximate the real engine by the one that has the same overall dimensions and shapes, but with its accessories such as alternator, cooling fan, pulleys, belts, etc. removed. Figure 4 shows such a simplified engine block with an overall length of 0.435 m in the z axis direction, an overall width of 0.46 m in the x axis direction, and an overall height of 0.63 m in the y axis direction.

Note that this simplified engine block is no less complex than the original one because it contains sharp edges and corners, and has abrupt changes in surface contour; whereas the original one has smooth and rounded edges and corners. The acoustic near field around these irregularities on the simplified engine block can be extremely complicated because of local acoustic reflection, diffraction, and scattering effects.

To simulate sound radiation from this simplified engine block, harmonic forces are assumed to act uniformly on two arbitrarily selected surfaces. The amplitude of the pressure acting on the protruding part of the front surface (marked in red in Fig. 4) is 2000 N/m², while that on the entire left surface (not shown in Fig. 4) is 800 N. The engine block is approximated by a solid piece of steel. Since in practice an engine is often mounted on rigid frames, the bottom surface of the engine block shown in Fig. 4 is assumed to be clamped with zero displacement and slope, while the remaining surfaces are free to move. The dynamic responses of this

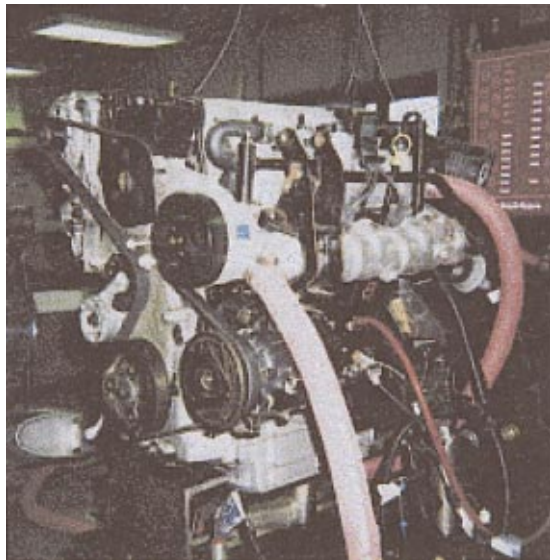


FIG. 3. Test setup for a full-size, four-cylinder engine for a passenger vehicle.

engine block are solved numerically using the commercial FEM software I-DEAS¹⁶ under various excitation frequencies. In carrying out numerical computations, the engine block is divided into 2616 elements with a total of 4228 nodes.

Once the normal component of the surface velocity distribution is determined, the surface acoustic pressures are solved using a standard BEM code.¹⁷ In this paper several excitation frequencies are considered. For simplicity, however, the numbers of segments and nodes discretized on the engine block surface are kept the same, at 368 and 1106, respectively, for all frequencies. Having obtained the normal component of surface velocity and surface acoustic pressure, the radiated acoustic pressure at any field point is calculated using the Helmholtz integral formulation.

Note that in acquiring the field acoustic pressures, we could set the locations on several small planar surfaces that encloses the simplified engine block at very close distances so as to improve the accuracy of the near-field acoustic pres-

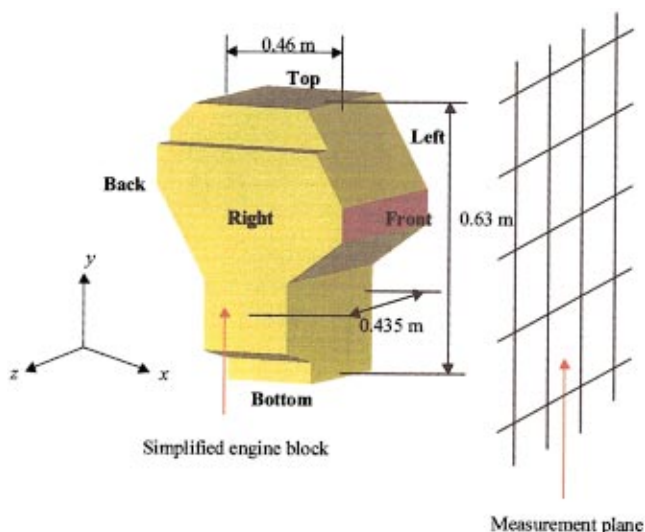
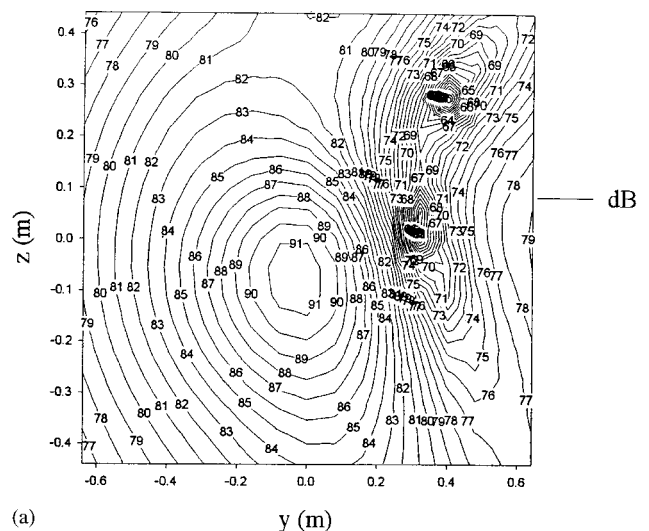
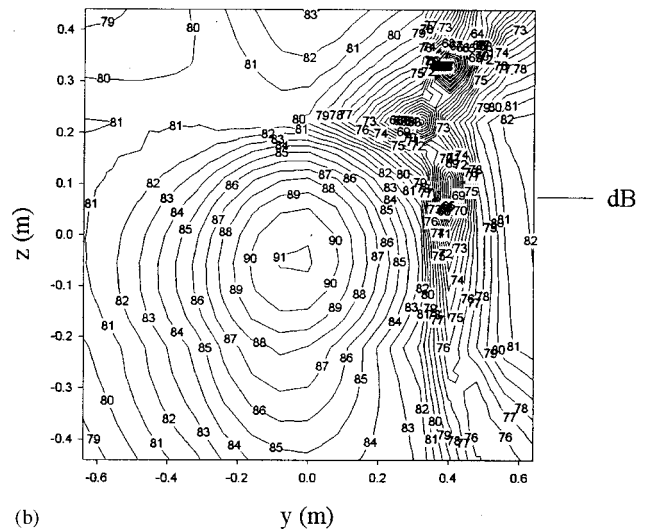


FIG. 4. Schematic of a simplified engine block and measurement locations.



(a)



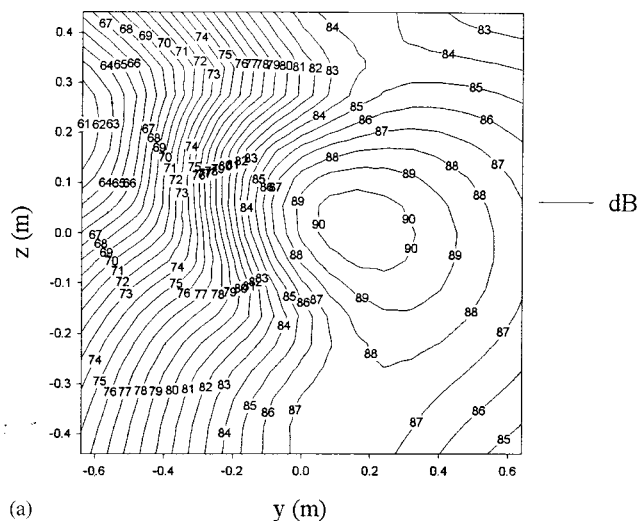
(b)

FIG. 5. Comparison of acoustic pressure amplitudes over a planar surface at $x=0.73$ m in front of the simplified engine block at 120 Hz. (a) BEM; (b) HELS method.

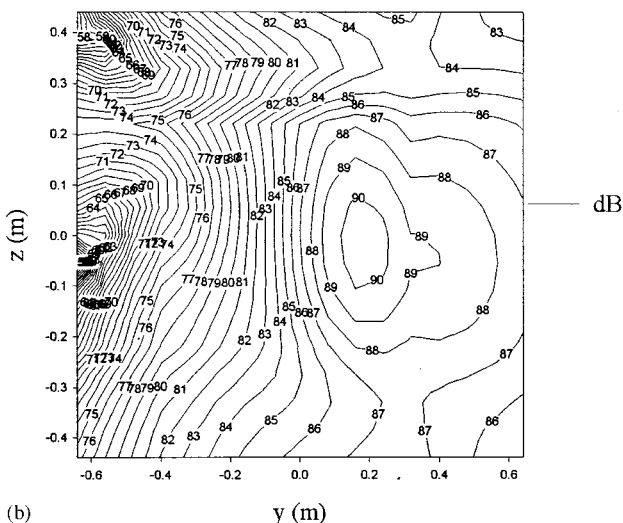
sure measurements. However, this kind of data acquisition would correspond to a prolonged process, which may not be realistic in practice. To test the effectiveness of the HELS method, we decide to take 36 measurements uniformly distributed over a planar surface at $x=+0.55$ m with respect to the center of the engine or equivalently, $d=0.32$ m away from the engine's front surface with a measurement interval of $\Delta=0.22$ m. This measurement plane happens to parallel with one of the surfaces on which harmonic forces are applied. In addition, we take three measurements over the back, top, left, and right sides of the engine at the same distance and interval, respectively, thus making a total of 48 measurements in the field. The reason for setting the measurement distance at $d=0.32$ m is to ensure that all the input data fall outside the minimum sphere.¹²

V. RECONSTRUCTION SCHEME

The measured acoustic pressures described in the preceding section are taken as the input to Eq. (9). Note that these data are not error-free, because they are obtained nu-



(a)



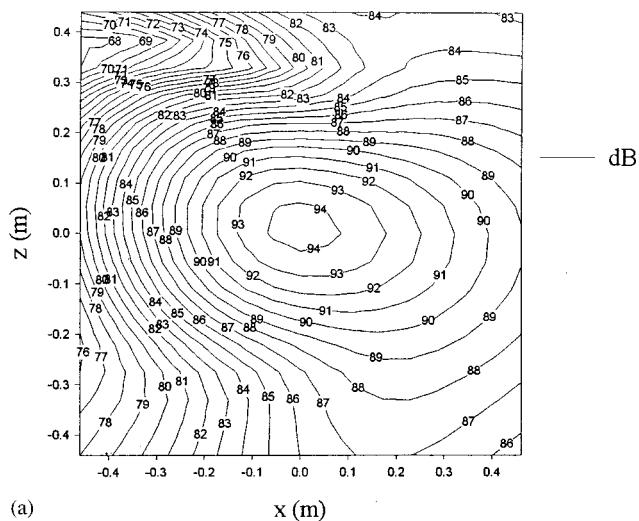
(b)

FIG. 6. Comparison of acoustic pressure amplitudes over a planar surface at $x = -0.73$ m behind the simplified engine block at 120 Hz. (a) BEM; (b) HELS method.

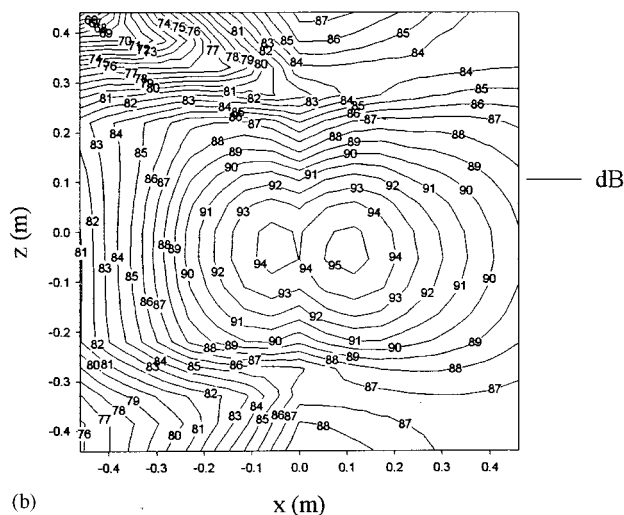
merically using FEM and BEM codes. Moreover, some near-field effects may have been lost since the acoustic pressures are taken at certain distances from the source. These erroneous data are then used to reconstruct the acoustic pressures in the entire field, including the engine surface which is predominated by the near-field effect. Hence the problem may be ill posed. Such an ill-posedness difficulty is inevitable in an inverse problem.

Because of the presence of this ill-posedness difficulty, increasing the number of basis functions will not necessarily improve the accuracy of reconstruction. Numerical tests have indicated that the accuracy of reconstruction on the source surface increases at first, but then decreases with the number of basis functions [see Fig. 2(a)]. This is because each addition of the basis functions may introduce new errors to the matrix equation [Eq. (9)]. As errors accumulate, the accuracy of reconstruction deteriorates. Eventually the solution to Eq. (9) may simply diverge beyond a certain value of J .

On the other hand, the accuracy of reconstruction on the measurement surface always increases with the number of



(a)



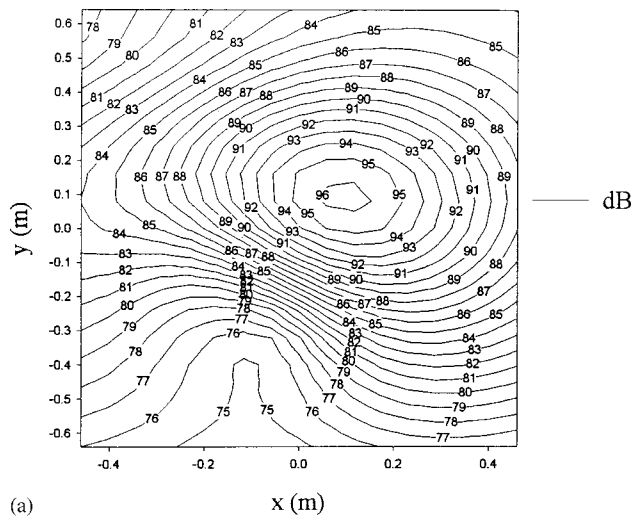
(b)

FIG. 7. Comparison of acoustic pressure amplitudes over a planar surface at $y = 0.815$ m above the simplified engine block at 120 Hz. (a) BEM; (b) HELS method.

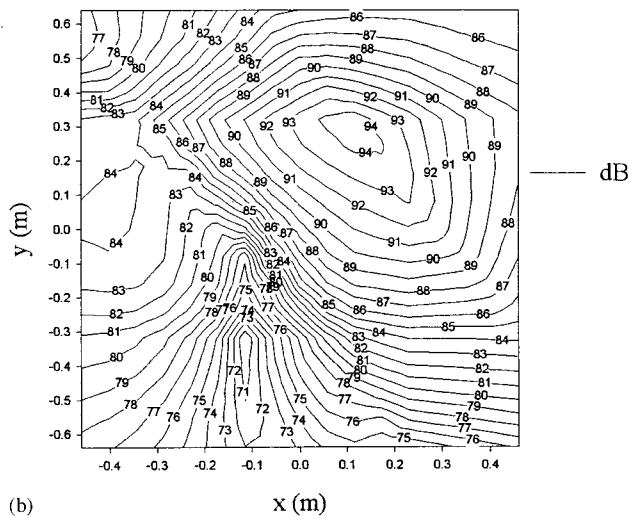
expansion terms [see Fig. 2(b)]. This is because the errors in reconstruction on the measurement plane have been minimized by the least-squares method. Since there is no prior knowledge of the surface acoustic pressures on a complex vibrating structure, it is impossible to know exactly where to truncate the expansion so as to obtain the best reconstruction. However, for the HELS method to become a practical noise diagnostic tool, it is imperative to develop a procedure to select an optimal number of expansion terms that may yield a satisfactory reconstruction of the acoustic pressure field. Such a procedure is described below.

It is emphasized that because of the nature of the problem, the optimal number of basis functions cannot be determined precisely. In this paper, we propose to optimize the number of expansion terms with respect to the acoustic pressures on the measurement surface.

First, we determine the maximum frequency that will fit the present measurement setup. Given the average overall dimension of the engine $a = (0.435 + 0.43 + 0.63)/3 = 0.508$ m, the measurement distance $d = 0.32$ m, and the measurement interval $\Delta = 0.22$ m, a reasonable value of f_{\max}



(a)



(b)

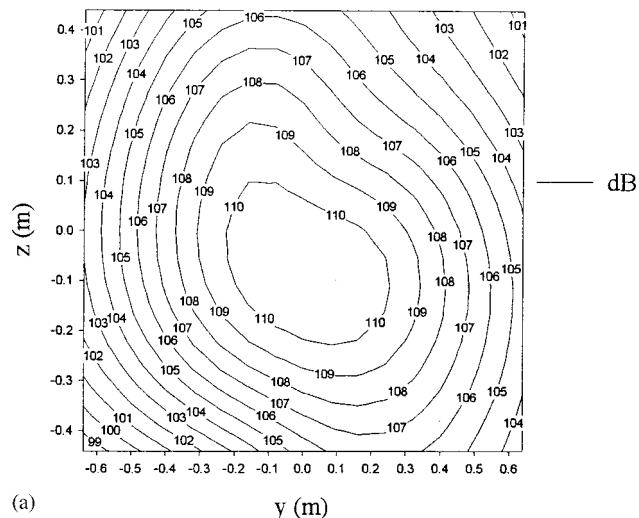
FIG. 8. Comparison of acoustic pressure amplitudes over a planar surface at $z = -0.715$ m on the left side of the simplified engine block at 120 Hz. (a) BEM; (b) HELS method.

would be 500 Hz with a wavelength $\lambda_{\min} = c/f_{\max} = 0.686$ m. Hence all three conditions for the measurement distance and the spatial resolution requirement are satisfied.

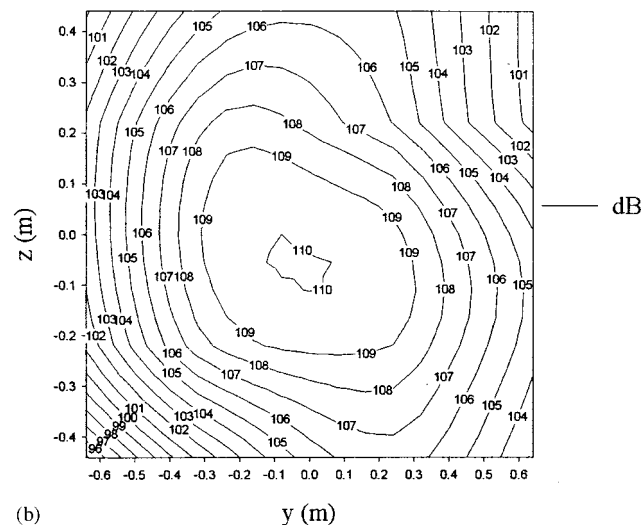
To find an optimal number of expansion terms in reconstruction with respect to the selected 48 measurements, we start from $N=1$, which represents a four-term expansion because $J=(N+1)^2=4$. The corresponding basis functions are: $h_0(kr)P_{0,0}(\cos \theta)$, $h_1(kr)P_{1,0}(\cos \theta)$, $h_1(kr) \times P_{1,1}(\cos \theta)\sin \phi$, and $h_1(kr)P_{1,1}(\cos \theta)\cos \phi$. Solving a 4×48 matrix equation [Eq. (9)] yields the coefficients C_j , $j=1$ to 4. The acoustic pressure field $\hat{p}(\mathbf{x})$ can now be reconstructed by using Eq. (5) and its accuracy is checked by calculating the $\|L\|^2$ -norm errors of the reconstructed acoustic pressures with respect to the measured ones $\hat{p}_m(\mathbf{x})$ at the same locations

$$\|L\|^2 = \sum_{i=1}^M |\hat{p}(\mathbf{x}_i) - \hat{p}_m(\mathbf{x}_i)|^2. \quad (11)$$

Next, we increase the number of expansion terms by one, $J=J+1$. Then we solve a $(J+1) \times M$ matrix equation



(a)



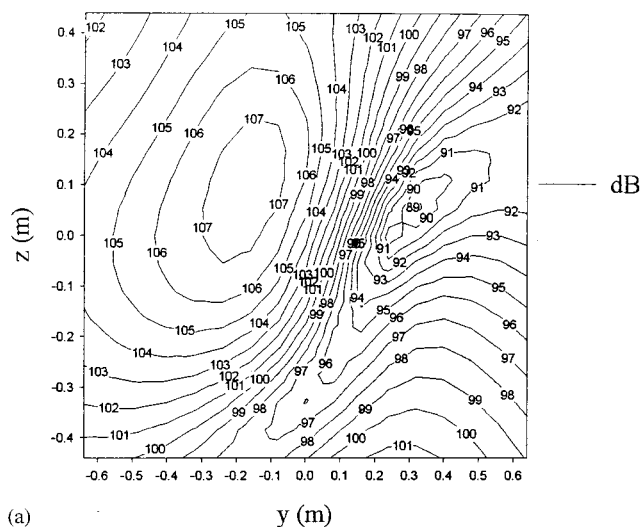
(b)

FIG. 9. Comparison of acoustic pressure amplitudes over a planar surface at $x = 0.73$ m in front of the simplified engine block at 360 Hz. (a) BEM; (b) HELS method.

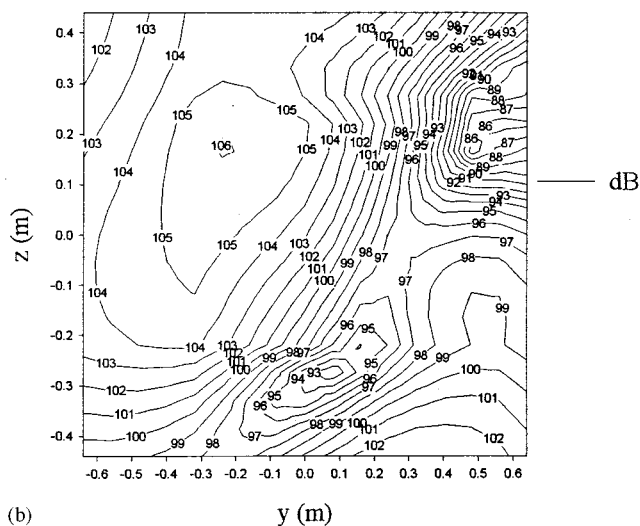
[Eq. (9)] for C_j , use Eq. (5) to reconstruct the acoustic pressures on the measurement plane, and calculate the $\|L\|^2$ -norm errors again. Because of the measurement uncertainties involved in the input data, the $\|L\|^2$ -norm errors may fluctuate with respect to the value of J . Hence curve-fitting is utilized to smooth the $\|L\|^2$ -norm error curve.

Since the accuracy of reconstruction on the measurement plane increases with J , the slope of this $\|L\|^2$ -norm error curve approaches zero, $\tan \alpha \rightarrow 0$, beyond certain values of J . However, the optimum value of J_0 which produces the best reconstruction on the source surface may occur at a nonzero slope, $\tan \alpha = \epsilon_0 > 0$. Consequently, the iterations are stopped as soon as $\tan \alpha \approx \epsilon_0$. The corresponding value of J_0 can be interpreted as being optimized with respect to the M measured acoustic pressures. Obviously, different number of measurements would yield different optimal value of J_0 .

This optimal value J_0 may be different for different frequencies as well. In general, the higher the frequency, the larger the value of J_0 . Consequently, the efficiency of numerical computations decreases as the frequency increases.



(a)



(b)

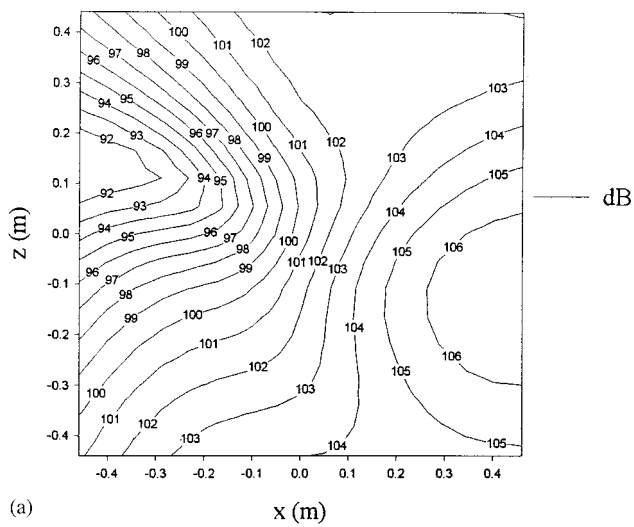
FIG. 10. Comparison of acoustic pressure amplitudes over a planar surface at $x = -0.73$ m behind the simplified engine block at 361 Hz. (a) BEM; (b) HELS method.

This inefficiency at high frequencies is not unique to the HELS method, but inherent in all expansion theories, as well as in the standard FEM/BEM codes. Therefore, other methods such as asymptotic approximations should be used to reconstruct the radiated acoustic pressure field in the high-frequency regime.

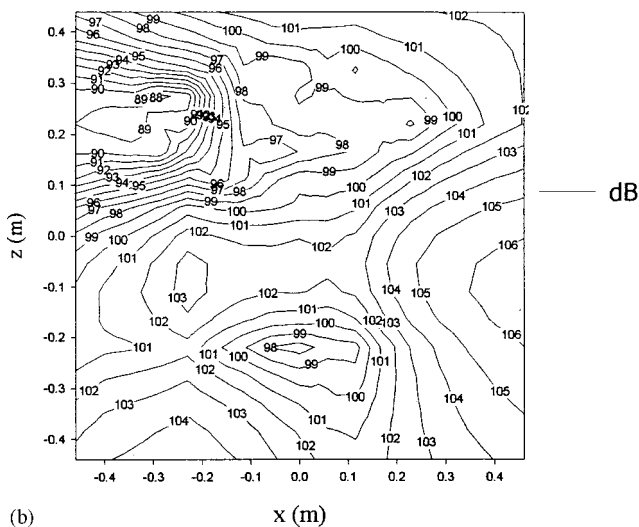
The optimization process as discussed above is repeated for all frequencies. The values of J_0 thus determined are then used to reconstruct the acoustic pressures anywhere, including the source surface.

In this investigation, the radiated acoustic pressure fields from a simplified engine block subjected to harmonic excitations at 120 and 361 Hz are reconstructed. Using the above reconstruction scheme, the optimal number of basis functions is found to be $J_0 = 18$ for both frequencies. Therefore, a set of 18×48 matrix equation [Eq. (9)] is solved in each case for the corresponding coefficients C_j , $j = 1$ to 18. Because of a relatively small size of the matrix, numerical computations are extremely fast.

Note that the engine block under consideration is excited



(a)



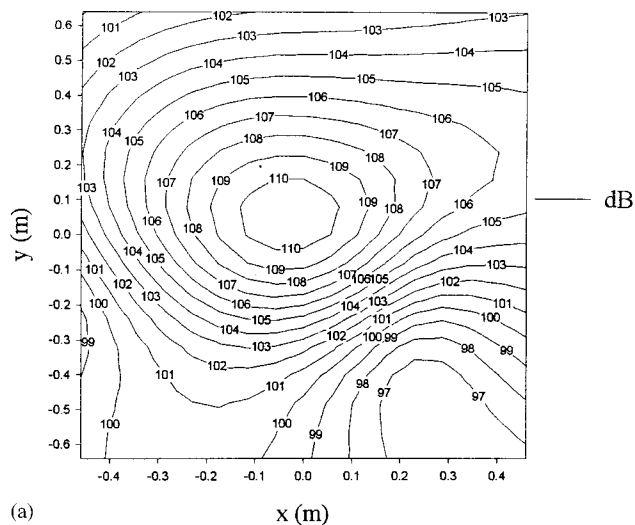
(b)

FIG. 11. Comparison of acoustic pressure amplitudes over a planar surface at $y = 0.815$ m above the simplified engine block at 361 Hz. (a) BEM; (b) HELS method.

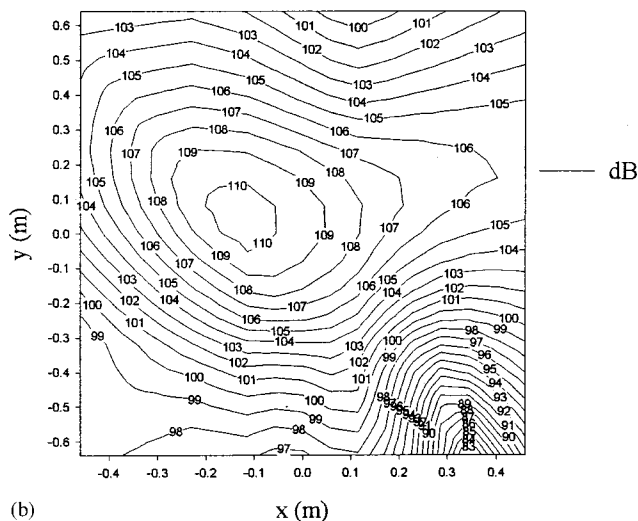
by harmonic forces distributed uniformly over two surfaces. Because of the size difference, the total force acting on the left surface is larger than that on the front surface. Therefore, it would be better to take at least as many measurements over the left surface as those over the front one so as to improve the accuracy of reconstruction. However, in engineering applications the excitation forces are usually unspecified, so the best measurement locations are unknown *a priori*. To simulate the real situation and to test the robustness of the HELS method, we focus (36) measurements on the front side of the engine block with (three) supplemental measurements over each of the back, top, left, and right sides, respectively. These input data are used to reconstruct the radiated acoustic pressure field.

VI. NUMERICAL RESULTS

This section displays two-dimensional contour plots of the reconstructed acoustic pressures radiated from a simplified engine block based on 48 measurements with 18 expansion functions in Eq. (5). It is emphasized that in this example a somewhat unfavorable condition is assumed. This is



(a)

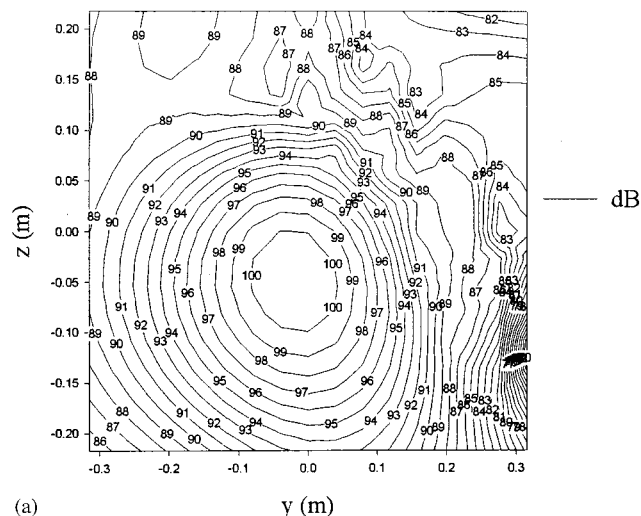


(b)

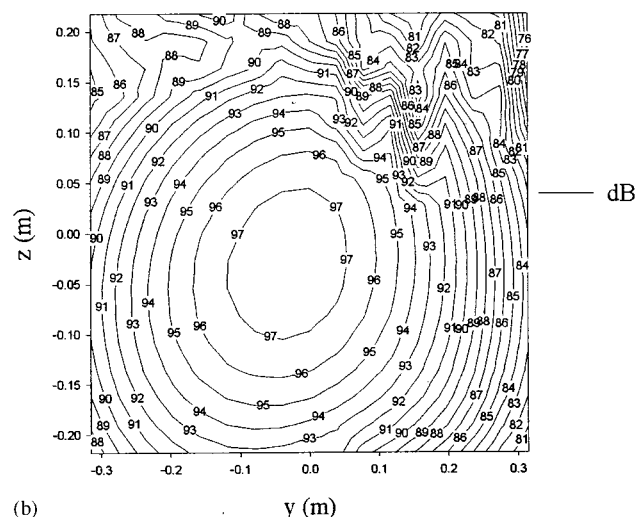
FIG. 12. Comparison of acoustic pressure amplitudes over a planar surface at $z = -0.715$ m on the left side of the simplified engine block at 361 Hz. (a) BEM; (b) HELS method.

because in reality the test environment may not allow for taking enough measurements at very close distances. Moreover, the vibrating structures often have complex shapes and the excitation forces are unspecified. These adverse circumstances may severely affect the accuracy of data acquisition. Consequently, it would be unrealistic to expect a precise reconstruction of the acoustic pressure field.

Figure 5 shows the comparison of the reconstructed acoustic pressure distribution over a planar surface four times the size of the front surface of the simplified engine block at $x = 0.73$ m and 120 Hz with the benchmark values using the BEM codes. Results show that not only a large pressure peak in the center, but also two relatively small ones toward the right are captured. The difference between the reconstructed and benchmark acoustic pressure peak amplitudes is about 1 dB. Considering the fact that reconstruction is done by taking only 48 measurements, these results seem quite good. However, this is not surprising because the reconstruction region is outside the minimum sphere where the acoustic pressure field can be adequately represented by a superposition of spherical harmonic functions.



(a)



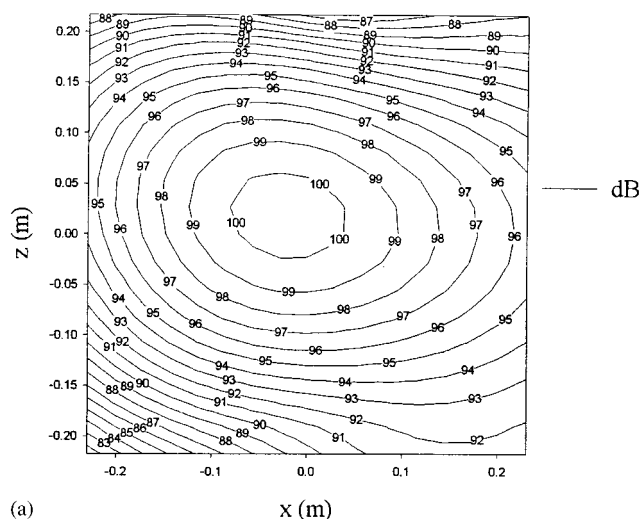
(b)

FIG. 13. Comparison of acoustic pressure amplitudes over the entire front surface of the simplified engine block at 120 Hz. (a) BEM; (b) HELS method.

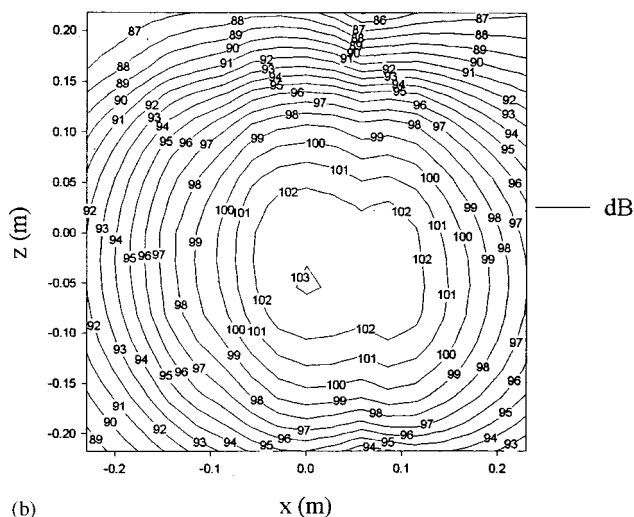
Figures 6–8 depict comparisons of the reconstructed and benchmark acoustic pressure distributions over similar planar surfaces behind, above, and on the left side of the simplified engine block at 120 Hz, respectively. While the accuracy in reconstruction using the HELS method is far from perfect, it nonetheless represents the only methodology known to the author that enables one to visualize the radiated acoustic pressure field around a complex vibrating structure based on a few measurements taken primarily on one side of the source.

Figures 9–12 illustrate the comparisons of the reconstructed acoustic pressure distributions over these same planar surfaces in front of, behind, above, and on the left side of the simplified engine block at 361 Hz with the corresponding benchmark values, respectively. Good agreements are obtained in these cases as well.

The situation becomes quite different, however, when reconstruction is attempted inside the minimum sphere that encloses a nonspherical surface. This is because the superposition of spherical harmonic functions may no longer be enough to reproduce the acoustic pressures in this region



(a)

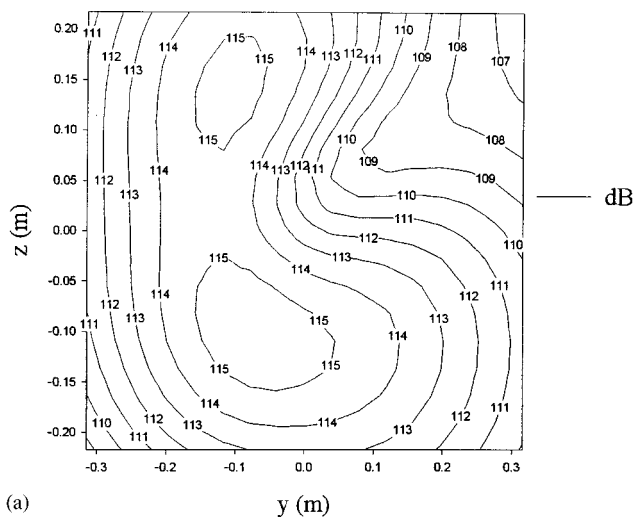


(b)

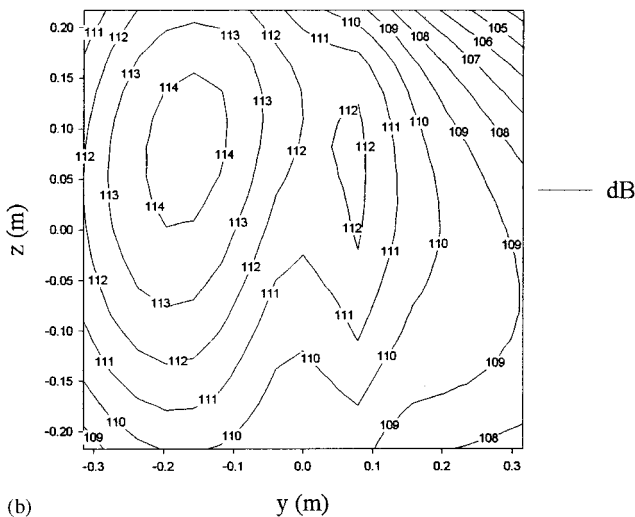
FIG. 14. Comparison of acoustic pressure amplitudes over the entire top surface of the simplified engine block at 120 Hz. (a) BEM; (b) HELS method.

where the near-field acoustic pressure is predominant. Moreover, part of the near-field effect may have been lost in the input data, thus the problem becomes highly ill posed. As a result, the HELS formulation can only yield at best an approximate solution.

Figure 13 shows the comparison of the reconstructed acoustic pressure distribution over the entire front surface of the engine block at 120 Hz with the benchmark values. Note that the front surface of the present engine block is nonplanar, containing abrupt changes in surface contour, and the two-dimensional contour plots shown in Fig. 13 are projections of the surface acoustic pressures onto a planar surface. Obviously, the accuracy of reconstruction is not as high as that shown in Fig. 5. Comparison of the reconstructed and benchmark acoustic pressure amplitudes at 120 Hz over the entire top engine block surface is demonstrated in Fig. 14. The differences between the reconstructed and benchmark acoustic pressure peak values are about 3 dB. Further, the locations of the reconstructed acoustic pressure peaks seem shifted slightly from those of the benchmark results. These discrepancies are caused not only by the errors in the input



(a)



(b)

FIG. 15. Comparison of acoustic pressure amplitudes over the entire back surface of the simplified engine block at 361 Hz. (a) BEM; (b) HELS method.

data and truncations of the expansion functions, but also by the fact that the engine block surface falls inside the minimum sphere.

Comparisons of the reconstructed acoustic pressure amplitudes with the benchmark results over the entire back and left surfaces of the simplified engine block at 361 Hz are displayed in Figs. 15 and 16, respectively. Once again, the accuracy of reconstruction on the engine block surface is not as high as that outside the minimum sphere as expected.

It is emphasized that the accuracy of reconstruction on the source surface can be enhanced by taking more measurements over a conformal surface at very close range.^{4,5,14} The present example is designed to examine the robustness and effectiveness of the HELS method on a general, complex vibrating structure under relatively unfavorable conditions.

VII. CONCLUDING REMARKS

The HELS method is used to reconstruct the acoustic pressure fields radiated from an arbitrarily shaped vibrating structure, based on a finite number of measurements in the field. To examine the effectiveness of the HELS method,

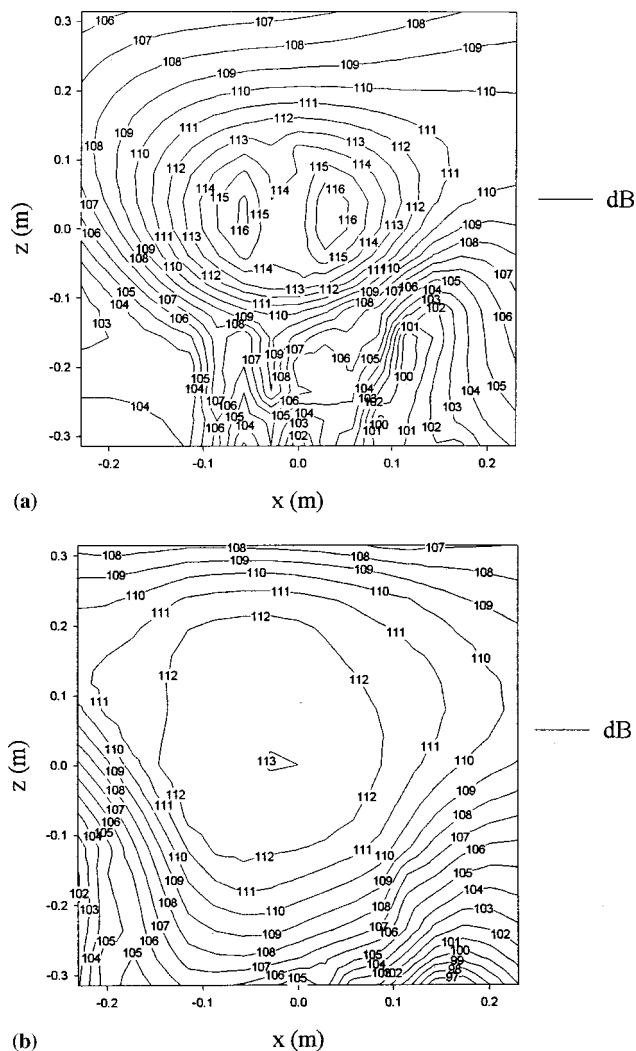


FIG. 16. Comparison of acoustic pressure amplitudes over the entire left surface of the simplified engine block at 361 Hz. (a) BEM; (b) HELS method.

adverse conditions are designed, which include abrupt changes in surface contour and sharp edges and corners on source geometry, and a relatively small number of measurements taken primarily on a finite planar surface on one side of the source. These unfavorable conditions are of significance because in practice a vibrating structure is always of arbitrary shape, and the location and number of measurements are often restricted. The HELS method seems to be able to work under these circumstances and produce reasonably good results at least in the low-to-mid frequency regime. Further, a systematic reconstruction scheme is developed to provide an optimal number of expansion terms with respect to the measured acoustic pressures. The test results show that the reconstructed acoustic pressures agree satisfac-

torily with the benchmark values. Moreover, the accuracy of reconstruction increases significantly as the reconstruction surface moves to the far field. The efficiency of numerical computations, however, may deteriorate at high frequencies. This high-frequency inefficiency is inherent in all expansion theories, as well as in the standard FEM/BEM codes. Therefore other methods such as asymptotic approximations must be developed for reconstructing acoustic pressures in the high-frequency regime.

ACKNOWLEDGMENTS

The author would like to express his most sincere appreciation to Dr. Yuhua Wu for his assistance in carrying out numerical computations in this investigation. This work was supported by National Science Foundation Grant No. CMS-9802847.

¹Z. Wang and S. F. Wu, "Helmholtz Equation Least-Squares (HELs) method for reconstructing the acoustic pressure field," *J. Acoust. Soc. Am.* **102**, 2020–2032 (1997).

²S. F. Wu and A. Wang, "Noise diagnostic system," U.S. Patent Number: 5712805 (1998).

³S. F. Wu and J. Yu, "Reconstructing Interior Acoustic Pressure Fields via Helmholtz Equation Least-Squares (HELs) Method," *J. Acoust. Soc. Am.* **104**, 2054–2060 (1998).

⁴K. Leach and S. F. Wu, "Visualization of sound radiation from a bowling ball," *Proceedings of the 1999 ASME Noise Control and Acoustics Division, NCA-Vol. 26*, pp. 209–215, Nashville, Tennessee, November 1999.

⁵N. Rayess and S. F. Wu, "Visualization of acoustic pressure radiation from complex vibrating structures," *Proceedings of the ASME Noise Control and Acoustics Division, NCA-Vol. 26*, pp. 387–392, Nashville, Tennessee, November 1999.

⁶M. A. Pinsky, *Partial Differential Equations and Boundary-Value Problems with Applications* (McGraw-Hill, New York, 1991), pp. 18.

⁷I. N. Vekua, "On completeness of a system of metaharmonic functions," *Dokl. Akad. Nauk SSSR* **90**, 715–718 (1953).

⁸S. Bergman, *Integral Operators in the Theory of Linear Partial Differential Equations* (Springer, Berlin, 1960), pp. 66–67.

⁹P. J. Davis and P. Rabinowitz, "Advances in orthonormalizing computation," in *Advances in Computers*, edited by F. L. Alt (Academic, New York, 1961), Vol. 2, pp. 55–133.

¹⁰G. W. Stewart, *Introduction to Matrix Computations* (Academic, New York, 1973), pp. 184–230.

¹¹D. S. Burnett, "A three-dimensional acoustic infinite element based on a prolate spheroidal multipole expansion," *J. Acoust. Soc. Am.* **96**, 2798–2816 (1994).

¹²E. Williams, *Fourier Acoustics: Sound Radiation and Nearfield Acoustical Holography* (Academic, San Diego, California, 1999), pp. 206–210.

¹³M. Bellanger, *Digital Processing of Signals: Theory and Practice* (Wiley, New York, 1984), pp. 18 and 19.

¹⁴N. Rayess and S. F. Wu, "Experimental validation of the HELS method on reconstructing radiated acoustic pressures from a complex vibrating structure," *J. Acoust. Soc. Am.* (to be published).

¹⁵Y. Bobrovnikii, "Problem of reconstructing the full vibration field from limited data," *J. Acoust. Soc. Am.* **105**, 968(A) (1999).

¹⁶I-DEAS User Manual, *Integrated Design Engineering Analysis Software*, SDRC, Milford, OH, 1988.

¹⁷S. F. Wu and Y. Wu, "Reconstruction of radiated acoustic pressure fields from a complex vibrating structure," *Proceedings of the 1998 ASME Noise Control and Acoustics Division, NCA-Vol. 25*, pp. 339–346, Anaheim, California, November 1999.

Active control of harmonic sound transmission into an acoustic enclosure using both structural and acoustic actuators

Sang-Myeong Kim^{a)} and Michael J. Brennan^{b)}

Institute of Sound and Vibration Research, University of Southampton, Southampton, Hampshire SO17 1BJ, United Kingdom

(Received 30 September 1998; accepted for publication 27 December 1999)

This paper describes an analytical and experimental investigation into the active control of harmonic sound transmission in a structural–acoustic coupled system. A rectangular enclosure is considered that has five acoustically rigid walls and a flexible plate on the remaining side through which a harmonic sound wave is transmitted into the enclosure. The control system is designed to globally reduce the sound field inside the enclosure, and the roles of structural and acoustic actuators are of particular interest. Three control configurations, classified by the type of actuators, are compared and discussed. They are: (i) use of a single point-force actuator, (ii) use of a single acoustic piston source, and (iii) simultaneous use of both a point-force actuator *and* an acoustic piston source. It is shown both analytically and experimentally that the point-force actuator is effective in controlling plate-dominated modes while the acoustic source is effective in controlling cavity-dominated modes. Since the transmitted sound field is governed by both plate- and cavity-dominated modes, the hybrid use of both types of actuators is shown to be a desirable configuration for the active control of sound transmission into a structural–acoustic coupled system. © 2000 Acoustical Society of America. [S0001-4966(00)04004-2]

PACS numbers: 43.40.Vn [PJR]

INTRODUCTION

Analytical studies of active control in vibro–acoustic systems have been conducted by many investigators with the aim of achieving physical insight so that effective control systems can be designed. Active control has been applied to the control of sound radiation from a plate^{1–4} and the sound transmission into a rectangular enclosure.^{5,6} Meirovitch and Thangjitham⁴ showed that more control actuators generally gives better control effects. Pan *et al.*⁶ used a point-force actuator to control sound transmission into an enclosure, and discussed the control mechanisms in terms of plate- and cavity-dominated modes. Although it is well-known that a point-force actuator and an acoustic piston source can be used to control well-separated vibration and acoustic modes, respectively,^{7,8} it is not clear which is the best actuator to use to actively control sound transmission in a structural–acoustic system.

The work presented here considers a theoretical and experimental investigation into the active control of sound transmission into a structural–acoustic coupled system using both structural and acoustic actuators. The theoretical formulation is an extension of the work presented in Ref. 9, where a theoretical framework for the analysis of structural–acoustic systems was presented. Following a general formulation of active control theory for structural–acoustic coupled systems in Sec. I, it is applied in Sec. II to a rectangular enclosure having one flexible plate through which ex-

ternal noise is transmitted. The acoustic potential energy inside the enclosure is adopted as the cost function to minimize, and perfect knowledge of the acoustic field is assumed. Three control systems classified by the type of actuators are compared and discussed. They are: (i) use of a single point-force actuator, (ii) use of a single acoustic piston source, and (iii) simultaneous use of both a point-force actuator *and* the acoustic piston source. Of particular interest are the roles of each actuator, which are discussed using the concept of “control spillover.”^{2,10} The simultaneous use of both structural and acoustic actuators is shown to be an effective way to control the sound transmission into a structural–acoustic coupled system. In Sec. III the experimental work is presented and the results are compared with the predictions made using the analytical model. Finally, the paper closes with some conclusions in Sec. IV. There is also one appendix which details the equations for the rectangular enclosure used in the simulations and the experimental work.

I. ANALYTICAL MODEL OF A STRUCTURAL–ACOUSTIC COUPLED SYSTEM

A. Model derivation

Consider an arbitrary-shaped enclosure surrounded by a flexible structure and an acoustically rigid wall as shown in Fig. 1. A harmonic plane wave is assumed to be incident on the flexible structure, and wave interference outside the enclosure between the incident wave and the radiated wave from the structure is neglected. Three separate sets of coordinate systems are used; coordinate \mathbf{x} is used for the acoustic field in the cavity, coordinate \mathbf{y} is used for the vibration of the structure, and coordinate \mathbf{r} is used for the sound field outside the enclosure. The cavity acoustic field and the flex-

^{a)}Present address: Mechatronics Dept., Kwang-Ju Institute of Science and Technology, 1 Oryong-dong, Puk-gu, Kwangju 500-712, Republic of Korea.

^{b)}Author to whom correspondence should be addressed; electronic mail: mjb@isvr.soton.ac.uk

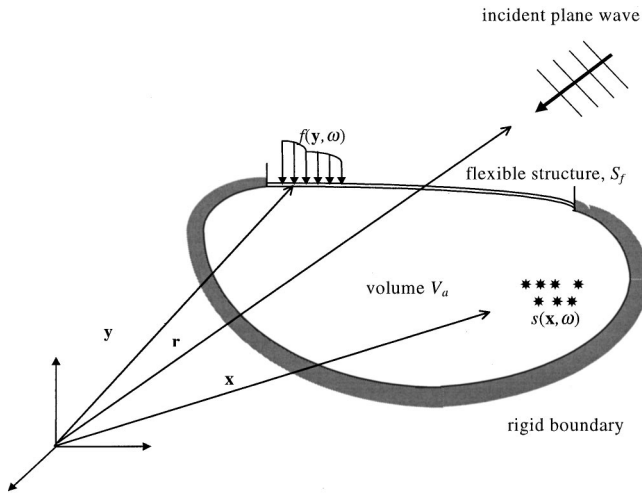


FIG. 1. A structural acoustic coupled system with the volume V_a and its flexible boundary surface of area S_f .

ible structure are assumed to be governed by the linear Helmholtz equation and the isotropic thin-plate theory,¹¹ respectively. The sign of the force distribution function and normal vibration velocity are set to be positive when they direct inward to the cavity, so that the structural contribution to acoustic pressure has the same sign as the acoustic source contribution to acoustic pressure.

It is also assumed that the coupled response of the system can be described by finite summations of the *uncoupled* acoustic and structural modes. The uncoupled modes are the *rigid-walled* acoustic modes of the cavity and the *in vacuo* structural modes of the structure. The acoustic pressure and structural vibration velocity normal to the vibrating surface are chosen to describe the responses of the coupled system.

The time-averaged acoustic potential energy in the cavity is adopted as the cost function for the global sound control, which is given by¹²

$$E_p = \frac{1}{4\rho_o c_o^2} \int_{V_a} |p(\mathbf{x}, \omega)|^2 dV, \quad (1)$$

where ρ_o and c_o , respectively, denote the density and the speed of sound in air, and $p(\mathbf{x}, \omega)$ is the sound pressure inside the cavity of volume V_a . The time-averaged vibration kinetic energy of the flexible structure, which will be used to judge the control effect on structural vibration, is given by⁷

$$E_k = \frac{\rho_s h}{4} \int_{S_f} |u(\mathbf{y}, \omega)|^2 dS, \quad (2)$$

where ρ_s is the density of the plate material and h is the thickness of the plate of area S_f . If the acoustic pressure and the structural vibration are assumed to be described by a summation of N and M modes, respectively, then the acoustic pressure at position \mathbf{x} inside the enclosure and the structural vibration velocity at position \mathbf{y} are given by

$$p(\mathbf{x}, \omega) = \sum_{n=1}^N \psi_n(\mathbf{x}) a_n(\omega) = \Psi^T \mathbf{a}, \quad (3)$$

$$u(\mathbf{y}, \omega) = \sum_{m=1}^M \phi_m(\mathbf{y}) b_m(\omega) = \Phi^T \mathbf{b}, \quad (4)$$

where the N length column vectors Ψ and \mathbf{a} consist of the array of uncoupled acoustic mode shape functions $\psi_n(\mathbf{x})$ and the complex amplitude of the acoustic pressure modes $a_n(\omega)$ respectively. Likewise, the M length column vectors Φ and \mathbf{b} consist of the array of uncoupled vibration mode shape functions $\phi_m(\mathbf{y})$ and the complex amplitude of the vibration velocity modes $b_m(\omega)$, respectively. The superscript T denotes the transpose.

The mode shape functions $\psi_n(\mathbf{x})$ and $\phi_m(\mathbf{y})$ satisfy the orthogonal property in each uncoupled system, and are normalized as follows:

$$V_a = \int_{V_a} \psi_n^2(\mathbf{x}) dV, \quad (5)$$

$$S_f = \int_{S_f} \phi_m^2(\mathbf{y}) dS, \quad (6)$$

where V_a and S_f are the volume of the enclosure and the area of the flexible structure, respectively. Since the acoustic mode shape functions are normalized as given by Eq. (5), the acoustic potential energy can be written as

$$E_p = \frac{V_a}{4\rho_o c_o^2} \mathbf{a}^H \mathbf{a}. \quad (7)$$

Similarly, from Eqs. (2) and (6), the vibration kinetic energy can be written as

$$E_k = \frac{M_s}{4} \mathbf{b}^H \mathbf{b}, \quad (8)$$

where the mass of the flexible plate is $M_s = \rho_s h S_f$, and superscript H denotes the Hermitian transpose.

For the active control of sound, the complex amplitude of acoustic pressure vector \mathbf{a} should be described in terms of excitation variables. The complex amplitude of the n th acoustic mode under structural and acoustic excitation is given by^{13,14}

$$a_n(\omega) = \frac{\rho_o c_o^2}{V_a} A_n(\omega) \left(\int_{V_a} \psi_n(\mathbf{x}) s(\mathbf{x}, \omega) dV + \int_{S_f} \psi_n(\mathbf{y}) u(\mathbf{y}, \omega) dS \right), \quad (9)$$

where $s(\mathbf{x}, \omega)$ denotes the acoustic source strength density function in the cavity volume V_a , and $u(\mathbf{y}, \omega)$ denotes the normal velocity of the surrounding flexible structure on surface S_f . The two integrals inside the brackets represent the n th acoustic modal source strength contributed from $s(\mathbf{x}, \omega)$ and $u(\mathbf{y}, \omega)$, respectively. The acoustic mode resonance term $A_n(\omega)$ is given by

$$A_1(\omega) = \frac{1}{1/T_a + j\omega}, \quad \text{when } n=1, \quad (10a)$$

$$A_n(\omega) = \frac{j\omega}{\omega_n^2 - \omega^2 + j2\zeta_n \omega_n \omega}, \quad \text{when } n \neq 1, \quad (10b)$$

where the time constant of the first acoustic mode T_a is determined from both the acoustic compliance of the cavity and the acoustic resistance due to the rate of air leakage through holes on the cavity walls,⁸ and ω_n and ζ_n are the natural frequency and damping ratio of the n th acoustic mode, respectively.

Substituting Eq. (4) into Eq. (9) and introducing the modal source strength $q_n = \int_{V_a} \psi_n(\mathbf{x}) s(\mathbf{x}, \omega) dV$, then we get

$$a_n(\omega) = \frac{\rho_o c_o^2}{V_a} A_n(\omega) \left(q_n(\omega) + \sum_{m=1}^M C_{n,m} b_m(\omega) \right), \quad (11)$$

where $C_{n,m}$ represents the geometric coupling relationship between the uncoupled structural and acoustic mode shape functions on the surface of the vibrating structure S_f and is given by¹⁴

$$C_{n,m} = \int_{S_f} \psi_n(\mathbf{y}) \phi_m(\mathbf{y}, \omega) dS. \quad (12)$$

If we use L independent acoustic control sources, q_n can be written as

$$\begin{aligned} q_n(\omega) &= \sum_{l=1}^L \frac{1}{S_{q,l}} \int_{V_a} \psi_n(\mathbf{x}_{c,l}) dV q_{c,l}(\omega) \\ &= \sum_{l=1}^L D_{q,nl} q_{c,l}(\omega), \end{aligned} \quad (13)$$

where $D_{q,nl} = (1/S_{q,l}) \int_{V_a} \psi_n(\mathbf{x}_{c,l}) dV$, and the l th control source strength $q_{c,l}(\omega)$ having an area of $S_{q,l}$ is defined at $\mathbf{x}_{c,l}$. Thus, the complex amplitude of acoustic modal pressure vector \mathbf{a} can be expressed as

$$\mathbf{a} = \mathbf{Z}_a (\mathbf{D}_q \mathbf{q}_c + \mathbf{C} \mathbf{b}), \quad (14)$$

where $\mathbf{Z}_a = (\rho_o c_o^2 / V_a) \mathbf{A}$ is the *uncoupled acoustic modal impedance matrix*, which determines the relationship between the acoustic source excitation and the resultant acoustic pressure in modal coordinates of the uncoupled acoustic system.⁹ The matrix \mathbf{A} is an $(N \times N)$ diagonal matrix in which each (n, n) diagonal term consists of A_n given in Eq. (10), the $(N \times M)$ matrix \mathbf{C} is the structural-acoustic mode shape coupling matrix, the $(N \times L)$ matrix \mathbf{D}_q determines geometrical coupling between the L acoustic source locations and the N acoustic modes, the L length vector \mathbf{q}_c is the complex strength vector of acoustic control sources, and \mathbf{b} is the complex vibration modal amplitude vector.

Since the flexible structure in Fig. 1 is assumed to be governed by the isotropic thin-plate theory, the complex vibration velocity amplitude of the m th mode can be expressed as¹³

$$\begin{aligned} b_m(\omega) &= \frac{1}{M_s} B_m(\omega) \left(\int_{S_f} \phi_m(\mathbf{y}) (f(\mathbf{y}, \omega) \right. \\ &\quad \left. + p^{\text{ext}}(\mathbf{y}, \omega) - p(\mathbf{y}, \omega)) dS \right), \end{aligned} \quad (15)$$

where again M_s is the total mass of the plate. Inside the integral $f(\mathbf{y}, \omega)$, $p^{\text{ext}}(\mathbf{y}, \omega)$, and $p(\mathbf{y}, \omega)$ denote the force distribution function, and the exterior and interior acoustic pres-

sure distributions on the surface S_f , respectively. Because of the sign convention used, there is a minus sign in front of $p(\mathbf{y}, \omega)$. The structural mode resonance term $B_m(\omega)$ can be expressed as

$$B_m(\omega) = \frac{j\omega}{\omega_m^2 - \omega^2 + j2\zeta_m \omega_m \omega}, \quad (16)$$

where ω_m and ζ_m are the natural frequency and the damping ratio of m th mode, respectively. Substituting Eq. (3) into Eq. (15), we get

$$\begin{aligned} b_m(\omega) &= \frac{1}{M_s} B_m(\omega) \left(g_{c,m}(\omega) + g_{p,m}(\omega) \right. \\ &\quad \left. - \sum_{n=1}^N C_{n,m}^T a_n(\omega) \right), \end{aligned} \quad (17)$$

where $g_{c,m}(\omega) = \int_{S_f} \phi_m(\mathbf{y}) f(\mathbf{y}, \omega) dS$, $g_{p,m}(\omega) = \int_{S_f} \phi_m(\mathbf{y}) \times p^{\text{ext}}(\mathbf{y}, \omega) dS$, and $C_{n,m}^T = C_{m,n}$. If we use K independent point-force actuators, the m th mode generalized force due to control forces, $g_{c,m}$, can be written by

$$\begin{aligned} g_{c,m}(\omega) &= \sum_{k=1}^K \int_{S_f} \phi_m(\mathbf{y}) \delta(\mathbf{y} - \mathbf{y}_{c,k}) dS f_{c,k}(\omega) \\ &= \sum_{k=1}^K D_{f,mk} f_{c,k}(\omega), \end{aligned} \quad (18)$$

where $D_{f,mk} = \int_{S_f} \phi_m(\mathbf{y}) \delta(\mathbf{y} - \mathbf{y}_{c,k}) dS$, and the k th control point force $f_{c,k}(\omega)$ is located at $\mathbf{y}_{c,k}$. Thus, the modal vibration amplitude vector \mathbf{b} can be expressed as

$$\mathbf{b} = \mathbf{Y}_s (\mathbf{g}_p + \mathbf{D}_f \mathbf{f}_c - \mathbf{C}^T \mathbf{a}), \quad (19)$$

where $\mathbf{Y}_s = (1/M_s) \mathbf{B}$ is the *uncoupled structural modal mobility matrix* which determines the relationship between structural excitation and the resultant structural velocity response in modal coordinates of the uncoupled structural system.⁹ The matrix \mathbf{B} is an $(M \times M)$ diagonal matrix in which each (m, m) diagonal term consists of B_m given in Eq. (16), \mathbf{C}^T is the transpose matrix of \mathbf{C} , the $(N \times K)$ matrix \mathbf{D}_f determines geometrical coupling between the K point-force locations and the M structural modes, \mathbf{g}_p is the generalized modal force vector due to the primary plane wave excitation, the K length vector \mathbf{f}_c is the complex vector of structural control point forces, and \mathbf{a} is the complex acoustic modal amplitude vector.

Combining Eqs. (14) and (19), we get

$$\mathbf{a} = \mathbf{Z} (\mathbf{C} \mathbf{Y}_s \mathbf{g}_p + \mathbf{D}_q \mathbf{q}_c + \mathbf{C} \mathbf{Y}_s \mathbf{D}_f \mathbf{f}_c), \quad (20)$$

$$\mathbf{b} = \mathbf{Y} (\mathbf{g}_p + \mathbf{D}_f \mathbf{f}_c - \mathbf{C}^T \mathbf{Z}_a \mathbf{D}_q \mathbf{q}_c), \quad (21)$$

where $\mathbf{Z} = (\mathbf{I} + \mathbf{Z}_a \mathbf{Y}_{cs})^{-1} \mathbf{Z}_a$ and $\mathbf{Y} = (\mathbf{I} + \mathbf{Y}_s \mathbf{Z}_{ca})^{-1} \mathbf{Y}_s$ where the matrices $\mathbf{Z}_{ca} = \mathbf{C}^T \mathbf{Z}_a \mathbf{C}$ and $\mathbf{Y}_{cs} = \mathbf{C} \mathbf{Y}_s \mathbf{C}^T$ are defined as the *coupled acoustic modal impedance matrix* and the *coupled structural modal mobility matrix*, respectively.⁹ The matrix \mathbf{Z}_{ca} determines the acoustic reaction force induced by structural vibration in modal coordinates, whereas \mathbf{Y}_{cs} determines the induced source strength on the flexible structure induced by acoustic excitation in modal coordinates. They are the result of the coupling between the structure and the cavity as

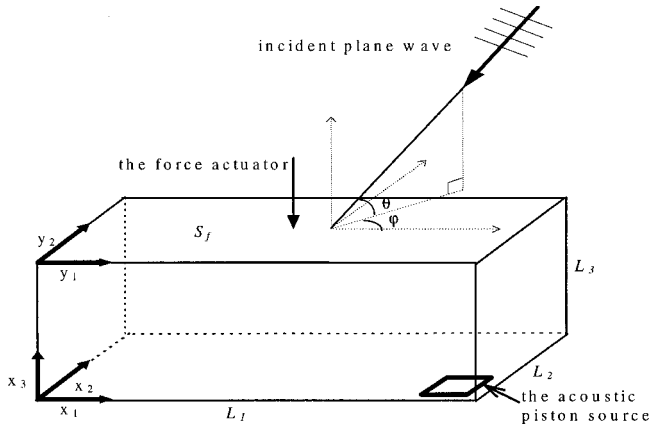


FIG. 2. A rectangular enclosure with one simply supported plate on the surface S_f on which external plane wave is incident with the angles of ($\varphi = 0^\circ$) and ($\theta = 45^\circ$).

the uncoupled impedance and mobility matrices are combined with the matrices \mathbf{C} and \mathbf{C}^T .

Equations (20) and (21) are compact descriptions of the classical theory for calculating the coupled responses in a structural–acoustic coupled system that is excited by both structural and acoustic sources.

B. Minimization of the acoustic potential energy

When both acoustic piston sources and vibration actuators are used to control the acoustic environment in the cavity, Eq. (20) can be rewritten as a function of the total control variable vector \mathbf{t}_c

$$\mathbf{a} = \mathbf{Z}(\mathbf{C}\mathbf{Y}_s\mathbf{g}_p + \mathbf{T}\mathbf{t}_c), \quad (22)$$

where $\mathbf{T} = [\mathbf{C}\mathbf{Y}_s\mathbf{D}_f \ \mathbf{D}_q]$, and

$$\mathbf{t}_c = \begin{Bmatrix} \mathbf{f}_c \\ \mathbf{q}_c \end{Bmatrix}.$$

Substituting Eq. (22) into Eq. (7), we get the acoustic potential energy in terms of the control variables and the primary incident wave excitation, which is given by

$$E_p = \frac{V_a}{4\rho_0 c_0^2} (\mathbf{C}\mathbf{Y}_s\mathbf{g}_p + \mathbf{T}\mathbf{t}_c)^H \mathbf{Z}^H \mathbf{Z} (\mathbf{C}\mathbf{Y}_s\mathbf{g}_p + \mathbf{T}\mathbf{t}_c). \quad (23)$$

Since Eq. (23) is the Hermitian quadratic form of complex variable vector \mathbf{t}_c , the optimal strength of the total control actuator variable vector \mathbf{t}_{co} , which minimizes the Hermitian quadratic form, is given by⁸

$$\mathbf{t}_{co} = \begin{Bmatrix} \mathbf{f}_{co} \\ \mathbf{q}_{co} \end{Bmatrix} = -(\mathbf{T}^H \mathbf{Z}^H \mathbf{Z} \mathbf{T})^{-1} \mathbf{T}^H \mathbf{Z}^H \mathbf{Z} \mathbf{C}\mathbf{Y}_s\mathbf{g}_p. \quad (24)$$

If only structural force actuators are used, the optimal strength of force actuator vector \mathbf{f}_{co} can be obtained by setting $\mathbf{T} = \mathbf{C}\mathbf{Y}_s\mathbf{D}_f$ and $\mathbf{t}_c = \mathbf{f}_c$. Likewise, if only acoustic piston sources are used, the optimal strength of acoustic piston source vector \mathbf{q}_{co} can be obtained by setting $\mathbf{T} = \mathbf{D}_q$ and $\mathbf{t}_c = \mathbf{q}_c$.

TABLE I. Material properties.

Material	Density (kg/m ³)	Phase speed (m/s)	Young's modulus (N/m ²)	Poisson's ratio (ν)	Damping ratio (ζ)
Air	1.21	340	0.01
Steel	7870	...	207×10^9	0.29	0.01
Al	2770	...	71×10^9	0.33	0.01

II. ACTIVE CONTROL OF SOUND TRANSMISSION INTO A RECTANGULAR ENCLOSURE

A. Model description and system characteristics

In this section, we investigate the active control of sound transmission into a rectangular enclosure as shown in Fig. 2. Three coordinates systems, \mathbf{x} , \mathbf{y} and \mathbf{r} , are used for specifying the cavity, the plate and the sound field outside the cavity, respectively. Five rigid walls and a simply supported flexible plate surround the cavity, and the aim is to attenuate the acoustic potential energy inside the cavity. A plane wave of pressure amplitude 1 Pa outside the enclosure is assumed to be incident on the plate with the angles of ($\varphi = 0^\circ$) and ($\theta = 45^\circ$) and is transmitted into the cavity by vibrating the plate. The dimensions of the cavity are $2.0 \times 0.6 \times 0.9$ m, and the thickness of the steel plate is 7 mm. The material properties of air and steel used in the simulations are listed in Table I. The modal damping ratios of the plate and the cavity were assumed to be 0.01.

In order to minimize the transmission of sound into the cavity, two kinds of actuators are used: a single point-force actuator on the plate and a single rectangular-type acoustic piston source on the cavity wall. The rectangular piston source is centered at (1.85, 0.15, 0) with an area of 0.15 by 0.15 m. This location was chosen because the sound pressure of each mode in a rectangular cavity is a maximum at the corners, and thus the control source is placed away from the acoustic nodal planes.¹² The nodal planes can be dramatically changed when the cavity and the structure are strongly coupled. However, the coupling in the enclosure studied is weak, so that the locations are able to excite all cavity-dominated modes of interest. For a similar reason, the point-force actuator is located at $(9/20L_1, L_2/2)$ on the plate, where there are no nodal lines within the frequency range of interest. Table II shows the natural frequencies of each uncoupled system and their geometric mode shape coupling coefficients, which are normalized by their maximum value. Some natural frequencies which are not excited by the given

TABLE II. The natural frequencies and geometric mode shape coupling coefficients of each uncoupled system.

Order	Plate		1	2	3	4	5	7	10
	Type		(1,1)	(2,1)	(3,1)	(4,1)	(5,1)	(6,1)	(7,1)
	Cavity	Freq. (Hz)	52	64	86	115	154	200	256
1	(0,0,0)	0	0.71	0	0.24	0	0.14	0	0.10
2	(1,0,0)	85	0	0.67	0	0.27	0	0.17	0
3	(2,0,0)	170	-0.33	0	0.60	0	0.24	0	0.16
4	(0,0,1)	189	-1.00	0	-0.33	0	-0.20	0	-0.14
5	(1,0,1)	207	0	-0.94	0	-0.38	0	-0.24	0
6	(2,0,1)	254	0.47	0	-0.85	0	-0.34	0	-0.22
7	(3,0,0)	255	0	-0.40	0	0.57	0	0.22	0

incident angle ($\varphi=0^\circ$) were omitted. The (m_1, m_2) and (n_1, n_2, n_3) indicate the indices of the m th plate mode and the n th cavity mode, respectively, and the corresponding uncoupled natural frequencies of the plate and the cavity are listed. A total of 15 structural and 10 acoustic modes was used for the simulations in the frequency range 0–300 Hz, and no significant difference was noticed in simulations with more modes. Equations for a rectangular panel and a rectangular cavity are given in the Appendix.

B. Active minimization of the acoustic potential energy

Three active control strategies classified by the configuration of actuators are considered in this section. They are: (i) use of a single force actuator, (ii) use of a single acoustic piston source, and (iii) simultaneous use of both the force actuator and the acoustic piston source. Although the formulation developed here is not restricted to a single actuator, single structural and acoustic actuators are used to simplify the problem so that the physical limitations of each case can be easily studied and effective guidelines for practical implementation can be suggested.

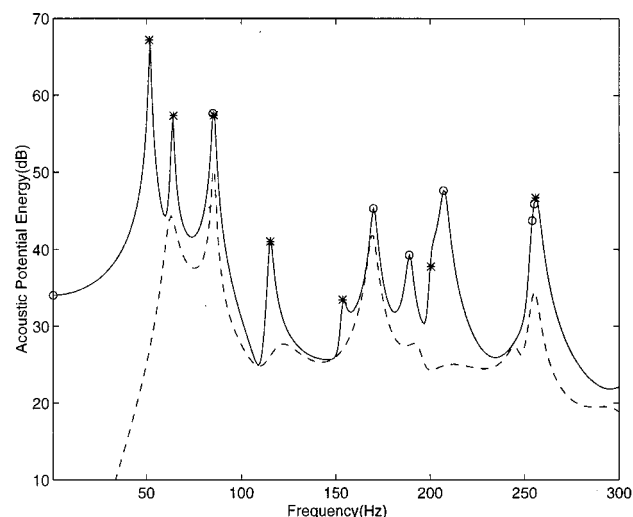
1. Control using a single force actuator

The structural actuator considered was a point-force actuator as shown in Fig. 2. The optimal control force can be calculated using Eq. (24), which can be combined with Eqs. (20) and (7) to give the acoustic potential energy of the cavity after control. This is shown in Fig. 3(a) with the acoustic potential energy without control. To show how this control system affects the vibration of the plate, the time-averaged vibration kinetic energy of the plate obtained from combining Eqs. (8) and (21) is plotted in Fig. 3(b). On each graph, uncoupled natural frequencies of the plate and the cavity are marked “*” and “○,” respectively. Because the coupling between the plate and the cavity is relatively weak, the acoustic potential energy in the uncontrolled state has peaks at both uncoupled plate and cavity natural frequencies. Likewise, the vibration response in the uncontrolled state is governed by the plate modes only.

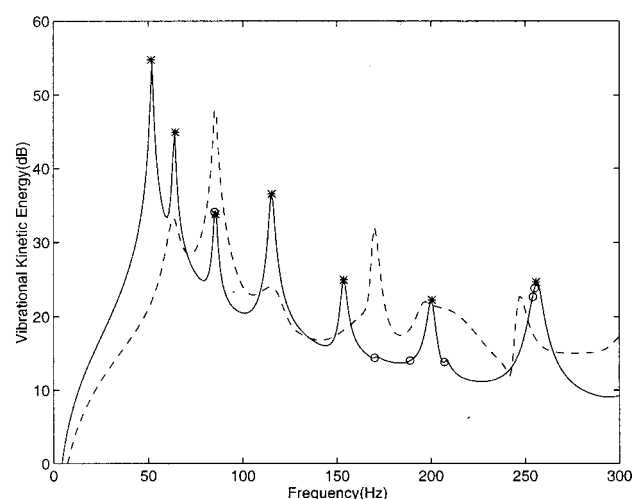
Examining Fig. 3(a) and (b), it can be seen that there is a large reduction of the acoustic potential energy, especially at the 1st, 2nd, 4th, and 5th plate modes corresponding to 52, 64, 115, and 154 Hz, respectively. This is because the sound field at these frequencies is governed by a plate vibration mode, and a single structural actuator is able to control the corresponding vibration mode so as to minimize sound transmission.

The structural actuator reduces the acoustic potential energy at cavity-dominated modes as well (especially the 2nd and 3rd cavity modes corresponding to 85 and 170 Hz, respectively); however, it increases the plate vibration significantly at these frequencies. This demonstrates that minimization of the acoustic potential energy does not always result in a reduction of structural vibration especially at cavity-dominated modes, and vice versa.

The physical limitation of the structural actuator in this case can be explained by the concept of *control spillover*.^{2,10} Control spillover is an inherent characteristic of the use of



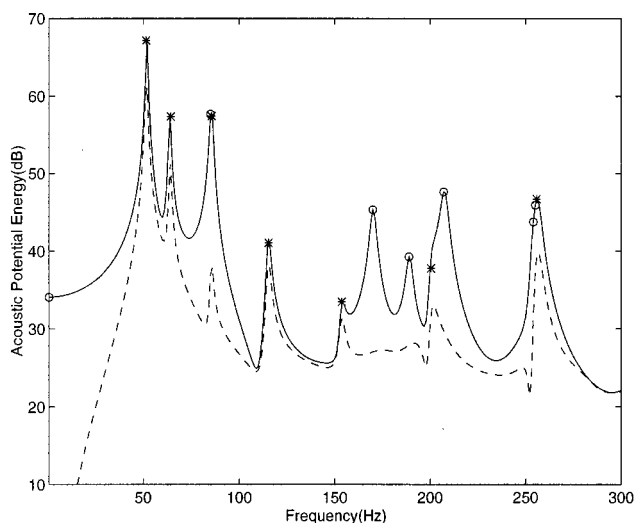
(a) acoustic potential energy of the cavity (dB ref.= 10^{-12} J)



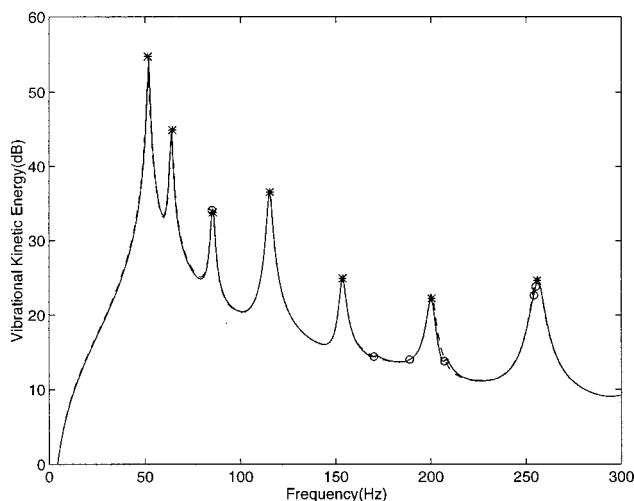
(b) vibration kinetic energy of the plate (dB ref.= 10^{-9} J)

FIG. 3. Effects of minimizing the acoustic potential energy using a point-force actuator (solid line: without control, dashed line: with control), where “*” and “○” are at uncoupled plate and cavity natural frequencies, respectively. (a) acoustic potential energy of the cavity; (b) vibration kinetic energy of the plate.

discrete actuators, which has an undesirable redundant effect. As can be seen from Eq. (A10), for example, a point-force actuator f_c located at (y_{1c}, y_{2c}) couples with every mode unless the location is at the nodal line. Thus, it is impossible to control exactly only one mode by using a discrete actuator because the actuator also excites other modes. To avoid such difficulty, continuous sensors and actuators have been introduced, for example.¹⁵ However, if a system has well separated modes and is dominated by a single mode at frequencies close to a natural frequency, a discrete actuator can be effective. On the other hand, if the response is governed by several modes, a single discrete actuator is not suitable because of control spillover. This can be seen in Fig. 3 at frequencies where a cavity mode dominates the acoustic response. The mechanism involved may be more easily understood by using an analogous mechanical lumped-parameter model of a structural–acoustic coupled system.¹⁶ Thus, for effective control of a cavity-dominated mode, a



(a) acoustic potential energy of the cavity (dB ref.= 10^{-12} J)



(b) vibration kinetic energy of the plate (dB ref.= 10^{-9} J)

FIG. 4. Effects of minimizing the acoustic potential energy using an acoustic piston source (solid line: without control, dashed line: with control), where “*” and “○” are at uncoupled plate and cavity natural frequencies, respectively. (a) Acoustic potential energy of the cavity; (b) vibration kinetic energy of the plate.

large number of structural actuators are required in general. An alternative to this is to use an acoustic piston source as well as a structural actuator. Before this is discussed, the efficacy of a single acoustic piston source is investigated.

2. Control using a single acoustic source

The single acoustic piston source is positioned as shown in Fig. 2, and the optimal control source strength and the energies of the system can be obtained in a similar way to the force actuator case. The acoustic potential energy of the cavity with and without control is shown in Fig. 4(a), and the corresponding vibration kinetic energy of the plate is shown in Fig. 4(b). Since a plate-dominated mode is generally coupled with several cavity modes, the acoustic source is not effective at the plate-dominated modes (especially at 52, 64, 115 Hz). However, it is more able to reduce transmitted sound at the cavity-dominated modes (especially at 0, 85, 170, and 189 Hz) than the structural actuator. Because the

coupling between the structure and the acoustic cavity is relatively weak, the acoustic source does not strongly influence the vibration of the plate and thus there is not much difference in the vibration kinetic energy with and without control. This demonstrates that the acoustic actuator is able to reduce the sound field globally without increasing or decreasing the plate vibration, which may be important in some applications.

Thus, it can be seen that a single structural actuator is effective in controlling sound transmission at frequencies where a plate-dominated mode governs the acoustic potential energy, but at frequencies where a cavity-dominated mode governs, a single acoustic source is more effective.

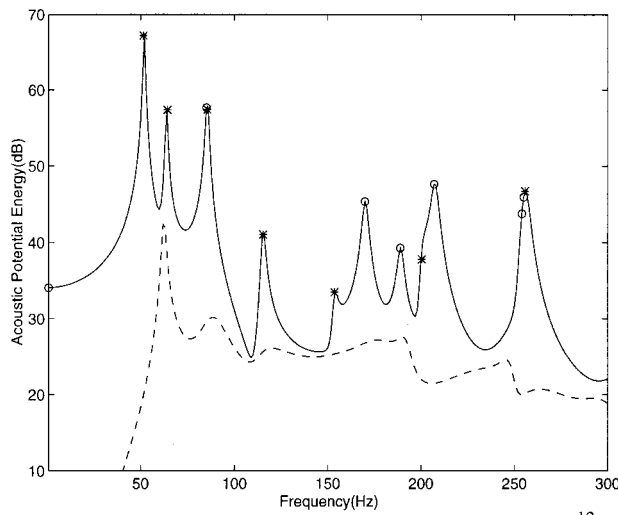
3. Control using both the piston source and the structural actuator

In this section, the effect of using a hybrid system consisting of both the point-force actuator and the acoustic piston source is investigated and compared with the single actuator approaches. As before, the optimal strength of the force actuator and the piston source can be obtained from Eq. (24); using Eqs. (7), (8), (20), and (21), the acoustic potential energy and the vibration kinetic energy with control can be determined. Figure 5(a) shows the acoustic potential energy of the cavity and Fig. 5(b) shows the vibration kinetic energy of the plate with and without the control actuators in operation. The main difference between this configuration and that with the force actuator is that, at frequencies where the cavity modes dominate, there is a large reduction in the acoustic potential energy without a significant increase in structural vibration. In addition, there is an improved performance, i.e., the reduction in acoustic potential energy is greater than when either a single structural or acoustic actuator is used.

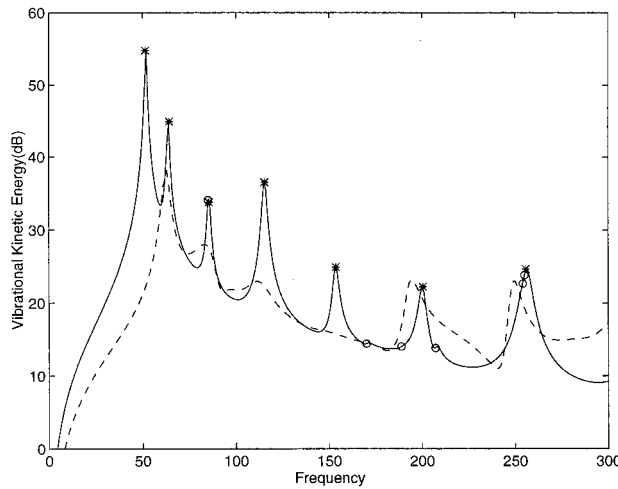
To investigate the control efforts of each control system, the amplitude of the force actuator and the source strength of the piston source are plotted in Fig. 6. When both actuators are used, there is a large decrease of the force amplitude at the well-separated cavity-dominated modes (especially at 85 and 170 Hz). This trend can also be seen in the case of the piston source strength, especially at the first structural natural frequency (52 Hz). By using the hybrid approach, better control effects on the vibration of the plate, the transmission noise reduction, and the control efforts of the actuators can be achieved.

III. EXPERIMENTAL WORK

To validate the analytical model and to support the conclusions from the simulations presented above, some experimental work was conducted. This is described in this section and it is compared with some predictions. Although real-time active control was not implemented, predictions of what could be achieved using such a system are made using measured frequency-response functions. Because we are interested in the active control of harmonic disturbances, there is not an issue of causality and the physical effects of control can be studied without including control issues.



(a) Acoustic potential energy of the cavity (dB ref.= 10^{-12} J)



(b) vibration kinetic energy of the plate (dB ref.= 10^{-9} J)

FIG. 5. Effects of minimizing the acoustic potential energy using both a point-force actuator and an acoustic piston source (solid line: without control, dashed line: with control), where “*” and “O” are at uncoupled plate and cavity natural frequencies, respectively. (a) acoustic potential energy of the cavity; (b) vibration kinetic energy of the plate.

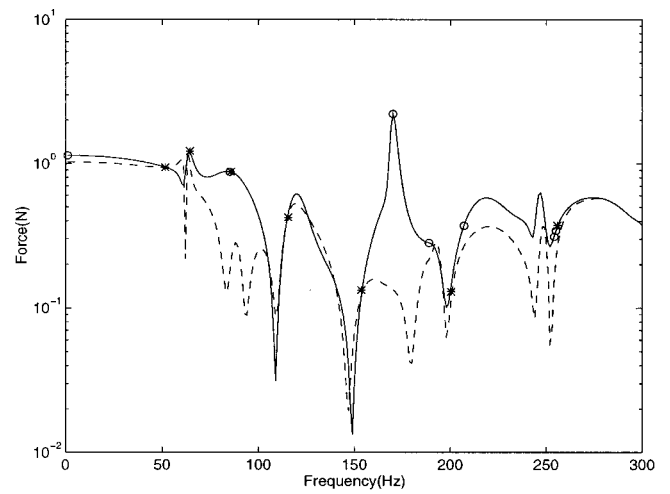
A. Optimal minimization of the approximate of acoustic potential energy

In practice, modal sensors measuring acoustic modal amplitudes in the previous section are not available. An alternative approach using microphones to measure acoustic pressure fluctuations at a discrete number of sensor locations was adopted, and the sum of the square pressure amplitudes at these locations was used as the cost function. The approximation of acoustic potential energy is given by¹²

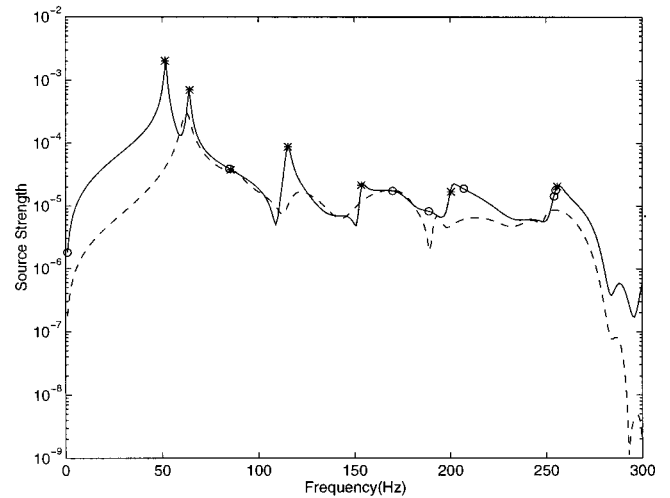
$$\hat{E}_p = \frac{V_a}{4\rho_o c_o^2 L} \sum_{n=1}^L |p(\mathbf{x}_l, \omega)|^2, \quad (25)$$

where $p(\mathbf{x}_l, \omega)$ is the complex pressure amplitude at the l th sensor location. Equation (25) can be written as

$$\hat{E}_p = \frac{V_a}{4\rho_o c_o^2 L} \mathbf{p}^H \mathbf{p}, \quad (26)$$



(a) strength of the force actuator



(b) strength of the piston source (unit: m^3/s)

FIG. 6. Comparison of control efforts of the three control strategies; using each actuator separately (solid line) and using both the force actuator and the piston source (dashed line), where “*” and “O” are at uncoupled plate and cavity natural frequencies, respectively. (a) strength of the force actuator; (b) strength of the piston source (unit: m^3/s).

where \mathbf{p} is the L -length vector whose l th component is $p(\mathbf{x}_l, \omega)$, and is given by¹²

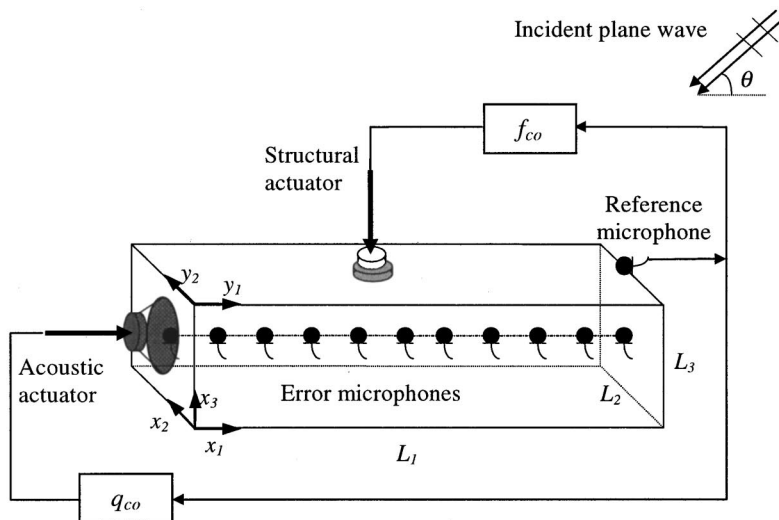
$$\mathbf{p} = \Psi_L^T \mathbf{a}, \quad (27a)$$

where the $(L \times N)$ -size matrix Ψ_L^T contains the N modal amplitudes at L sensor locations. Substituting Eq. (20) into Eq. (27a) we get

$$\mathbf{p} = \Psi_L^T \mathbf{Z} (\mathbf{C} \mathbf{Y}_s \mathbf{g}_p + \mathbf{D}_q \mathbf{q}_c + \mathbf{C} \mathbf{Y}_s \mathbf{D}_f \mathbf{f}_c). \quad (27b)$$

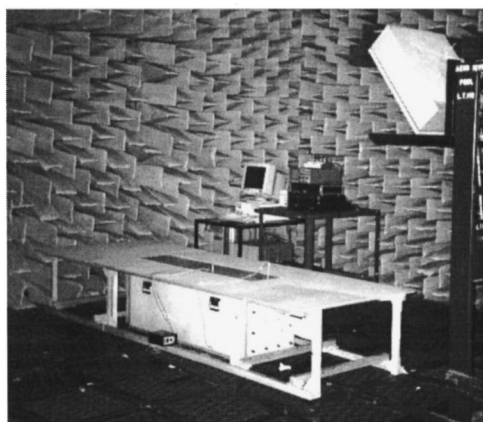
The first term $\Psi_L^T \mathbf{Z} \mathbf{C} \mathbf{Y}_s \mathbf{g}_p$ represents the primary plant response, and the second and third terms $\Psi_L^T \mathbf{Z} \mathbf{D}_q$ and $\Psi_L^T \mathbf{Z} \mathbf{C} \mathbf{Y}_s \mathbf{D}_f$ represent the frequency-response functions of the acoustic and structural secondary plants subject to the acoustic and structural actuators \mathbf{q}_c and \mathbf{f}_c , respectively. The three terms in Eq. (27b) can be obtained either theoretically for simple systems or experimentally for general systems.

By combining Eqs. (22) and (27b), the approximation of acoustic potential energy in Eq. (26) is again a Hermitian

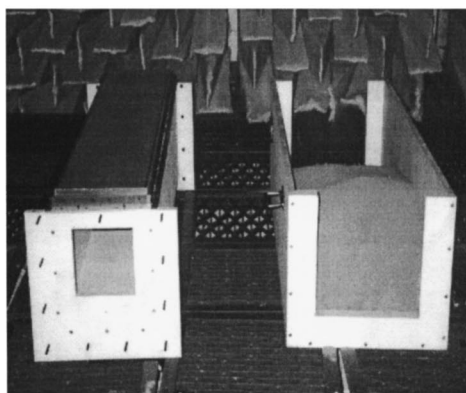


(a) Schematic diagram of the experiment

FIG. 7. Experimental setup. (a) Schematic diagram of the experiment; (b) experimental setup; (c) the rectangular enclosure.



(b) Experimental set-up



(c) The rectangular enclosure

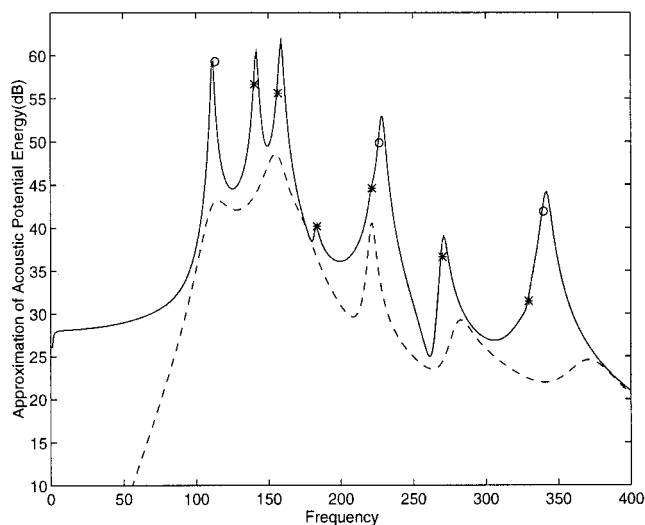
quadratic form of complex variable vector \mathbf{t}_c . The solution can be obtained from Eq. (24) by setting $\mathbf{Z} = \Psi_L^T (\mathbf{I} + \mathbf{Z}_a \mathbf{Y}_{cs})^{-1} \mathbf{Z}_a$, instead of $\mathbf{Z} = (\mathbf{I} + \mathbf{Z}_a \mathbf{Y}_{cs})^{-1} \mathbf{Z}_a$, which was used for modal sensors.

B. Experimental setup

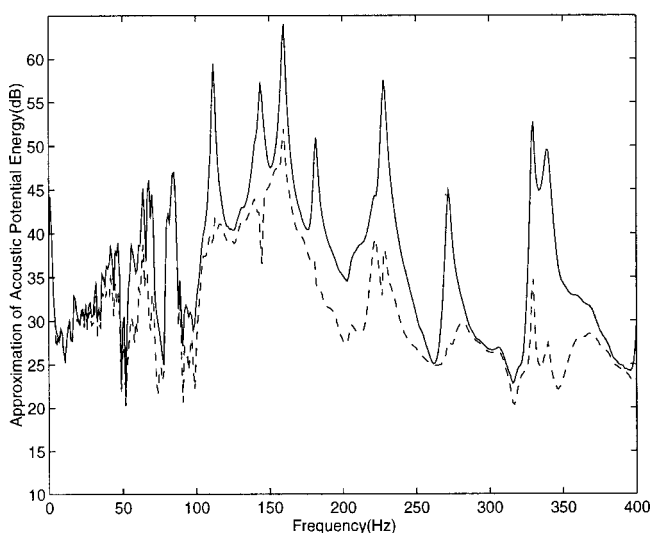
An experiment was performed using a duct-like rectangular enclosure where the acoustic field inside the cavity was one-dimensional along the direction of the duct within the frequency range of interest (up to 400 Hz). The duct-like enclosure was different from the general model used in simulations described in the previous section, and was chosen for the convenience of experimental work, particularly, reducing the number of microphones that was required to approximate the acoustic potential energy. A schematic diagram of the experiment is shown in Fig. 7(a) where the controllers f_{co} and q_{co} were used to globally minimize the approximation of the acoustic potential energy measured by the 11 microphones located equidistant along the center line of the duct. The duct-like rectangular enclosure had five acoustically rigid walls on its base and sides and a simply supported flexible plate on top. On the flexible plate, an incident plane wave was generated with angles of $(\varphi = 0^\circ)$ and $(\theta = 45^\circ)$, where φ is the angle from the coordinate \mathbf{y}_1 on the plate θ is

the angle of incidence from the plate. The dimensions of the cavity were $L_1 \times L_2 \times L_3$, where $L_1 = 1.5$ m, $L_2 = 0.3$ m, and $L_3 = 0.4$ m, and the thickness of the aluminum plate was 5 mm. The reference signal required for feed-forward control was measured by a microphone located at the right-hand edge of the aluminum plate. A loudspeaker of radius 0.15 m was installed at the left end with its center at $(0, L_2/2, L_3/2)$, and an inertial actuator was used as a point-force actuator and was located at $(9L_1/20, L_2/2)$ on the plate.

The experiment was set up in an anechoic chamber whose cutoff frequency is approximately 100 Hz, and is shown in Fig. 7(b). An elevated large speaker was used to generate the external acoustic incident wave, and the top plate of the enclosure was fitted inside an acoustic baffle made of a wood panel of area 3.5×0.9 m with thickness 12 mm. To make the acoustically rigid boundary condition, the walls were made of 25-mm-thick plywood which were surrounded by 75-mm-deep sand layers packed by an extra container as shown in Fig. 7(c). To make the simply supported boundary condition for the plate, steel strips with 1.25-mm thickness were bolted around the peripheral of the plate. The design utilizes the characteristics of the thin strip that is relatively rigid to in-plane motions but flexible to rotation. Some preliminary experiments were performed by Kim,¹⁷ and the



(a) theory



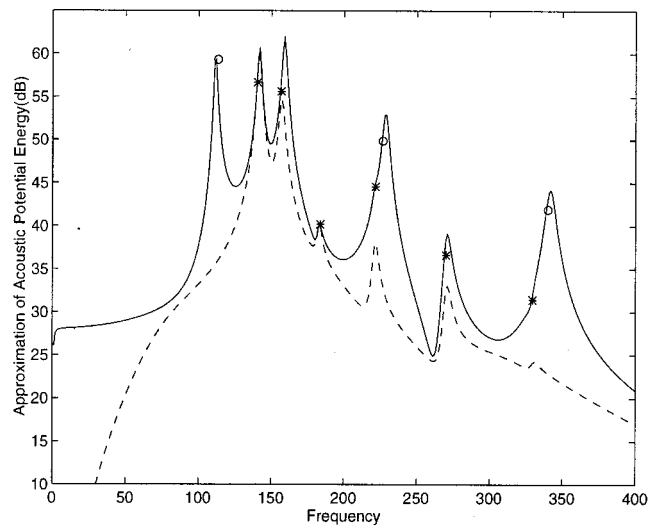
(b) experiment

FIG. 8. The control performance of using the structural actuator; the original cost function (solid), and after control (dashed). (a) Theory; (b) experiment.

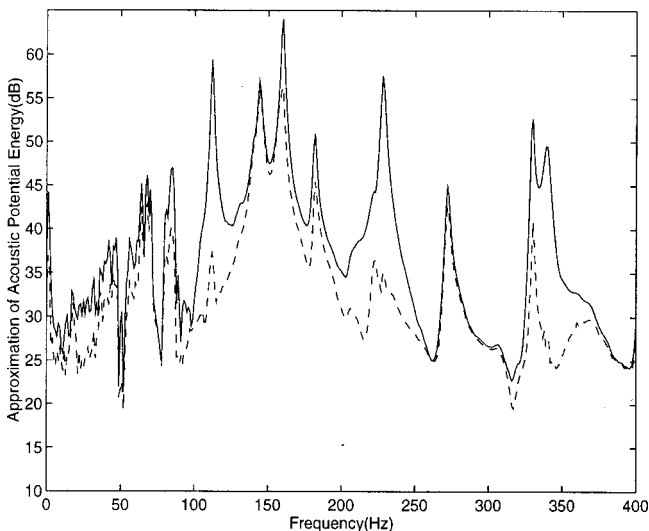
results showed that the experimental model had relatively good simply supported boundary conditions.

C. Experimental results

The experimental work was mainly to identify the frequency-response functions of the primary and secondary plants given in Eq. (27b). The primary plant was identified using the error microphone responses inside cavity with respect to the incident wave measured by a microphone positioned outside the cavity on the plate. The acoustic and structural secondary plants were identified by the error microphone responses to the source strength of the speaker and the force of the inertial actuator, respectively. Low-passed random noise with a cutoff frequency at approximately 300 Hz was used to excite the three actuators for the system identification of the plants: the external speaker for the primary plant, the internal speaker for the acoustic secondary plant, and the force actuator for the structural secondary plant. A low-passed random signal was used to reduce



(a) theory

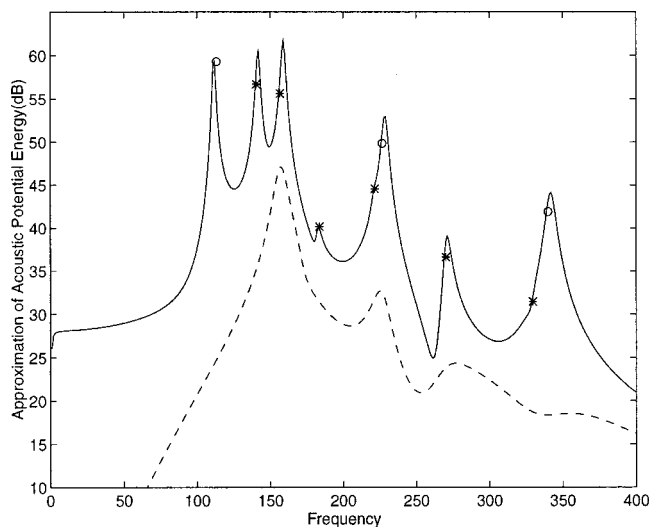


(b) experiment

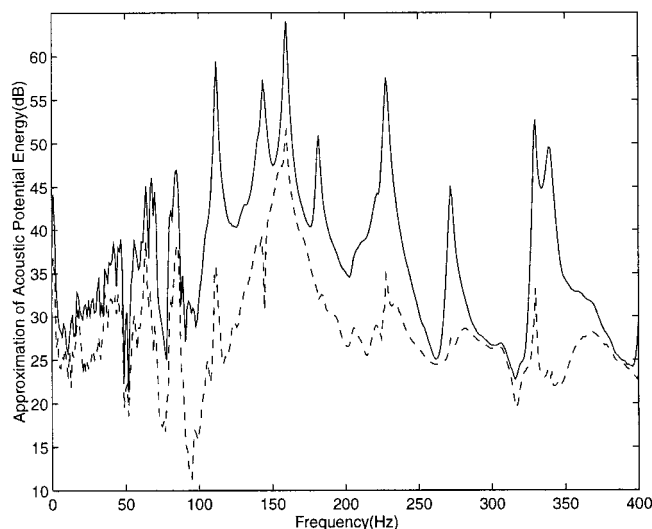
FIG. 9. The control performance of using the acoustic actuator; the original cost function (solid), and after control (dashed). (a) Theory; (b) experiment.

the excitation of higher frequency modes. Following the solution procedure described in Sec. III A, the optimal force and source strength of the actuators were calculated both from the theoretical model and experimental data. The material properties of air and the aluminum plate are given in Table I, and the time constant of the first acoustic mode used in the simulations was assumed to be 0.2 s. A total of four acoustic and six structural modes was assumed to contribute to the coupled responses within the frequency range of interest. Simulations with the colinear microphone array showed a good approximation of the acoustic potential energy under the frequency range of interest.¹⁷

The simulated and experimental control performances for a single secondary force actuator are shown in Fig. 8(a) and (b). The cost function was the approximation of acoustic potential energy as discussed in Sec. III A. The noisy experimental results below about 100 Hz were due to poor coherence between the external and internal microphones. Furthermore, due to the physical difficulty of realizing an acoustic plane wave at a fixed incident angle of $\theta = 45^\circ$, the relative



(a) theory



(b) experiment

FIG. 10. The control performance of using both actuators; the original cost function (solid), and after control (dashed). (a) Theory; (b) experiment.

amplitudes of the peaks are slightly different from the predicted results, especially at the 3rd (180.5 Hz) and 6th (328.5 Hz) plate-dominated modes. Apart from this, the experimental results agree reasonably well with the predictions. The peaks in each graph correspond to the plate- and cavity-dominated modes, and the uncoupled natural frequencies of the plate and the cavity are marked “*” and “O,” respectively in the simulations. It can be seen that, due to the structural–acoustic coupling, the peaks have shifted slightly away from the uncoupled natural frequencies of the plate and the cavity. The results in Fig. 8(a) and (b) demonstrate that the structural actuator is effective in controlling plate-dominated modes (especially at 143, 160, 180.5, and 271 Hz). It also offers large reductions at the cavity-dominated modes in this particular model studied, but as mentioned in Sec. II this is achieved at the cost of increased vibration of the plate.

With the speaker used to control the sound transmission, the acoustic potential energy for the simulation and experi-

ment are shown in Fig. 9(a) and (b), respectively. Both simulation and experimental results agree well each other, and demonstrate that the acoustic actuator is effective in controlling cavity-dominated modes (especially at 113 and 227.5 Hz). Figure 10(a) and (b) show the control performance when both the force actuator and the speaker were used for the control. Again, the simulation and experimental results are in good agreement, showing large reductions in the acoustic potential energy at both plate- and cavity-dominated modes. Although vibration kinetic energy of the plate was not measured, simulation results of the vibration kinetic energy, and indeed the control effort for each case, showed similar trends to that described in Sec. II.

IV. CONCLUSIONS

The active control of the transmission of harmonic sound into a coupled structural–acoustic system has been considered. The roles of each actuator have been of particular interest throughout the work, and been discussed using the concept of “control spillover.” The results obtained both theoretically and experimentally demonstrate that a single point-force actuator is effective in controlling well-separated plate-dominated modes, while a single acoustic piston source is effective in controlling well-separated cavity-dominated modes. Since the transmitted sound field is governed by both plate- and cavity-dominated modes, the hybrid use of both types of actuators is a desirable configuration for the active control of sound transmission into a structural–acoustic coupled system. By using the hybrid approach, improved control effects on the plate vibration, further reduction in sound transmission, and reduced control efforts of the actuators can be achieved.

APPENDIX: EQUATIONS FOR THE ANALYSIS OF A RECTANGULAR ENCLOSURE

For a rigid-walled rectangular enclosure, the acoustic (n_1, n_2, n_3) mode shape functions normalized as given by Eq. (5) is given by:¹²

$$\psi_n(\mathbf{x}) = \sqrt{e_1 e_2 e_3} \cos\left(\frac{n_1 \pi x_1}{L_1}\right) \cos\left(\frac{n_2 \pi x_2}{L_2}\right) \cos\left(\frac{n_3 \pi x_3}{L_3}\right), \quad (\text{A1})$$

where n_1, n_2 , and n_3 are integers and L_1, L_2 , and L_3 are the dimensions of the rectangular enclosure in the $\mathbf{x}_1, \mathbf{x}_2$, and \mathbf{x}_3 coordinate directions. The normalization factors are given by $e_i = 1$ if $n_i = 0$ and $e_i = 2$ if $n_i \geq 1$ where the sub-script i can be 1, 2, and 3. The corresponding acoustic natural frequency is given by

$$\omega_n = \pi c_o \sqrt{\left(\frac{n_1}{L_1}\right)^2 + \left(\frac{n_2}{L_2}\right)^2 + \left(\frac{n_3}{L_3}\right)^2}, \quad (\text{A2})$$

where c_o is the speed of sound in air. For a simply supported isotropic rectangular plate of dimensions $(L_1 \times L_2)$, the plate (m_1, m_2) mode shape function normalized by Eq. (6) can be written as

$$\phi_m(\mathbf{y}) = 2 \sin\left(\frac{m_1 \pi y_1}{L_1}\right) \sin\left(\frac{m_2 \pi y_2}{L_2}\right), \quad (\text{A3})$$

where m_1 and m_2 are positive integers. The corresponding structural natural frequency is given by

$$\omega_m = \left(\frac{D}{\rho_s h}\right)^{1/2} \left[\left(\frac{m_1 \pi}{L_1}\right)^2 + \left(\frac{m_2 \pi}{L_2}\right)^2 \right], \quad (\text{A4})$$

where ρ_s is the density of the plate material, and the bending stiffness D is given by

$$D = \frac{Eh^3}{12(1-\nu^2)}, \quad (\text{A5})$$

where E is Young's modulus, ν is Poisson's ratio of plate, and h is the thickness of the plate. For the rectangular enclosure with the simply supported flexible plate wall, the coupling coefficient $C_{n,m}$ between the n th acoustic mode (n_1, n_2, n_3) and m th structural mode (m_1, m_2) is given by⁵

$$C_{n,m} = \begin{cases} -1^{n_3} 2S_f \sqrt{e_1 e_2 e_3} \frac{m_1 m_2 (-1^{n_1+m_1}-1)(-1^{n_2+m_2}-1)}{\pi^2 (n_1^2 - m_1^2)(n_2^2 - m_2^2)}, & n_1 \neq m_1, n_2 \neq m_2, \\ 0, & \text{otherwise.} \end{cases} \quad (\text{A6})$$

An incident plane wave can be considered as the primary structural excitation for the noise transmission problem. The incident plane wave outside the enclosure may be described as

$$p_i(\mathbf{r}, t) = P_i e^{j(\omega t - \mathbf{k} \cdot \mathbf{r})}, \quad (\text{A7})$$

where \mathbf{k} is a wave number vector, \mathbf{r} is the location vector of the observation point, and P_i is a complex amplitude of the incident plane wave. When a harmonic plane wave is incident on the simply supported plate on a baffle, the generalized force on the m th structural mode due to the incident plane wave is required for the analysis. If weak coupling is assumed between the structure and the external acoustic field, the acoustic pressure term radiated from structural vibration can be neglected. Thus, the generalized force is given by¹⁸

$$g_{p,m} = 4S_f P_i I_{m_1} I_{m_2}, \quad (\text{A8})$$

where I_{m_1} and I_{m_2} are due to the geometric coupling between the plane wave and the m th mode, and given by

$$I_{m_1} = \frac{m_1 \pi [1 - (-1)^{m_1} e^{-j \sin \theta \cos \varphi (\omega L_1 / c_o)}]}{[m_1 \pi]^2 - [\sin \theta \cos \varphi (\omega L_1 / c_o)]^2} \quad (\text{A9a})$$

$$I_{m_1} = (j/2) \text{sgn}(\sin \theta \cos \varphi)$$

if

$$m_1 \pi = \pm \sin \theta \cos \varphi (\omega L_1 / c_o)$$

$$I_{m_2} = \frac{m_2 \pi [1 - (-1)^{m_2} e^{-j \sin \theta \sin \varphi (\omega L_2 / c_o)}]}{[m_2 \pi]^2 - [\sin \theta \sin \varphi (\omega L_2 / c_o)]^2}$$

$$I_{m_2} = (j/2) \text{sgn}(\sin \theta \sin \varphi) \quad (\text{A9b})$$

if

$$m_2 \pi = \pm \sin \theta \sin \varphi (\omega L_2 / c_o).$$

If a point-force actuator f_c is used as a control actuator on the simply supported rectangular plate of dimensions ($L_1 \times L_2$), the generalized force is given by¹

$$D_{f,m} = 2 \sin\left(\frac{m_1 \pi y_{1c}}{L_1}\right) \sin\left(\frac{m_2 \pi y_{2c}}{L_2}\right), \quad (\text{A10})$$

where the location (y_{1c}, y_{2c}) is the position of the point-force actuator.

¹T. Wang, C. R. Fuller, and K. Dimitriadis, "Active control of noise transmission through rectangular plates using multiple piezoelectric or point force actuators," J. Acoust. Soc. Am. **90**, 2820–2830 (1991).

²M. E. Johnson and S. J. Elliott, "Active control of sound radiation using volume velocity cancellation," J. Acoust. Soc. Am. **98**, 2174–2186 (1995).

³C. R. Fuller, C. H. Hansen, and S. D. Snyder, "Active control of sound radiation from a vibrating rectangular panel by sound sources and vibration inputs: An experimental comparison," J. Sound Vib. **145**, 195–215 (1991).

⁴L. Meirovitch and S. Thangjitham, "Active control of sound radiation pressure" Trans. ASME, J. Vib. Acoust. **112**, 237–244 (1990).

⁵S. D. Snyder and N. Tanaka, "On feedforward active control of sound and vibration using vibration error signals," J. Acoust. Soc. Am. **94**, 2181–2193 (1993).

⁶J. Pan, C. H. Hansen, and D. A. Bies, "Active control of noise transmission through a panel into a cavity. I. Analytical study," J. Acoust. Soc. Am. **87**, 2098–2108 (1990).

⁷C. R. Fuller, S. J. Elliott, and P. A. Nelson, *Active Control of Vibration* (Academic, New York, 1996).

⁸P. A. Nelson and S. J. Elliott, *Active Control of Sound* (Academic, New York, 1992).

⁹S. M. Kim and M. J. Brennan, "A compact matrix formulation using the impedance and mobility approach for the analysis of structural-acoustic systems" J. Sound Vib. **233**, 97–113 (1999).

¹⁰M. J. Balas, "Feedback control of flexible systems," IEEE Trans. Autom. Control. **23**(4), 673–679 (1978).

¹¹P. M. Morse and K. U. Ingard, *Theoretical Acoustics* (McGraw-Hill, New York, 1968), pp. 554–576.

¹²A. J. Bullmore, P. A. Nelson, A. R. D. Curtis, and S. J. Elliott, "The active minimization of harmonic enclosed sound fields. Part II. A computer simulation" J. Sound Vib. **117**, 15–33 (1987).

¹³E. H. Dowell, G. F. Gorman III, and D. A. Smith, "Acoustoelasticity: general theory, acoustic modes and forced response to sinusoidal excitation, including comparisons with experiment," J. Sound Vib. **52**(4), 519–542 (1977).

¹⁴F. Fahy, *Sound and Structural Vibration, Radiation, Transmission and Response* (Academic, New York, 1985).

- ¹⁵C. K. Lee and F. C. Moon, "Modal sensors/actuators," ASME J. Appl. Mech. **57**, 434–441 (1990).
- ¹⁶S. M. Kim and M. J. Brennan, "Modelling a structural-acoustic coupled systems with an equivalent lumped parameter mechanical systems," Trans. ASME, J. Vib. Acoust. **121**, 453–459 (1999).
- ¹⁷S. M. Kim, Ph.D. thesis, University of Southampton, "Active control of sound in structural-acoustic systems" (1998).
- ¹⁸L. A. Roussos, NASA Technical Report 2398, "Noise transmission loss of a rectangular plate in an infinite baffle" (1985).

Effects on sleep disturbance of changes in aircraft noise near three airports

Sanford Fidell,^{a)} Karl Pearsons, Barbara G. Tabachnick, and Richard Howe
BBN Technologies, A Unit of GTE Internetworking, 21128 Vanowen Street, Canoga Park, California 91303

(Received 19 August 1998; revised 20 January 2000; accepted 26 January 2000)

Field measurements were conducted of potential sleep disturbance associated with changes in nighttime aircraft noise exposure near three airports. One study was conducted near Stapleton International Airport (DEN) and Denver International Airport (DIA) in anticipation of the closure of the former and opening of the latter. Sleep behavior was monitored in 57 homes located near runway ends at the two airports. A second study was conducted in the vicinity of DeKalb-Peachtree Airport (PDK), a large general aviation airport that expected increased nighttime flight operations due to the Olympic Games in July and August of 1996. Similar methods of measuring nighttime noise levels and sleep disturbance in the two studies were maintained over the course of 2717 and 686 subject-nights of observations, respectively. No major differences in noise-induced sleep disturbance were observed as a function of changes in nighttime aircraft noise exposure. © 2000 Acoustical Society of America. [S0001-4966(00)01005-5]

PACS numbers: 43.50.Qp [MRS]

INTRODUCTION

The prevalence of aircraft noise-induced sleep disturbance in airport communities remains a matter of considerable interest to regulatory, advisory, and standards agencies, since a fully credible dosage-effect relationship useful for assessing nighttime aircraft noise effects on residential populations remains elusive. The findings of recent large-scale field studies (e.g., those of Ollerhead *et al.*, 1992; and of Fidell *et al.*, 1995a) cast doubt on the likelihood of adverse public health consequences of familiar nighttime aircraft noise exposure of adapted populations residing in airport neighborhoods. They also encourage controversy about definitions of sleep disturbance, data analysis methods, and the reconciliation of findings of laboratory and field measurements of noise-induced sleep disturbance.

Little of what is known about the prevalence of noise-induced sleep disturbance in airport neighborhoods (cf. Pearsons *et al.*, 1995) addresses the time course of response to changes in nighttime aircraft noise exposure. The recent studies of Ollerhead *et al.* (1992) and of Fidell *et al.* (1995a) were made under generally stable aircraft noise exposure conditions among airfield-vicinity residents familiar with both their sleeping quarters and routine neighborhood noise sources. The paucity of information about adaptation to changes in nighttime noise is due in large part to the rarity of appropriate settings for field study of habituation to major changes in aircraft noise-induced sleep disturbance.

The primary goal of the present studies was to document potential changes in sleep disturbance associated with changes in aircraft noise levels. The first study began in January 1994 in anticipation of changes in aircraft noise exposure associated with the (expected) imminent closure of Stapleton International Airport (DEN) and the opening of the newly constructed Denver International Airport (DIA). Sev-

eral postponements of the opening of DIA eventually led to four rounds of data collection. The second study was conducted over a 6-week period in residences near DeKalb-Peachtree Airport (PDK). Data collection took place during the 18 days prior to the Atlanta Olympic Games, the 17 days during the Olympics, and 1 week after the closing ceremonies.¹

I. METHOD

Data collection methods for both studies were similar to those described by Fidell *et al.* (1995a). Except as otherwise noted below, instrumentation was employed as shown in Fig. 2 of Fidell *et al.* (1995a) to simultaneously monitor indoor and outdoor noise events. As described by Fidell *et al.* (1995a), a noise event was defined as a sequence of noise levels that began when an A-weighted threshold level was exceeded for at least 2 s and ended when the noise level dropped more than 2 dB below the threshold. The threshold was set on a site-specific basis to maximize collection of noise intrusions (primarily aircraft) in the presence of indoor ambient noise, without exhausting the storage capacity of the noise monitors too rapidly.

Each test participant used a push button attached by a cable to a palmtop computer at bedside for behavioral confirmation of awakening. Test participants were instructed to push the button upon awakening for any reason whatsoever during the night.

A. Study 1: Observations of noise-induced sleep disturbance near Stapleton International and Denver International Airports

1. Study sites and data collection schedules

Observations of noise exposure and sleep disturbance were made in single-family detached homes near Stapleton International (DEN) and Denver International (DIA) during

^{a)}Electronic mail: fidell@bbn.com

TABLE I. Summary of data collection conditions in studies 1 and 2.

Site	Data collection round	Number of homes	Number of test participants	Number of subject-nights of data collection
Study 1				
DEN	Before closure (February/March 1994)	15	30	677
DIA	Before opening (April/May 1994)	14	29	712
DIA	Spanning opening (February/March 1995)	13	30	848
DEN	After closure (April 1995)	15	28	480
	Total	57 (38 different homes)	117 (77 different people)	2717
Study 2				
PDK	Before Olympics (2–16 July 1996)			294
PDK	During Olympics (17 July–4 August 1996)	12	22	295
PDK	After Olympics (5–11 August 1996)			97
	Total	12	22	686

four rounds of data collection. Preference in selection of data collection sites was given, to the extent possible, to homes with greater proximity to runways.

The initial data collection plans were based on the announced closing of DEN in March 1994. Data collection began in two neighborhoods to the immediate south and east of DEN 2 weeks prior to the announced closing date. Data collection continued for an additional 2 weeks after the closing data had been postponed. Noise exposure and sleep disturbance were measured in 15 residences located across residential streets from the airport fence to the south and east of Runway 26 L.

A residential neighborhood as close as possible to DIA was selected for a second round of data collection after the opening date for DIA was first postponed. This neighborhood was located approximately 4 km from the north end of Runway 34.² Data collection began 3 weeks before the second announced opening date (15 May 1994) and continued for an additional 2 weeks after the opening was once again postponed. Noise exposure and sleep disturbance were measured in 14 test participants' homes near DIA. Test participants were paid adult volunteers in good health who responded to canvassing letters.

A third round of data collection was conducted in the same neighborhood as the second round of data collection shortly before the final announced opening date of DIA, 28 February 1995. Data were collected from a total of 30 test participants in 13 residences, including several new participants. A fourth round of data collection began during the first week of April 1995 at homes in two neighborhoods near DEN. Data were collected from a total of 15 residences and 28 participants for a period of 3 weeks.

2. Noise measurements

Indoor noise measurements were made continuously during the first two rounds of data collection for the 4-week data collection period. Noise-monitoring instrumentation³ inside test participants' sleeping quarters in the first round of data collection recorded average sound levels⁴ every 60 s, as well as 1-s time histories of noise events. Outdoor noise measurements were made with five unattended noise monitors in the vicinity of test participants' residences programmed with the same noise event classification parameters as indoor noise monitors. The noise-monitoring equipment was reprogrammed in the last two rounds of data collection to record average sound levels every 2 s.

3. Definition of noise events

An outdoor noise event in the first two rounds of data collection was considered to have occurred when the noise level exceeded 70 dB for at least 2 s. An indoor noise event was considered to have occurred when the noise level exceeded 60 dB for at least 2 s. During the last two rounds of data collection, an outdoor noise event was considered to have occurred when the noise level exceeded 60 dB for at least 2 s, whereas an indoor noise event was considered to have occurred when the noise level exceeded 50 dB for at least 2 s.

Since both indoor and outdoor noise levels were monitored by unattended instrumentation, their sources are not known with certainty. Outdoor microphones were located in such a manner, and noise event classification criteria were so set, however, that the great majority of outdoor noise events was almost certainly aircraft. Many of the indoor noise

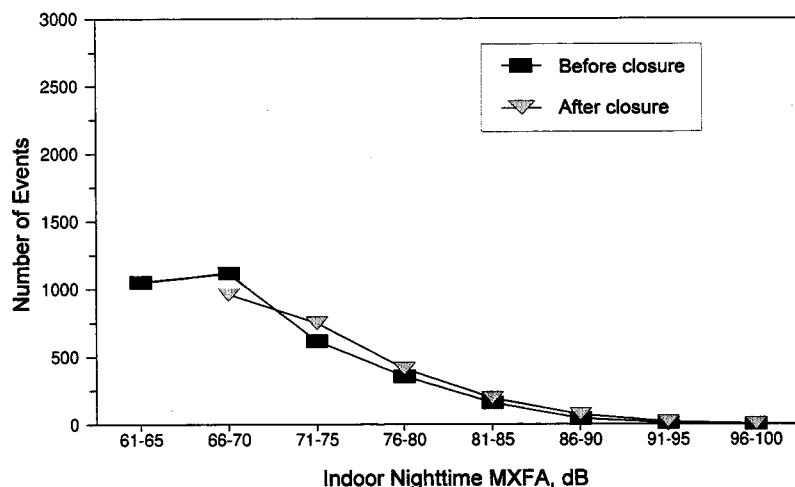


FIG. 1. Distribution of noise events recorded inside test participants' sleeping quarters at DEN from 2200 to 0700 hours.

events were attributable to aircraft, but some were created by other sources as well. In any event, indoor and outdoor levels were analyzed separately.

4. Response measurements

Nighttime motility was recorded with wrist-worn actimeters. All 30 test participants in the first round of data collection living near DEN were provided with the same (Gaehwiler) actimeters employed in the study of Ollerhead *et al.* (1992). In addition, six test participants were actigraphs manufactured by Ambulatory Monitoring, Inc. (AMI).⁵ The latter actigraphs were also used to measure motility in the homes of six test participants near DIA during the second round of data collection. All test participants in the third and fourth rounds of data collection were provided with AMI actigraphs to measure motility.

B. Study 2: Observations of noise-induced sleep disturbance near DeKalb-Peachtree Airport

1. Study sites and data-collection schedules

Observations of noise exposure and sleep disturbance were made in 12 single-family detached homes to the north of PDK. Data collection began on 2 July 1996, 15 days prior

to the opening ceremony of the Olympic Games, continued through the Games, and ended one week after their conclusion.

2. Test participants

Observations of sleep disturbance of a total of 25 different people were made in the course of this study. Procedures for recruiting test participants near PDK were similar to those of study 1.

3. Noise measurements

Unattended noise-monitoring instrumentation was configured in this study as in the latter rounds of data collection in study 1. Indoor noise measurements were made continuously for the 6 weeks. Continuous 2-s time histories were recorded, as were hourly average sound levels. Levels associated with noise events were stored in the noise monitors during the entire measurement period. Noise events were defined as a time series of noise levels that began when a preset threshold was exceeded for at least 10 s, and that continued as long as the level remained more than 2 dB below the preset threshold. The threshold for indoor locations in this study was an A-weighted sound level of 50 dB.

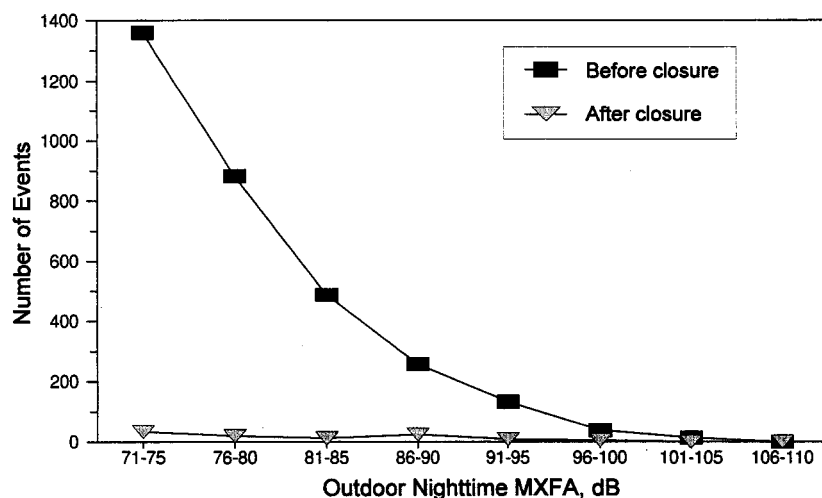


FIG. 2. Distribution of outdoor noise events at DEN between 2200 and 0700 hours.

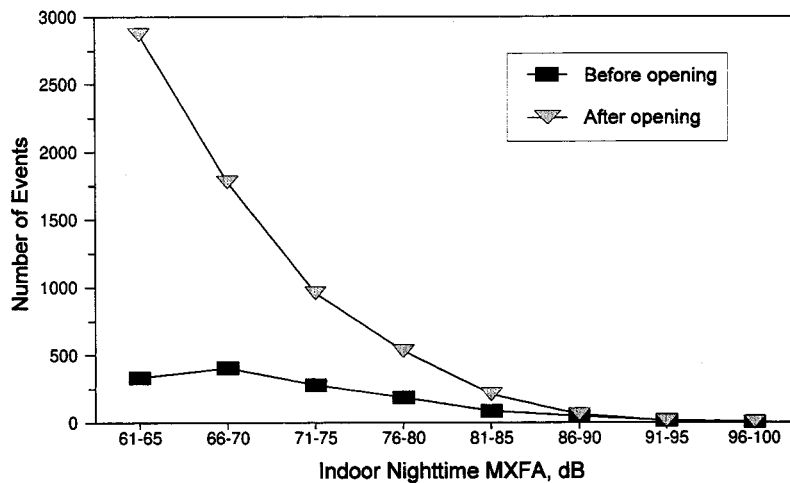


FIG. 3. Distribution of noise events recorded inside test participants' sleeping quarters at DIA from 2200 to 0700 hours.

Outdoor noise measurements were made in the vicinity of all test participants' homes with two Larson-Davis 820 noise monitors, with the same parameters used to collect indoor noise data. The outdoor noise-event threshold was set to an A-weighted level of 60 dB.

4. Response measurements

Sleep disturbance measurements in this study were identical to those of study 1. Recording accelerometers (AMI actigraphs) configured in the zero-crossing mode measured nighttime motility in 30-s epochs.

II. RESULTS

A. Study 1: Summary of noise measurements and sleep disturbance responses

Table I summarizes the conditions under which 2717 subject-nights of data were collected in the first study, and 686 subject-nights in the second study. Figures 1–4 summarize the distributions of indoor and outdoor nighttime noise-event levels in each round of data collection. The noise levels are reported in terms of maximum sound level using fast response (125 ms) with A-weighting (MXFA). As expected, outdoor nighttime noise-event levels decreased after flight

operations ceased at DEN. Outdoor nighttime noise-event levels increased, although less dramatically, near test participants' homes after flight operations began at DIA. Note that no data are shown for the lowest indoor noise level interval for the "after closure" condition. Nonaircraft noise events in the homes of one or more test participants in this condition appear to have generated artifactual "intrusions." Indoor nighttime noise-event levels as measured in sleeping quarters were much less affected by the changes in aircraft operations at DEN. However, indoor noise events greatly increased after opening of DIA.

Figure 5 compares the average motility of individual test subjects attributable and unattributable to aircraft noise, while Fig. 6 compares the average rate of behavioral awakening responses attributable and unattributable to aircraft noise. The pattern of findings summarized in these figures indicates that motility and awakenings were little affected by the changes in nighttime aircraft flight operations.

Table II summarizes the number of awakenings confirmed by button pushes averaged over the two sites. Few of the responses were attributable to noise events.

B. Study 1: Inferential analysis

All dosage–response relationships were constructed from noise events occurring between 2200 and 0700 hours,

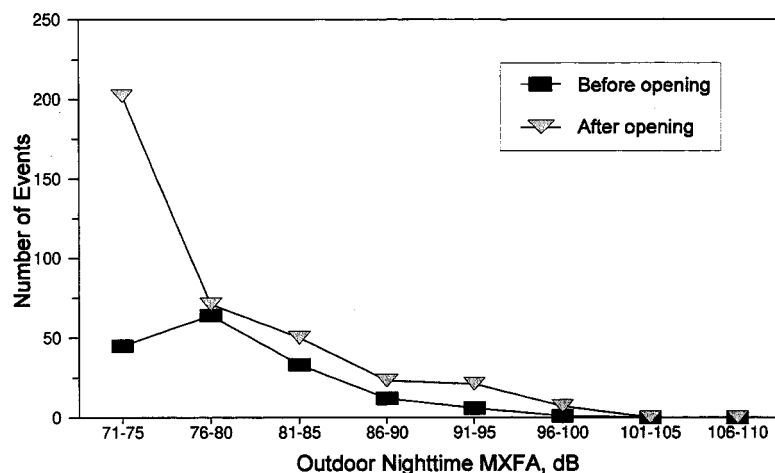


FIG. 4. Distribution of outdoor noise events at DIA between 2200 and 0700 hours.

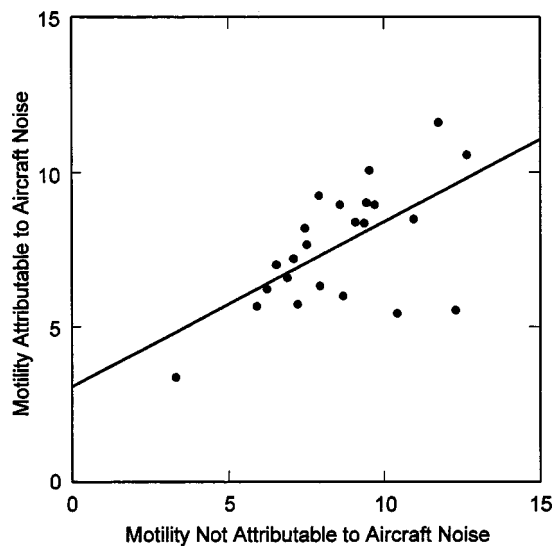


FIG. 5. Average motility in each 30-s epoch attributable to (ordinate) and not attributable to (abscissa) aircraft noise. Each data point represents responses of a single participant throughout study 1.

because earlier time periods in the evening and later time periods in the morning contained too high a density of (largely nonaircraft) noise events for reliable association with individual responses. Dosage–response relationships were constructed for five indicators of sleep disturbance:

- (1) behavioral awakening responses (button pushes),
- (2) arousals defined by Ollerhead *et al.* (1992)⁶ criteria for the Gaehwiler actimetric data,
- (3) arousals defined by Cole *et al.* (1992) criteria for the AMI actigraphic data,
- (4) motility as recorded by the Gaehwiler actimeters, and
- (5) motility as recorded by AMI actigraphs.

Analyses of awakenings associated with independently confirmed aircraft noise events at DIA were possible only for the month prior to opening and the month following the

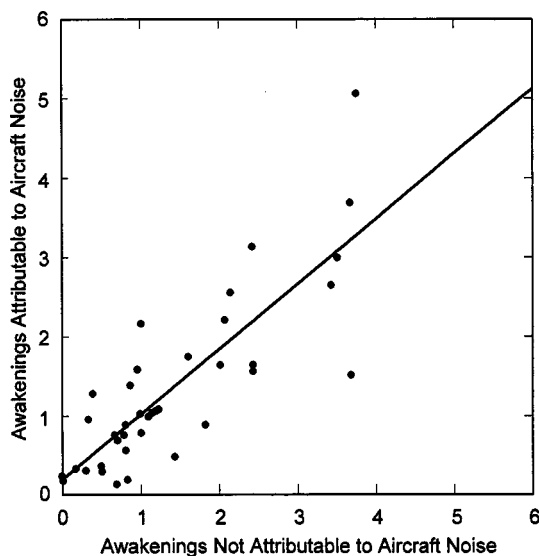


FIG. 6. Average number of behavioral awakening responses per night attributable to (ordinate) and not attributable to (abscissa) aircraft noise. Each data point represents responses of a single participant throughout study 1.

opening of the airport. The independent (predictor) variable for all dosage–response relationships was either indoor or outdoor sound-exposure levels (SEL), quantized in 3-dB intervals. Data points reflect the proportion of noise events in each noise level interval that produced a response. Data were combined for all test participants and all data collection sessions for behavioral awakening and AMI actigraph responses. Gaehwiler actimeter recordings were available only at DEN for the data collection session before airport closure. Table III shows the definitions of awakening, arousal, and motility adopted for the various data collection devices.

Correlations for the various dosage–response relationships are summarized in Table IV. Polynomial regressions revealed no higher-order relationships among noise metrics and any measures of sleep disturbance. Four of the dosage–response relationships, all based on SEL of noise events measured indoors, were statistically reliable. The SEL value of indoor noise events successfully predicted (1) behavioral awakening responses, (2) motility as recorded by the Gaehwiler actimeters, (3) motility as recorded by the AMI actigraphs, and (4) AMI actigraphic arousals as defined by Cole *et al.* (1992). None of the sleep disturbance measures varied reliably with SEL of noise events measured outdoors, nor did they vary reliably with SEL of confirmed aircraft noise events only.

Figure 7 shows that the probability of occurrence of at least one actimetric response recorded by a Gaehwiler actimeter within 5 minutes of the start of a noise event was strongly related to indoor SEL, $r(9)=0.90$, $p<0.001$. The data set in which this relationship was observed was composed of noise events recorded for the participants at DEN before airport closure. The slope of the regression equation shown in Fig. 7 is rather shallow: each 1-dB increase in SEL raised the probability of an actimetric blip (per the definition of Ollerhead *et al.*, 1992) by about 1.23%.

Figure 8 shows that the probability of occurrence of an average number of zero crossings greater than 0 as measured by the AMI actigraph also was reliably related to indoor SEL [$r(9)=0.84$, $p<0.025$]. The data set in which this relationship was observed was based on six participants in the first rounds of data collection at DEN and DIA, and all participants for remaining data collection periods. Each 1-dB increase in SEL raised the probability of occurrence of a motility indication by about 0.4%. The difference between correlations with SEL for the two actimetric criteria ($r=0.84$ and 0.90) was not statistically reliable.

Figure 9 shows that the indoor SEL of noise events predicted behavioral awakening responses moderately well [$r(10)=0.68$, $p<0.025$]. The probability of awakening increased by about 0.25% for each 1-dB increase in SEL.

Arousals as scored by the Cole *et al.* (1992) actigraphic criterion also were predicted reasonably well by the indoor SEL of noise events [$r(9)=0.62$, $p<0.025$], as shown in Fig. 10. Motility measurements were collected from six participants in the first rounds of data collection at DEN and DIA, and from all participants during the remaining data collection periods. The probability of arousal increased about 0.28% for each 2-dB increase in SEL of indoor noise events.

TABLE II. Summary of behavioral awakening responses for all subject-nights at DEN and DIA.

	Variable		
	Average number of behaviorally confirmed awakenings per night	Average number of spontaneous awakenings per night	Average number of noise-related awakenings per night
Attributable to aircraft			
Mean	1.53	1.27	0.26
Standard deviation	1.99	1.80	0.59
Range	0–23	0–21	0–4
Not attributable to aircraft			
Mean	1.69	1.49	0.19
Standard deviation	2.12	1.98	0.56
Range	0–19	0–16	0–5
Averaged over all nights			
Mean	1.61	1.39	0.22
Standard deviation	2.06	1.90	0.57
Range	0–23	0–21	0–5

None of the correlations in Table IV differs reliably from the others ($p > 0.05$).

C. Study 2: Summary of noise measurements and sleep disturbance responses

Informed consent was obtained from 25 residents of 14 homes to participate in this study. Data screening reduced the number of participants considered in the data analyses to 22 residents of 12 homes. Indoor noise event levels were measured in sleeping quarters, while outdoor event levels were estimated from the measurements made at the closer of the two outdoor measurement locations to a participant's house. Table I (in Sec. II A) shows the numbers of subject-nights of observations made in the periods before, during, and after the Olympic Games.

Figures 11 and 12 show the distributions of maximum levels of noise events recorded between 2200 and 0700 hours indoors and outdoors near PDK before, during, and after the Olympics. The number of indoor noise events observed prior to the Olympics exceeded that observed during and after the Olympics, as shown in Fig. 11. The number of noise event levels in each 5-dB interval measured outdoors during the Olympics was greater than that observed both before and after the Olympics, as shown in Fig. 12.

Table V summarizes the number of awakenings confirmed by button pushes. A total of 540 behavioral awakening responses (button pushes) was observed during the 294 subject-nights of data prior to the Olympics, for an average of 1.8 per night. During the Olympics, 370 behavioral awakenings were observed during the 295 subject-nights of data, for an average of 1.3 per night. After the Olympics, 98 behavioral awakening responses were observed during the 97 subject-nights of data, yielding an average of 1 per night.

D. Study 2: Inferential analyses

Three indicators of sleep disturbance were constructed for dosage–response analyses: behavioral awakening, motility, and arousal. The independent variable for all dosage–response analyses was either indoor or outdoor SEL, as in study 1.

Correlations for the various dosage–response relationships are summarized in Table VI. Polynomial regression failed to reveal any statistically reliable quadratic relationships. Two of the six dosage–response relationships were statistically reliable. The SEL value of indoor noise events successfully predicted arousals as defined by Cole *et al.* (1992), and the SEL value of outdoor noise events success-

TABLE III. Definitions of awakening and motility adopted for various data collection devices in study 1.

Indication of sleep disturbance	Recording device	Criterion of effect
Awakening	Push button	Occurrence of behavioral response within 5 minutes of start of noise event.
Arousal	Gaehwiler actimeter	Estimated awakening identical to that of Ollerhead <i>et al.</i> (1992).
Arousal	AMI actigraph	Estimated awakening as defined by Cole <i>et al.</i> (1992), using base algorithm without iteration.
Motility	Gaehwiler actimeter	Any activity occurring in any of the ten 30-s epochs following the start of a noise event.
Motility	AMI actigraph	Any activity occurring in any of the ten 30-s epochs following the start of a noise event.

TABLE IV. Summary of dosage–response correlations for events occurring between 2200 and 0700 hours. (Data aggregated over DEN and DIA for button-push responses and U.S.-made actigraph. Data available only at DEN for Swiss-made actimeter.)

Measure of sleep disturbance	Criterion for sleep disturbance	Number of indoor noise events	Number of outdoor noise events	Noise measurement type		
				Indoor criterion	Outdoor criterion	Outdoor confirmed aircraft
Motility	Gaehwiler actimeter (time above threshold)	1519	6915	0.90 ^a	ns ^b	nd ^c
Arousal	AMI actigraph (zero crossings)	466	1535	0.84 ^a	ns	ns
	Ollerhead (Gaehwiler actimeter)	1519	6915	ns	ns	nd
Awakening	Cole (AMI actigraph)	466	1535	0.62 ^a	ns	ns
	Behavioral awakening response	2169	8572	0.68 ^a	ns	ns

^a $p < 0.025$, one-sided.

^bns: not significantly different from a correlation of 0.

^cnd: no data.

fully predicted behavior awakenings as confirmed by button pushes. Reliable dosage–response relationships between indoor noise levels and behavioral awakening and motility could not be constructed from the data of this study.

Figure 13 shows that indoor SEL of noise events predicted arousal moderately well [$r(8) = 0.64$, $p = 0.033$]. The probability of arousal increased by about 5% for each 10-dB increase in SEL. Figure 14 shows that the outdoor SEL of noise events predicted behavioral awakening responses fairly

well [$r(9) = 0.72$, $p = 0.013$]. The probability of awakening increased by about 1.3% for each 10-dB increase in SEL.

III. DISCUSSION

A. Relationship between outdoor and indoor aircraft noise-event levels

Observed changes in the outdoor aircraft noise environment were only loosely reflected in changes of nighttime

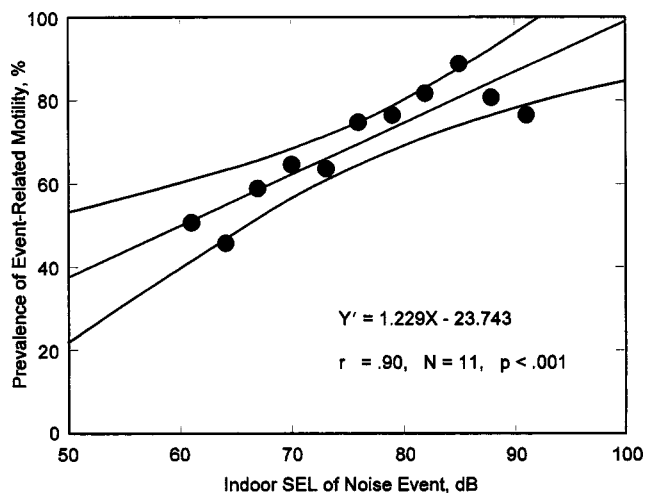


FIG. 7. Prevalence of actimetric blips (defined by Ollerhead's criterion) at DEN before airport closure, aggregated over test participants in 3-dB increments of indoor SEL values of noise events. Curved lines bound the 95%-confidence interval.

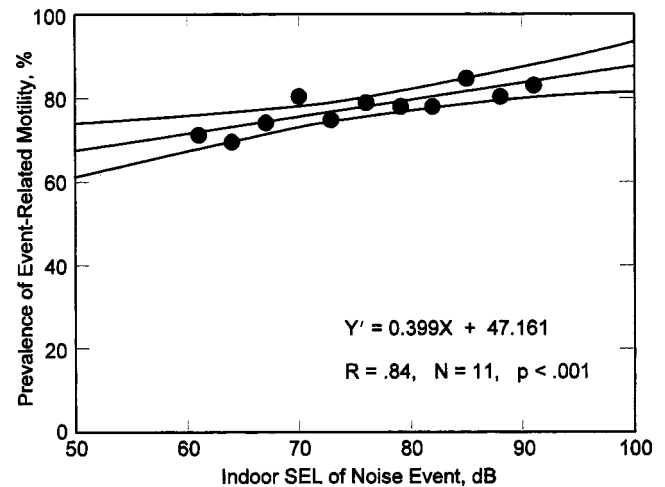


FIG. 8. Prevalence of actigraphic threshold crossings (defined by Cole's criterion) at DEN and DIA, aggregated over test participants in 3-dB increments of indoor SEL values of noise events. Curved lines bound the 95%-confidence interval.

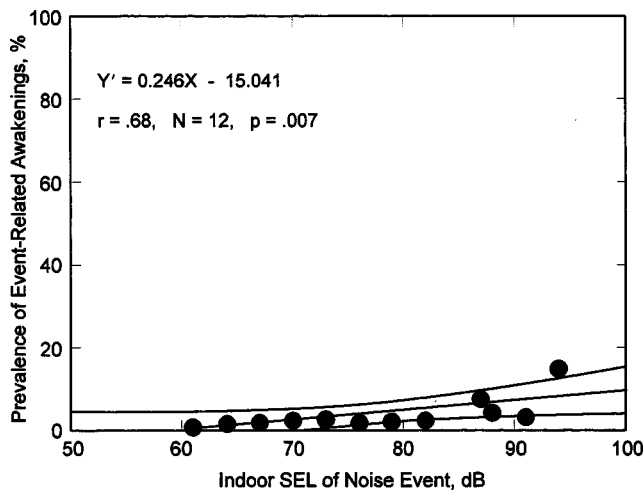


FIG. 9. Prevalence of behavioral awakening responses at DEN and DIA aggregated over test participants in 3-dB increments of indoor noise measurements. Curved lines bound the 95%-confidence interval.

noise event levels as measured in sleeping quarters. This finding is consistent with Schultz's (1982) observations that differences between indoor and outdoor single-event levels "typically fluctuate wildly over a range of as much as 30 dB," and that "the outdoor noise stimulus may have little or nothing to do with the noise actually heard indoors."

B. Dosage-response analysis of five diverse field studies

Figure 15 plots the dosage-response relationship between SEL and behavioral awakenings inferred from the present data, along with data from six field studies reviewed by Pearsons *et al.* (1995), the data from Fidell *et al.* (1995a), and the data from Ollerhead *et al.* (1992).⁷ The relationship shown for the combined data is reliable, but accounts for only about 20% of the variance in the data set. Each 10-dB increase in SEL increases the prevalence of awakening by only about 1.3%.

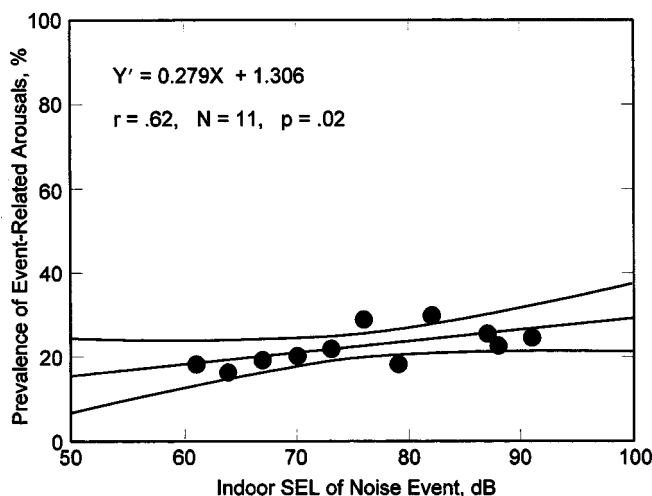


FIG. 10. Prevalence of arousal responses by U.S. actigraphic criterion at DEN and DIA aggregated over test participants in 3-dB increments of indoor noise measurements. Curved lines bound the 95%-confidence interval.

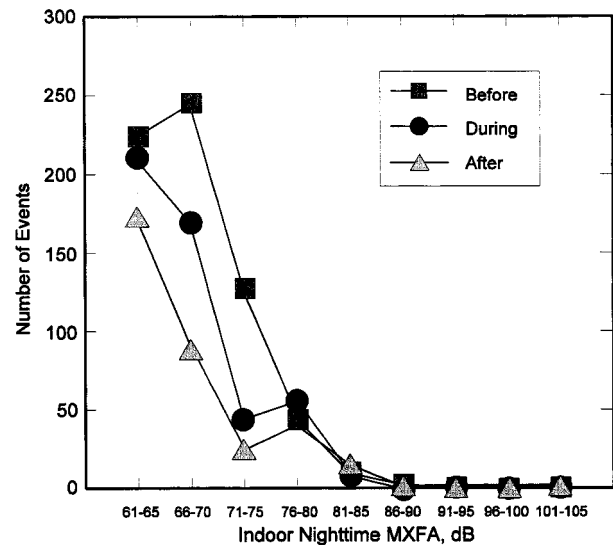


FIG. 11. Distribution of noise-event levels recorded inside test participants' sleeping quarters between 2200 and 0700 hours in three measurement periods.

The combined dosage-response relationship shows much greater variability at higher than at lower noise levels. For example, the range of prevalence of awakening at 60 dB is from 0% to about 2%. The range at 100 dB is from 0% to more than 20%. Even long-duration and/or high-level noise events sometimes failed to awaken test participants.

C. Dosage-response analysis of three similar field studies

The data of the studies described in this paper were combined with those of a prior study by Fidell *et al.* (1995a) in which similar definitions of noise events and of awakening were applied. Correlations for the six dosage-response relationships are summarized in Table VII. Four of the six dosage-response relationships were statistically reliable. The SEL value of indoor noise events successfully predicted behavioral awakenings as confirmed by button pushes.

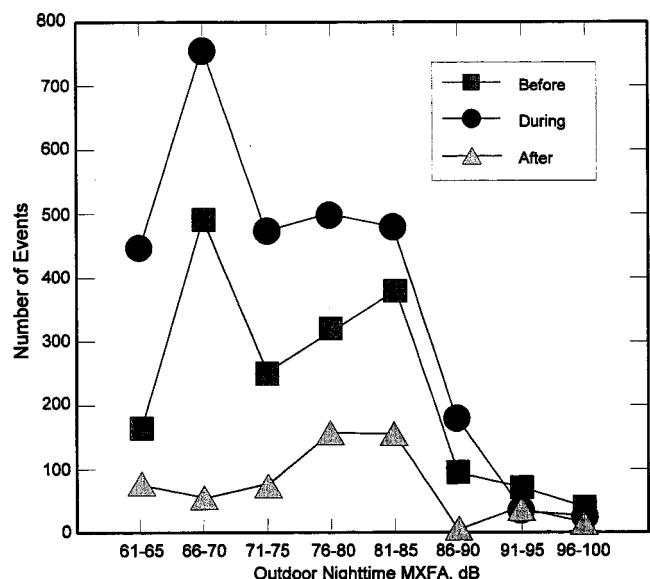


FIG. 12. Distribution of outdoor noise-event levels between 2200 and 0700 hours in three measurement periods.

TABLE V. Summary of behavioral awakening responses for all subject-nights during study 2.

Time frame	Variable		
	Average number of behaviorally confirmed awakenings per night	Average number of spontaneous awakenings per night	Average number of noise-related awakenings per night
Before and after Olympics			
Mean	1.63	1.20	0.43
Standard deviation	1.73	1.45	0.89
Range	0–10	0–8	0–7
During Olympics			
Mean	1.25	0.83	0.42
Standard deviation	1.40	1.05	0.86
Range	0–6	0–6	0–5
Averaged over all nights			
Mean	1.47	1.05	0.42
Standard deviation	1.61	1.31	0.88
Range	0–10	0–8	0–7

Figure 16 shows that indoor SEL of noise events predicted behavioral awakening moderately well [$r(15)=0.75$, $p<0.001$]. The probability of arousal increased by about 3% for each 10-dB increase in SEL. Polynomial regression revealed statistically reliable quadratic and cubic relationships ($p<0.001$), with multiple R increasing to 0.89 with inclusion of the quadratic trend and to 0.96 with inclusion of the cubic trend. For convenience of comparison with other dosage–response relationships, however, only the linear fit is shown in the figure.

Figure 17 shows that indoor SEL of noise events predicted arousal moderately well [$r(11)=0.66$, $p=0.007$]. The probability of arousal increased by about 2% for each 10-dB increase in SEL. Figure 18 shows that indoor SEL of noise events predicted motility moderately well [$r(11)=0.76$, $p=0.001$]. The probability of arousal increased by about 4% for each 10-dB increase in SEL. Polynomial regression revealed no statistically reliable quadratic or cubic relationships in either of these predictive relationships with indoor SEL of noise events.

Figure 19 shows that outdoor SEL of noise events predicted behavioral awakening moderately well [$r(16)=0.68$,

$p=0.001$]. However, the probability of arousal increased by less than 1% for each 10-dB increase in SEL. Polynomial regression revealed statistically reliable quadratic and cubic relationships ($p<0.005$), with multiple R increasing to 0.81 with inclusion of the quadratic trend and to 0.90 with inclusion of the cubic trend.

IV. SUMMARY OF FINDINGS

The present findings may appear counterintuitive, in that they suggest that the sleep of residents of neighborhoods near airports is not highly sensitive to nighttime disturbance by aircraft noise. Instead, the results indicate that relatively few nighttime noise intrusions disturb sleep, and that residential populations near airports seem well-adapted to nighttime noise intrusions. The general agreement among large-scale field studies of noise-induced sleep disturbance, involving thousands of subject-nights of observations, also suggests that further studies of a similar nature are likely to yield similar results.

The observed pattern of findings may be due in part to self-selection of airport vicinity residents for tolerance to

TABLE VI. Summary of dosage–response correlations for events occurring between 2200 and 0700 hours (data aggregated over all nights) for study 2.

Measure of sleep disturbance	Criterion for sleep disturbance	Number of indoor noise events	Number of outdoor noise events	Noise measurement type	
				Indoor criterion	Outdoor criterion
Motility	Actigraphic zero crossings	1487	5293	ns ^b	ns
Arousal	Cole <i>et al.</i> , 1992	1329	4693	0.64 ^a	ns
Awakening	Behavioral awakening response	1487	5293	ns	0.77 ^a

^a $p<0.025$, one-sided test.

^bns: Not significantly different from a correlation of 0.

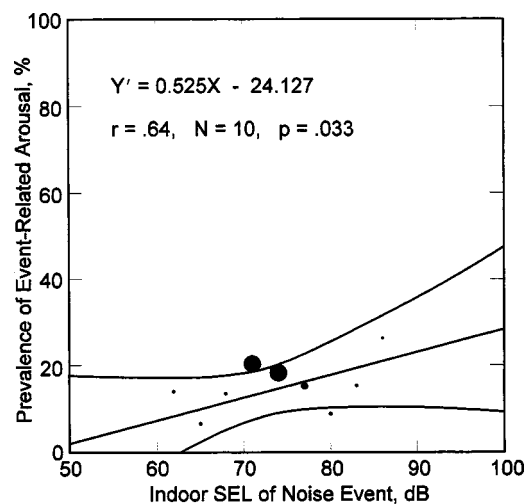


FIG. 13. Prevalence of arousal aggregated by test participants in 3-dB intervals of indoor noise measurements (for study 2). Curved lines bound the 95%-confidence interval. Larger data points indicate relatively greater numbers of events.

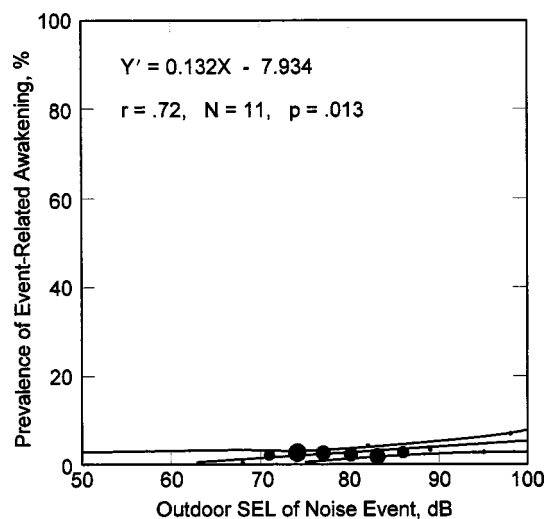


FIG. 14. Prevalence of behavioral awakening responses aggregated by test participants in 3-dB intervals of outdoor noise measurements (for study 2). Curved lines bound the 95%-confidence interval. Larger data points indicate relatively greater numbers of events.

nighttime noise intrusions, and in part to other nonacoustic factors, such as the meaning of the intruding noises. Those designing further field studies of sensitivity to noise-induced sleep disturbance may wish to consider controlled manipulations of such nonacoustic factors, including attitudinal factors and short-term habituation to novel (nonaircraft) noise sources.

V. CONCLUSIONS

Because no effort was made to rigorously define the complete population exposed to nighttime noise exposure, nor to obtain a representative sample of any wider population, conclusions drawn from the current studies apply strictly to observations made in the course of the present study only.

A. Study 1

The following are among the major findings of study 1:

- (1) The current findings closely resemble those of prior field studies of noise-induced sleep disturbance.
- (2) Numbers of outdoor nighttime noise events decreased greatly at DEN upon closure of the airport, but increased much less dramatically at DIA after opening of the airport.
- (3) Numbers of indoor nighttime noise events varied much less at DEN before and after its closing, but indoor nighttime noise events increased greatly after opening of DIA.
- (4) The average number of behavioral awakening responses per night was 1.8 at DEN and 1.5 at DIA. The number of

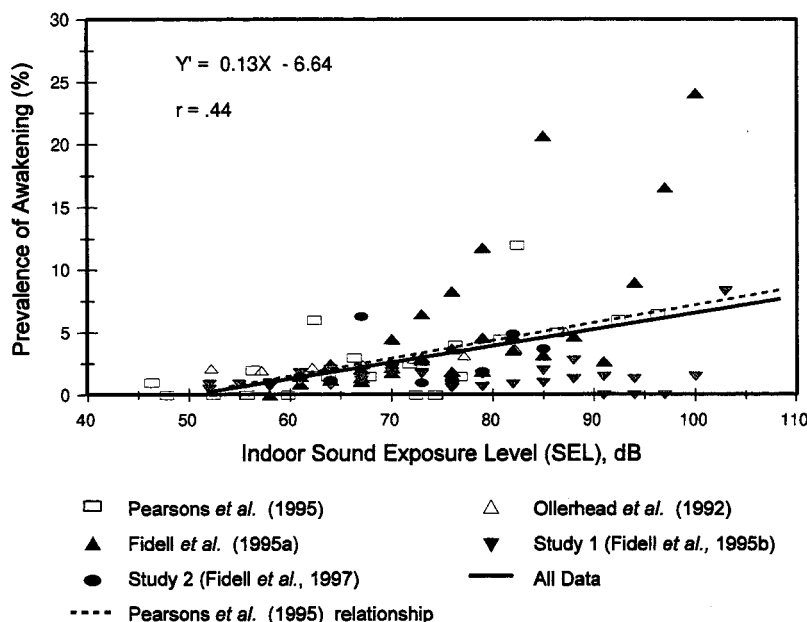


FIG. 15. Composite of data from current studies with findings of prior sleep disturbance field studies. (Data of Ollerhead *et al.*, 1992 adjusted to reflect indoor noise measurements.)

TABLE VII. Summary of dosage–response correlations for events occurring between 2200 and 0700 hours. (Data aggregated over all nights.)

Measure of sleep disturbance	Criterion for sleep disturbance	Number of indoor noise events	Number of outdoor noise events	Noise measurement location	
				Indoor criterion, based on 19 158 events	Outdoor criterion, based on 24 543 events
Motility	Ollerhead <i>et al.</i> , 1992 (zero crossings)	19 158	16 669	0.76 ^a	ns ^b
Arousal	Cole <i>et al.</i> , 1992 (time above threshold)	6 715	17 449	0.66 ^a	ns
Awakening	Behavioral awakening response	8 892	24 543	0.75 ^a	0.68 ^a

^a $p < 0.005$, one-sided test.

^bns: Not significantly different from a correlation of 0.

spontaneous behavioral awakening responses (unassociated with noise events) was 1.5 per night at DEN and 1.3 at DIA.

- (5) Statistically reliable relationships were observed between sound exposure levels of individual noise intrusions as measured inside sleeping quarters and several measures of sleep disturbances. These were:
 - (i) SEL of individual noise intrusions accounted for about 81% of the variance in motility as measured by the Gaehwiler actimeter. The linear relationship between the percentage of test participants exhibiting motility following a noise event was % motility = $1.23L_{AE} - 23.74$.⁸
 - (ii) SEL of individual noise intrusions accounted for about 71% of the variance in motility as measured by the AMI actigraph. The linear relationship between

the percentage of test participants exhibiting motility following a noise event was % motility = $0.4L_{AE} + 47.16$.

- (iii) SEL of individual noise intrusions accounted for about 45% of the variance in behavioral awakening responses. The linear relationship between the percentage of test participants exhibiting a behavioral awakening response following a noise event was % noise-induced awakening = $0.25L_{AE} - 15.04$.
- (iv) SEL of individual noise intrusions accounted for about 38% of the variance in arousals as measured by the AMI actigraph and defined and processed in accordance with the criteria of Cole *et al.* (1992). The linear relationship between the percentage of test participants exhibiting arousal following a noise event was % arousal = $0.28L_{AE} + 1.31$.

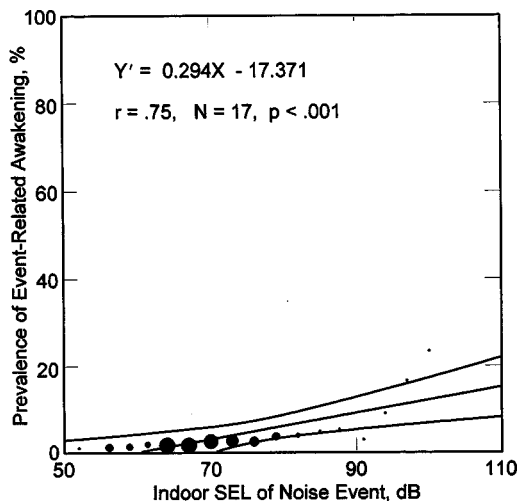


FIG. 16. Prevalence of behavioral awakening responses from three Fidell studies aggregated by test participants in 3-dB intervals of indoor noise measurements. Curved lines bound the 95%-confidence interval for the linear fit. Larger data points indicate relatively greater numbers of events.

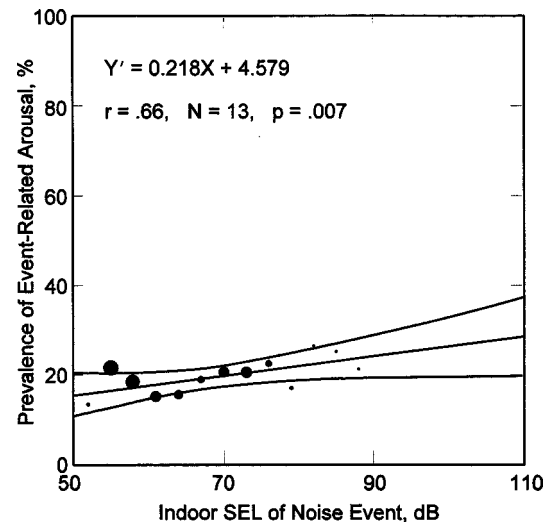


FIG. 17. Prevalence of arousal responses aggregated by test participants in 3-dB intervals of indoor noise measurements. Curved lines bound the 95%-confidence interval for the linear fit. Larger data points indicate relatively greater numbers of events.

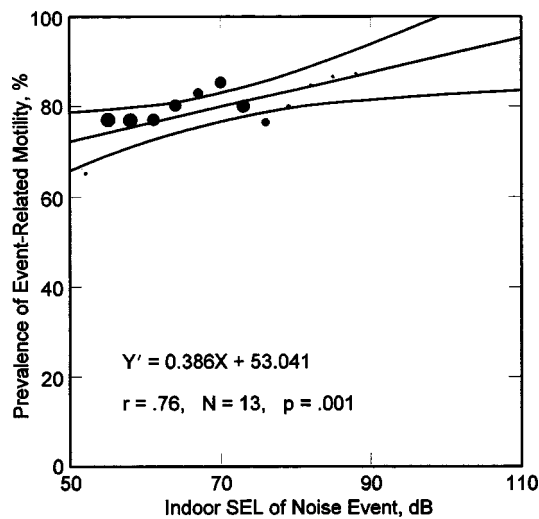


FIG. 18. Prevalence of motility responses aggregated by test participants in 3-dB intervals of indoor noise measurements. Curved lines bound the 95%-confidence interval for the linear fit. Larger data points indicate relatively greater numbers of events.

B. Study 2

Relationships between sleep disturbance and SEL values in the current findings are not fully comparable to those of prior studies. In prior studies (e.g., Fidell *et al.*, 1995a, 1995), a 2-s minimum was established for definition of a noise event. Minimum noise-event durations were set at 10 s in the current study to exclude more nonaircraft noise events. Since events with durations of less than 10 s were excluded, fewer total events were collected.

Likewise, the presence of masking by noise from mid-summer use of ventilating and air-conditioning systems in Atlanta may have been responsible for the failure to observe a reliable dosage-response relationship in the current data between awakening and indoor SEL levels, as in prior studies.

Outdoor noise event levels, defined by a level threshold,

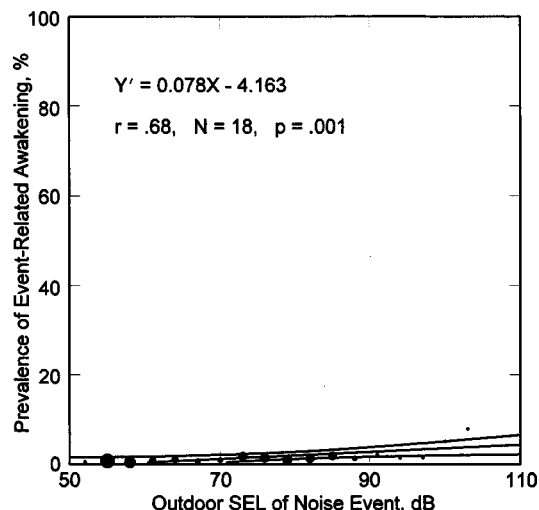


FIG. 19. Prevalence of behavioral awakening responses aggregated by test participants in 3-dB intervals of outdoor noise measurements. Curved lines bound the 95%-confidence interval for the linear fit. Larger data points indicate relatively greater numbers of events.

were reliably related to participants' behaviorally confirmed awakening as recorded by button pushes (% noise-induced awakening = $0.13L_{AE} - 7.93$), but not to motility rates as measured actigraphically, or actigraphically defined arousal. This is only partially consistent with Ollerhead's (1992) finding of a dosage-response relationship between outdoor noise levels and motility. However, both the actigraphs used to measure motility in the current study and the noise measures employed were different from those used by Ollerhead.

A reliable dosage-response relationship was found between indoor SEL of noise events and arousal as determined by Cole *et al.*'s (1992) algorithm applied to the actigraph data (% arousal = $0.52L_{AE} - 24.13$), but not between indoor SEL and behavioral awakening as recorded by button pushes or motility as measured actigraphically. This is not entirely consistent with prior findings of small but highly reliable relationships between SEL measured indoors and several indicators of sleep disturbance, including awakening and motility as well as arousal.

ACKNOWLEDGMENTS

The authors are grateful for the cooperation of test participants in the observations of in-home sleep disturbance reported here. These studies were sponsored by the Structural Acoustics Branch of NASA Langley Research Center, with Dr. Keven Shepherd acting as Technical Monitor. The field data collection at PDK was conducted under USAF Contract No. F41624-96-C9003. Dr. John Ollerhead of National Air Traffic Services, Civil Aviation Authority, Professor James Horne of Loughborough University, and Dr. Mark Rosekind of NASA Ames Research Center provided valuable instrumentation, advice, and assistance. The U.S. Air Force's Armstrong Laboratory at Wright-Patterson Air Force Base made available noise-monitoring instruments, and the U.S. Army's Construction Engineering Research Laboratory in Champaign, Illinois made available additional instrumentation for recording behavioral awakening responses. Ms. Nicole Porter brought to the author's attention a minor inconsistency in converting maximum A-weighted sound levels to SEL values (corrected in this article) in comparisons of the current findings with those of earlier U.K. findings.

¹Additional detail regarding data collection procedures, data reduction, and analyses of these two data sets is provided in Fidell *et al.* (1995b, 1998).

²Distances from runway ends to test participants' homes do not uniquely determine indoor single-event sound-exposure levels of aircraft operations. Many other factors also affect indoor aircraft noise-event levels, including differences in source levels, slant ranges to flight paths, types of operations, runway assignments, window size and opening, and so forth.

³Larson-Davis Model 820 and 870 noise monitors were used for unattended detection and recording of noise events.

⁴The abbreviations "TAVA" (or "TAV") and "MXFA" for time-averaged and maximum A-weighted, fast time sound levels follow the conventions of ANSI Standard S1.1 (1994), "Acoustical Terminology for Levels of Acoustical Measures."

⁵The Gaehwiler actimeter measures total time above an acceleration threshold, with a fixed sensitivity of 0.1 g over a bandwidth of 0.25 to 3 Hz. The AMI actimeter can measure both time above an acceleration threshold and the number of zero crossings (reversals of wrist movement directions). Its sensitivity is adjustable to 0.01 or 0.5 g over the range of 0.16 to 10 Hz.

⁶Ollerhead *et al.* (1992) define a motility "blip" as any nonzero value for the time above an acceleration threshold of wrist motion during a 30-s analysis epoch occurring after sleep onset, which was in turn defined as

starting 5 minutes into the first 7-minute-long movement-free period of the night. Cole's criterion for an awakening takes into account the duration of motility.

⁷The outdoor noise measurements reported by Ollerhead *et al.* (1992) are adjusted by 15 dB to represent them as indoor sound levels.

⁸The symbol for sound-exposure level in formulas is L_{AE} (see ANSI Standard S1.1-1994).

ANSI (1994). "American National Standard Acoustical Terminology" (American National Standards Institute, New York).

Cole, R., Kripke, D. F., Gruen, W., Mullaney, D., and Gillin, J. C. (1992). "Automatic sleep/wake identification from wrist activity," *Sleep* (N.Y.) **15**, 461–469.

FICON (Federal Interagency Committee on Noise) (1992). Final Report, Airport Noise Assessment Methodologies and Metrics, Washington, D.C. See also FICON (Federal Interagency Committee on Noise) (1992). Federal Agency Review of Selected Airport Noise Analysis Issues, Committee Recommendations, Washington, D.C.

Fidell, S., Howe, R., Tabachnick, B., Pearsons, K., Silvati, L., Sneddon, M., and Fletcher, E. (1998). "Field studies of habituation to change in night-

time aircraft noise and of sleep motility measurement methods," USAF Contract No. F41624-96-C-9003, BBN Report No. 8195.

Fidell, S., Pearsons, K., Tabachnick, B., Howe, R., Silvati, L., and Barber, D. (1995a). "Field study of noise-induced sleep disturbance," *J. Acoust. Soc. Am.* **98**, 1025–1033.

Fidell, S., Howe, R., Tabachnick, B., Pearsons, K., and Sneddon, M. (1995b). "Noise-Induced Sleep Disturbance in Residences Near Two Civil Airports," NASA Langley Research Center Contract No. NAS1-20101, NASA Contractor Report 198252, Hampton, VA.

Ollerhead, J. B., Jones, C. J., Cadoux, R. E., Woodley, A., Atkinson, B. J., Horne, J. A., Pankhurst, F., Reyner, L., Hume, K. I., Van, F., Watson, A., Diamond, I. D., Egger, P., Holmes, D., and McKean, J. (1992). *Report of a Field Study of Aircraft Noise and Sleep Disturbance* (Department of Safety, Environment and Engineering, London).

Pearsons, K., Barber, D., Tabachnick, B., and Fidell, S. (1995). "Analysis of the predictability of noise-induced sleep disturbance," *J. Acoust. Soc. Am.* **97**, 331–338.

Schultz, T. J. (1982). *Community Noise Rating*, 2nd ed. (Applied Science, London).

Surface area measurement utilizing an acoustic bridge

Ippei Torigoe

Faculty of Engineering, Kumamoto University, Kumamoto, Japan

Yasushi Ishii

Measurement Science Laboratory, Kawasaki, Japan

(Received 26 December 1998; accepted for publication 28 December 1999)

A new method is proposed for measuring the surface area of an object. The acoustic conductance of a cavity is proportional to the surface area of the cavity inner wall. The surface area of an object thus can be known from the measurement of the acoustic impedance of a chamber in which the object is placed. In order to measure the acoustic impedance accurately, the proposed method employs the acoustic bridge technique. The experimental device is composed of the following elements so arranged that their electric equivalents form a bridge circuit: a measuring chamber in which an object under test is placed and whose volume can be adjusted; a reference chamber whose inner surface area can be varied; a loudspeaker (the signal source) mounted between the two chambers; and a bypass channel at the midpoint of which a microphone (the null detector) is installed. This bridge balances when the volume and the inner surface area of each chamber become equal. The surface area of the object can then be known from the inner surface area of the reference chamber. Several experiments were performed with this device and the success of the proposed method was verified. © 2000 Acoustical Society of America. [S0001-4966(00)01404-1]

PACS numbers: 43.58.Bh [SLE]

INTRODUCTION

Surface area is a significant characterizing parameter in various fields. For example, the surface area of acoustic materials is closely related to their properties. The cost of plating depends directly on the surface area of the objects. In physiology the surface area of animals or plants is needed in order to estimate the basal metabolic rate. Another notable example is the increasing requirement in the field of public health administration. The amount of substance, which might be a carcinogen or an endocrine disrupter, eluted from an object into the surrounding medium is proportional to the area of contact between the object and the medium. In the safety test it is thus required to measure the surface area of various kinds of objects such as food containers, packages, tableware, toys, dolls, parts of coffee makers. In spite of these requirements, a complete solution for the surface area measurement has not been established. The surface area of an object of a simple shape can be calculated geometrically from the dimensions of the object. As for powders and porous materials the specific surface area is measured utilizing gas adsorption. For an object of a complex shape, however, the measurement of the dimensions is usually expensive and sometimes even difficult. The adsorption method, also, is usually not applicable to nonporous objects. The aim of this paper is to advance an entirely new device which measures the surface area of an object not geometrically, but on the basis of a physical principle.

The heat flow between an object and the surrounding medium is proportional to the surface area of the object. This heat conduction law can be applied to the measurement of the surface area of an object. We proposed one of the possible embodiments of this principle—a method utilizing sound—and investigated theoretical aspects of this method.^{1,2} If its wall is precisely adiabatic, the acoustic im-

pedance of a cavity is capacitive. For actual cavities, however, a real component arises from heat conduction at the wall. And the magnitude of this component is a function of the area of the cavity wall. Therefore the surface area of a cavity, or of an object placed in a known chamber, can be obtained from the measurement of the acoustic impedance.

The real component of the acoustic impedance of a chamber is but a small fraction of the whole content. Thus it is not easy to measure that component accurately eliminating effects of all influence quantities. In order to attain this goal, the bridge technique is employed. Two arms of the acoustic bridge are chambers; a measuring chamber in which an object is placed and a reference chamber whose inner surface area can be adjusted. The other two arms are composed of a channel bypassing the two chambers. The acoustic impedance of the measuring chamber is obtained from that of the reference chamber when the bridge is balanced. The canceling of effects of influence quantities with the bridge technique results in the accurate measurement of the acoustic impedance of a chamber.

I. PRINCIPLE

Let us begin by outlining briefly the fundamental principle. The sound attenuation due to heat conduction has received considerable attention and was investigated by several researchers.^{3–7} The following is an outline of the analysis sketched from the view point of the surface area measurement¹ and is consistent with the results of those pioneering researchers.

For sufficiently low frequencies, a cavity of an adiabatic wall is equivalently represented as a capacitor. The temperature of the air in a cavity fluctuates around the quiescent value in accordance with the pressure fluctuation. For actual cavities, the wall is nearly isothermal since the wall, in com-

parison with air, is highly conductive and has a far larger heat capacity. The heat conduction occurs between the air and the wall to yield a thermal boundary layer above the wall. The temperature distribution in the boundary layer can be determined by solving the equation of energy conservation and Fourier's heat conduction law. If the wall is plane and of large extent, and if the time dependence for pressure and temperature is sinusoidal, the excess temperature distribution θ is given by^{1,3,6}

$$\theta(x, t) = \frac{P}{\rho c_p} \{1 - e^{-(1+i)x/\delta}\}, \quad (1)$$

where

$$\delta = \sqrt{\frac{2\kappa}{\rho\omega c_p}} \quad (2)$$

is the thickness of the thermal boundary layer, x is the distance measured normal to the wall surface, $i = (-1)^{1/2}$, t is the time, P is the pressure variation, ρ is the density of air, κ is the thermal conductivity of air, c_p is the specific heat at constant pressure of air, and ω is the angular frequency. The influence of the isothermal wall on the distribution of the air temperature extends to a distance of the order of δ from the wall. For air at the atmospheric pressure and at the room temperature, δ is about 0.4 mm, for instance, at 40 Hz.

Let us next consider the acoustic impedance of a cavity of an isothermal wall. The state of the ideal gas is described by the equation of state

$$\frac{P}{\rho} = mRT, \quad (3)$$

where P is the pressure, ρ is the density, T is the temperature, R is the gas constant, and m is the molecular weight. Neglecting the higher orders of variation, we have the following equation for sinusoidal time variation of the state:

$$-\frac{i\omega\rho}{\rho_0} = -\frac{i\omega p}{P_0} + \frac{i\omega\theta}{T_0}, \quad (4)$$

where ρ is the density variation and the subscript 0 denotes the quiescent value. Integrating the left hand term of Eq. (4) over the cavity volume, and using the continuity equation

$$-\frac{i\omega\rho}{\rho_0} = \text{div } \mathbf{u}, \quad (5)$$

we find

$$-\int \int \int_V \frac{i\omega\rho}{\rho_0} dV = \int \int \int_V \text{div } \mathbf{u} dV = \int \int_S u_n ds = -U, \quad (6)$$

where \mathbf{u} is the particle velocity, V is the cavity volume, S is the cavity inner surface, u_n is the particle velocity component normal to S , and U is the volume velocity applied to the cavity. Let us next turn to the right hand side of Eq. (4). If the characteristic dimension of the cavity is large enough in comparison with δ , the excess temperature distribution is given by Eq. (1) almost everywhere in the boundary layer. Substituting Eq. (1) into θ and assuming that the pressure

varies in phase within the cavity, we can integrate the right hand side of Eq. (4) over the cavity volume;

$$-\frac{i\omega V_0}{\gamma P_0} \left\{ 1 + (\gamma - 1)(1 - i) \frac{S_0 \delta}{2V_0} \right\} p, \quad (7)$$

where γ is the ratio of specific heat, V_0 is the volume of the cavity, and S_0 is the inner surface area of the cavity. Equating Eq. (6) to Eq. (7), we find the acoustic impedance of the cavity to be given by

$$\frac{P}{U} = \frac{\gamma P_0}{i\omega V_0} \left\{ 1 + (1 - i)(\gamma - 1) \frac{S_0 \delta}{2V_0} \right\}^{-1}. \quad (8)$$

Since the term due to heat conduction (the second term in the braces) is small, Eq. (8) is expressed approximately as

$$\begin{aligned} \frac{P}{U} &\approx \frac{\gamma P_0}{i\omega V_0} \left\{ 1 - (1 - i)(\gamma - 1) \frac{S_0 \delta}{2V_0} \right\} \\ &= \frac{\gamma P_0}{\omega V_0} \left\{ (\gamma - 1) \frac{S_0 \delta}{2V_0} \right\} - i \frac{\gamma P_0}{\omega V_0} \left\{ 1 - (\gamma - 1) \frac{S_0 \delta}{2V_0} \right\}. \end{aligned} \quad (9)$$

When compared with the acoustic impedance of an adiabatic cavity

$$\frac{\gamma P_0}{i\omega V_0} = -i \frac{\gamma P_0}{\omega V_0},$$

we can see the following; owing to heat conduction at the wall, a real component arises in the acoustic impedance and the imaginary component is also affected. In addition, as Eqs. (8) and (9) indicate, the acoustic conductance is proportional to the inner surface area of the cavity and the acoustic resistance, also, is in approximate proportion to the inner surface area. On the basis of Eq. (8) or (9), the surface area of a cavity can be obtained from the acoustic impedance measurement.

In the deduction of Eq. (8) we have assumed that the characteristic dimension of the cavity is large compared with δ and thus the effects of edges, corners, and curved surfaces can be neglected. These effects as well as the influence of the thermal properties of the cavity wall were discussed in previous reports.^{1,2}

II. EXPERIMENTS

A. Experimental device

The experimental device shown in Fig. 1 and Fig. 2 was built. The device consists of two $80 \times 80 \times 70$ mm³ chambers; a measuring chamber in which an object is placed and a reference chamber. On the top wall of each chamber, bellows and a micrometer head are installed. The volume of the measuring chamber can be adjusted by feeding forward and backward the micrometer head, whereas its inner surface area remains always constant. The pleat pitch of the bellows is about 4 mm. As we shall see in the next section, this value is large enough from the viewpoint of the gap influence to the acoustical surface area. Therefore, the effect on the surface area of the pitch change due to expansion or compression of bellows can be neglected.

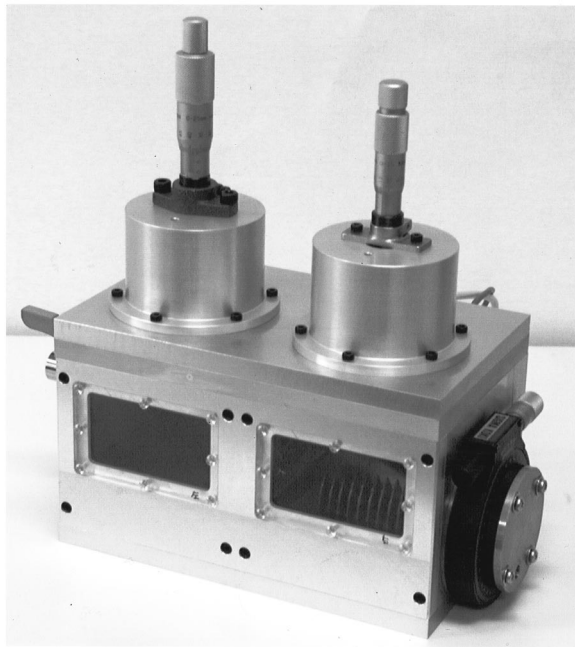


FIG. 1. Photograph of the experimental device.

The reference chamber is filled with silicon oil to the level of the bottom of the measuring chamber. On the side wall of the reference chamber a rotary stage is attached, and to this stage a rotor shaped like the rotating electrode of a

variable capacitor [Fig. 2(b)] is fixed concentrically. The rotor is positioned so that the axis thereof is included in the silicon oil surface. When the rotary stage is rotated forward, the rotor fins which were initially soaked in the oil bath emerge from the oil, and when rotated backward the fins are submerged below the oil level. In this way, the inner surface area of the reference chamber can be varied by the rotation of the rotary stage without the volume being varied.

Between the two chambers, a loudspeaker (diameter 40 mm) is mounted and driven sinusoidally at a frequency of 40 Hz to create sinusoidal volume perturbations in the two chambers that are equal in magnitude but opposite in sign. These volume perturbations lead to complementary pressure fluctuations in the two chambers. A 10 mm×3 mm rectangular groove of length 160 mm is cut on the backside wall to provide a bypass channel which connects the two chambers. At the midpoint of this channel a condenser microphone is installed to pick up the sound pressure there. Using a lock-in amplifier, the output of the microphone is converted to two dc voltages, which are used as an indicator of the bridge state.

Figure 3 illustrates the electrical equivalent for the experimental device. The measuring chamber is represented as a variable capacitor with a small serial resistance, and the reference chamber is represented by a capacitor and a variable resistance. Strictly speaking, as is seen from Eq. (8), the compliance of a cavity is affected by its surface area and the

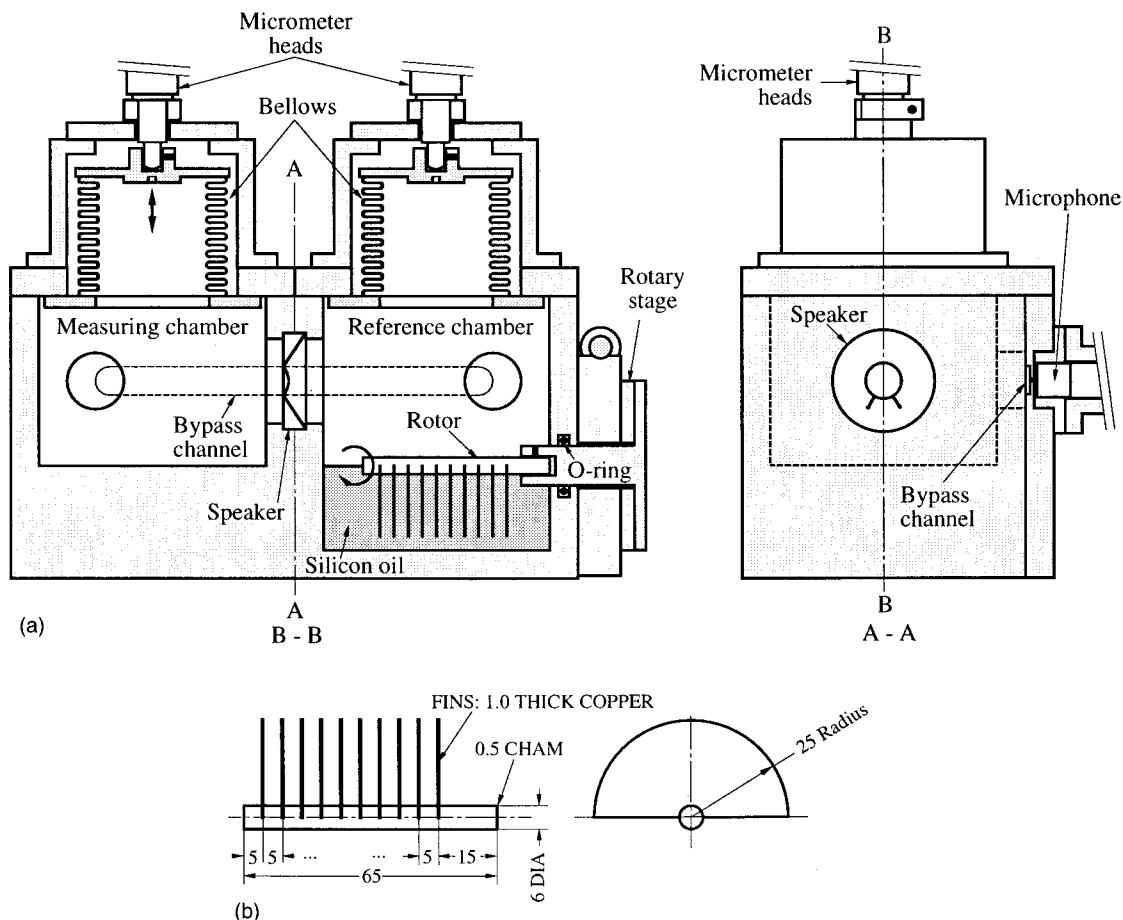


FIG. 2. (a) Experimental device. (b) The rotor (units: mm).

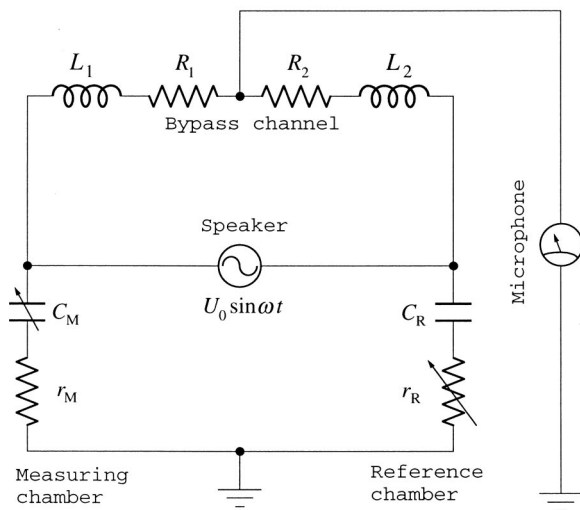


FIG. 3. Electrical equivalent for the experimental device.

resistance is also affected by its volume. However, since these effects are small, the electrical equivalents are depicted approximately for illustrative purposes. The speaker is represented as a current source whose intensity is equal to the volume velocity generated by it. The impedance of the speaker, together with its electroacoustic mechanism, is abstracted because they do not affect the balance condition of the bridge. The electrical equivalent for the bypass channel is a series of inductance and resistance. The balance equation for this bridge is

$$(R_1 + i\omega L_1) \left(r_R + \frac{1}{i\omega C_R} \right) = (R_2 + i\omega L_2) \left(r_M + \frac{1}{i\omega C_M} \right). \quad (10)$$

Since the microphone, the null detector, is installed at the midpoint of the bypass channel, Eq. (10) reduces to

$$r_R = r_M, \quad C_R = C_M, \quad (11)$$

which means that the acoustic bridge is balanced when the volume and the inner surface area of each chamber become equal.

The steps of the measurement of the surface area with this device are as follows. (i) First, the bridge is balanced by adjusting the volume of the empty measuring chamber and the inner surface area of the reference chamber. (ii) Second, the top wall of the measuring chamber is lifted, and the object to be measured is put into the chamber. (iii) Finally, the volume of the measuring chamber and the inner surface area

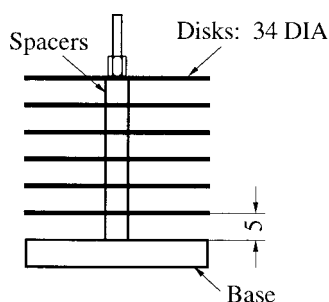


FIG. 4. Schematic diagram of the disk stack (units: mm).

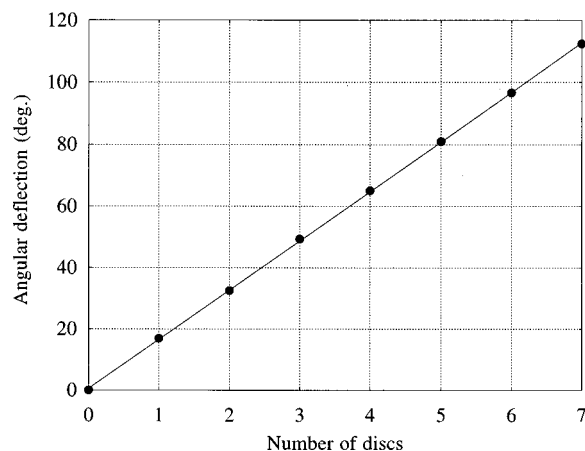


FIG. 5. Relationship between the angular deflection and the number of disks ($f=40$ Hz).

of the reference chamber are readjusted to the balance. The surface area of the object can then be known from the angular deflection of the rotor. Moreover, if need be, the volume of the object can be obtained from the displacement of the micrometer head of the measuring chamber.

B. Results

First, the surface area of copper disks of diameter 34 mm was measured. The disks were stacked on a base concentrically with equal separation (Fig. 4) and put in the measuring chamber. Figure 5 shows the relationship between the angular deflection of the rotor and the number of disks. In this experiment the disk stack is placed at one fixed position in the measuring chamber.

Although Fig. 5 suggests that the angular deflection is proportional to the surface area of an object, it remains to be tested whether the angular deflection has a dependence upon the object position or not, that is to say, a dependence upon the cavity shape. Hence, the influence of the object position, to put it another way, the influence of the medium compressibility was investigated. The angular deflection was read in two cases that the disk stack was placed at A and at B in the chamber (Fig. 6). The difference of the angular deflection between A and B was plotted in Fig. 7 as a function of

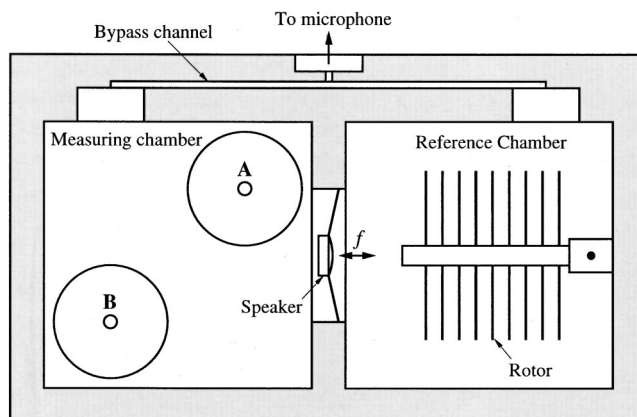


FIG. 6. Positions of the disk stack in the experiment of Fig. 7.

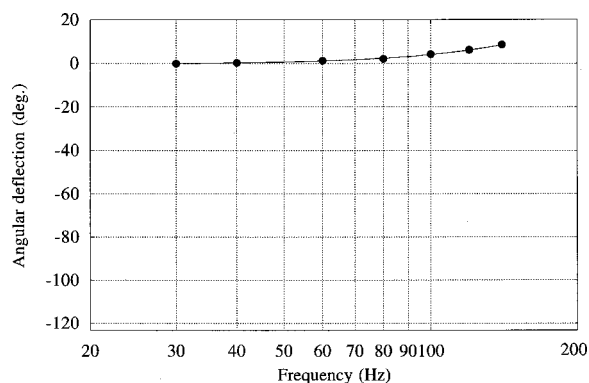


FIG. 7. Influence of the object position on the surface area measurement.

frequency. Generally speaking, when the wavelength is much greater than the dimension of the cavity, the angular deflection is expected to have little dependence upon the cavity shape. The results agree with this theoretical expectation and indicate that at frequencies lower than 40 Hz the influence of the object position is smaller than 0.3%. Furthermore, measurements in which disks were put vertically on the chamber bottom in several axis directions were made. These experiments did not demonstrate a dependence upon the object position for frequencies below 40 Hz, either.

Next, the surface areas of objects of several shapes were measured. Copper disks, steel balls, brass cylinders, and aluminum foils were measured and the results are shown in Fig. 8. The abscissa is the surface area which is calculated from the object dimensions and, for objects containing edges, corrected considering the edge effect.² The ordinate represents the measured surface area into which the angular deflection reading is converted; the scale was determined from a single-point calibration using a sample disk. Since the influences of the curved surface of objects are of a comparable order with the measurement uncertainty, the results are plotted without correcting them. Figure 8 indicates that the experimental device is capable of measuring the surface area of objects irrespective of their shape.

Finally, a theoretical prediction about the influence of

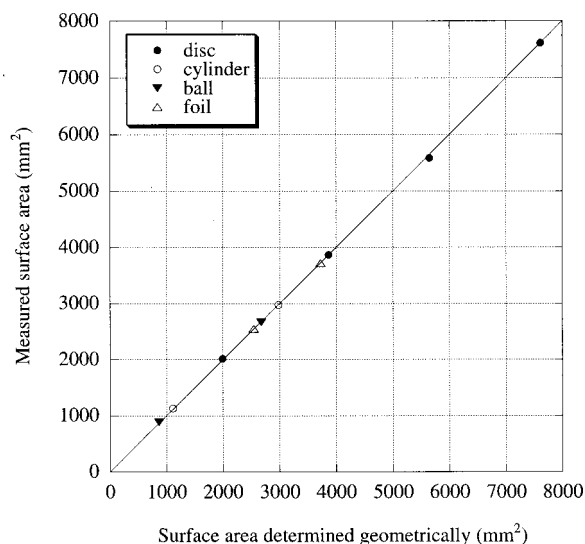


FIG. 8. Measured surface area of objects of various shapes ($f=40$ Hz).

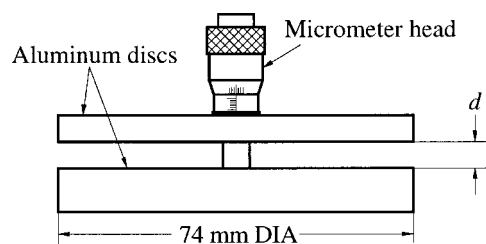


FIG. 9. Sample object used in the test of the gap influence.

gap on surface area measurement was verified. A sample shown in Fig. 9 was placed in the measuring chamber; the distance d of the gap between the two disks can be varied by feeding forward and backward the micrometer head. Figure 10 is a plot of the augmentation of measured surface area from the value at $d=0$. In the limit that the gap distance becomes far greater than the thickness of the thermal boundary layer δ , the augmentation ought to converge a constant value. The solid line in the figure represents the theoretical value.¹ The results show that the measured value agrees fairly well with the theoretical prediction.

C. Error analysis

In the deduction of Eq. (8) we have assumed the following: (i) the medium is an ideal gas; (ii) the cavity inner wall is rigid; (iii) the characteristic dimension of the cavity is large compared with δ and thus the effects of edges, corners and curved surfaces can be neglected; (iv) the thermal conductivity and the heat capacity of the wall are infinite; (v) the dimension of the cavity is small enough compared with the sound wavelength and the pressure changes in phase within the cavity. Departures in a real device from these conditions can affect the surface area measurement. In comparison with errors described later, the error derived from the departure of real air can be neglected at the atmospheric pressure, and that derived from the finite wall impedance is also negligible [(i), (ii)]. The influence of the shape of the cavity wall and the effect of edges and corners were discussed in the previous reports^{1,2} [(iii)]. The influence of the thermal properties of the cavity wall was also discussed there and found to be negligible under the conditions of our experiments [(iv)]. As to (v), we can roughly evaluate the error from the results of the experiments on object position dependence; in the ex-

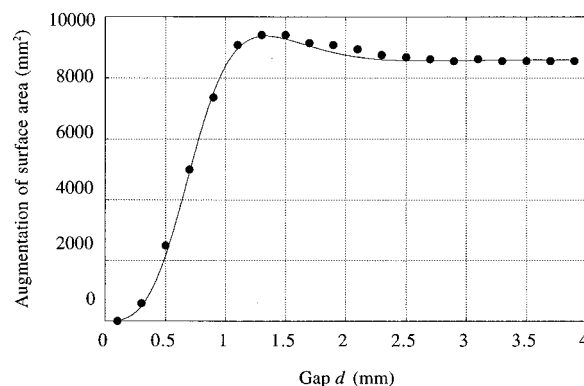


FIG. 10. Augmentation of the measured surface area versus the gap between the two disks ($f=40$ Hz).

perimental device the error due to the medium compressibility is smaller than 0.3% at a frequency of 40 Hz.

In the determination of the surface area on the basis of Eq. (8), the error in each of the variables is propagated to that in the measured surface area. More basic quantities such as temperature also affect the measurement through these variables. In order to eliminate these influences the bridge technique has been employed in the experimental device. If a perfect symmetry of the acoustic bridge is attained, all the influence will cancel out and the surface area of an object can be determined only from the inner surface area of the reference chamber. In a real device, however, imbalance in machining is unavoidable and gives rise to the departure of the acoustic bridge from ideal operation. For example, if the cross sectional area differs slightly for each side of the bypass channel, when the condition such as temperature changes, the ratio between the real and imaginary part of the acoustic impedance varies in a slightly different way between the two sides, which results in a drift of the balance point. Variations in conditions are accounted for by a calibration process that precedes each measurement. After the bridge is balanced with the baseline configuration (empty chamber) a calibration is run with a known object, in other words, a comparative test procedure is adopted. As long as the short term deflection of the surface area from the baseline is considered, all the condition, being fixed, will not affect the measurement. The uncertainty in the determination of the surface area is thus composed of random errors only.

Random errors are caused by unpredictable fluctuations of the measurement conditions: noises in the electronic circuit, the ambient pressure fluctuation, mechanical vibrations, the temperature fluctuation in the measuring chamber introduced by the fluctuation of room air temperature, the fluctuation of gas content (the humidity in particular) in the measuring chamber, etc. The measurement uncertainty due to random errors is estimated from the experimental standard deviation and is 61 mm^2 in the experiment of Fig. 8.

Let us finally evaluate the measurement uncertainty in the experiment of Fig. 8.⁸ The uncertainty due to random errors (type A uncertainty) is 64 mm^2 , which is obtained by combining the uncertainty in individual measurements with that in the baseline. On the other hand, the major sources of the systematic error are the medium compressibility, the calibration error, and the curvature of object surface. The error due to the medium compressibility is expected to be 0.3% at the maximum. Assuming the uniform distribution of the error, since we have no further information on its distribution, the corresponding standard uncertainty is estimated as $0.3/\sqrt{3}\%$.

The calibration error arises from the error in the calibration coefficient which relates the angular deflection reading to the surface area. In this experiment the uncertainty in the calibration coefficient is 1.0%. The influences of the curved surface are predicted as follows:¹ +4.8% for the smaller ball, +2.1% for the smaller cylinder, -2.9% for the foils, +2.7% for the larger ball, +1.5% for the larger cylinder, and negligible for the disks. After all, the combined standard uncertainty in the measurement of the experiment of Fig. 8, apart from the influence of the curved surface, ranges from 74 mm^2 (8.5%) to 148 mm^2 (2%).

III. CONCLUSIONS

A device was proposed which measures the surface area of an object from the acoustic impedance of a chamber in which the object is placed. In order to determine the acoustic impedance exactly the acoustic bridge technique was employed. An experimental device whose electrical equivalent was represented as a bridge circuit was built and several experiments were performed with this device. The results were in good agreement with the theoretical predictions and indicated that the experimental device was capable of measuring the surface area of various objects.

Because of the precursory nature of this study, we have concentrated on the elementary examination of the device, and the experiments presented above are far from complete. It should also be added that there are many other possible embodiments of the basic measuring principle. Future work in this study, thus, should include (i) the design of a device for practical applications, (ii) broadening the scope of experiments, and (iii) detailed tests on the effect of influence quantities.

¹I. Torigoe and Y. Ishii, "Surface area measurement utilizing sound," Transactions of the Society of Instrument and Control Engineers **34**, 182-187 (1998).

²I. Torigoe, "Edge correction of the acoustical surface area," Transactions of the Society of Instrument and Control Engineers **34**, 1739-1741 (1998).

³J. W. S. Rayleigh, *The Theory of Sound* (Dover, New York, 1945), Vol. II, pp. 319-326.

⁴H. Von Pfriem, "Zur thermischen Dämpfung in Mikrophonluftspalten," Akust. Z. **5**, S.103 (1940).

⁵H. Von Pfriem, "Zur thermischen Dämpfung in Kugelsymmetrische schwingenden Gasblasen," Akust. Z. **5**, S.202 (1940).

⁶A. K. Nielsen, "Acoustic resonators," Transactions of the Danish Academy of Technical Sciences **10** (1949).

⁷R. F. Lambert, "A study of the factors influencing the damping of an acoustical cavity resonator," J. Acoust. Soc. Am. **25**, 1068-1083 (1953).

⁸BIMP, IEC, IFCC, ISO, IUPAP, OIML, "Guide to the expression of uncertainty in measurement," ISO (1993).

Inverse problem solution techniques as applied to indirect *in situ* estimation of fish target strength

Andrzej Stepnowski and Marek Moszyński

Technical University of Gdańsk, Acoustics Department, 80-952 Gdańsk, Poland

(Received 13 November 1998; revised 23 December 1999; accepted 27 January 2000)

In situ indirect methods of fish target strength (TS) estimation are analyzed in terms of the inverse techniques recently applied to the problem in question. The solution of this problem requires finding the unknown probability density function (pdf) of fish target strength from acoustic echoes, which can be estimated by solving the integral equation, relating pdf's of echo variable, target strength, and beam pattern of the echosounder transducer. In the first part of the paper the review of existing indirect *in situ* TS-estimation methods is presented. The second part introduces the novel TS-estimation methods, viz.: Expectation, Maximization, and Smoothing (EMS), Windowed Singular Value Decomposition (WSVD), Regularization and Wavelet Decomposition, which are compared using simulations as well as actual data from acoustic surveys. The survey data, acquired by the dual-beam digital echosounder, were thoroughly analyzed by numerical algorithms and the target strength and acoustical backscattering length pdf's estimates were calculated from fish echoes received in the narrow beam channel of the echosounder. Simultaneously, the estimates obtained directly from the dual-beam system were used as a reference for comparison of the estimates calculated by the newly introduced inverse techniques. The TS estimates analyzed in the paper are superior to those obtained from deconvolution or other conventional techniques, as the newly introduced methods partly avoid the problem of ill-conditioned equations and matrix inversion.

© 2000 Acoustical Society of America. [S0001-4966(00)02505-4]

PACS numbers: 43.60.Pt, 43.30.Gv, 43.80.Ev [JCB]

INTRODUCTION

Reliable estimates of the average target strength (TS) of individual fish in logarithmic domain or backscattering cross-section (σ_{bs}) and acoustical backscattering length (\mathcal{L}_{BS}) in absolute domain are indispensable to convert echo integration data to absolute estimates of fish density and hence population estimates.^{1,2} Additionally, variations in mean target strength are thought to be among the dominant sources of nonsurvey errors in acoustic population estimates.³

There are three principal approaches to fish target strength estimation, viz.: theoretical, *ex situ*, and *in situ* measures on free-swimming fish in their natural habitat.

As is known^{4,5} TS data obtained from the theoretical and *ex situ* methods are often not reliable and consistent with *in situ* results as many factors are likely to influence fish target strength (migration, aspect, behaviors, physiological state, etc.) and may differ in respect to time and place. Therefore, the measurement of fish TS *in situ* is thought to be the most reliable and recommended approach.^{6,7}

Figure 1 shows that *in situ* methods of TS measurement require removing the effect of unknown random location of a fish in the acoustic beam $b(\theta, \varphi)$ —i.e., beam pattern factor,^{1,6} which can be achieved either directly from each individual echo, or indirectly by processing a collection of echoes. Direct methods are generally more complex and costly than indirect techniques because they require a special configuration of the transducer providing multiple beam (dual-, split-, or quasi-ideal-beam) and multi channel echo sounder receiver.^{3,8}

Indirect methods are attractive because they use the

same single-beam echo sounder as used for routine echo integration surveys, but they require knowledge of the transducer's beam pattern and assume uniform distribution of fish in sampled volume of water that might be often not the case.^{3,6} Indirect methods can be either parametric (with respect to the target strength pdf), or nonparametric.

I. REVIEW OF EXISTING INDIRECT *IN SITU* TS-ESTIMATION METHODS

Indirect TS-estimation methods had been introduced by Craig and Forbes⁹ who applied statistical correction of the measured echo level pdf to the target strength pdf, using directivity pdf, represented by circular areas covered by cross-sections of the beam pattern. Matrix inversion of the set of linear equations obtained in this way leads to the solutions (TS pdf estimate) often unreliable and strongly dependent on the actual TS pdf.

Ehrenberg,⁶ was the next who formulated the inverse problem of TS estimation in terms of a Volterra integral equation of the first kind and used an n th degree polynomial approximation to solve it for the unknown σ_{bs} pdf. The unknown polynomial coefficients were evaluated simultaneously by a least square fit. The drawback of the method is its ability to generate ill-conditioned simultaneous equations for the higher degree polynomials.

This method was modified by Robinson¹⁰ by subdividing the $p(\sigma_{bs})$ estimate space into a number of subintervals and fits low order polynomials (of $n < 3$) to the unknown σ_{bs} pdf. For the actual pdf's with a larger standard deviation (> 2.5 dB) the simultaneous equations to be solved may become ill-conditioned, as in the Ehrenberg method, which

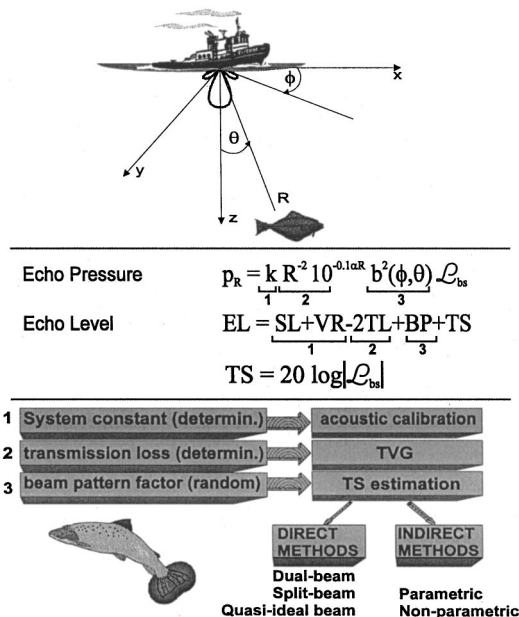


FIG. 1. The concept of *in situ* fish target strength estimation by direct and indirect methods.

may result in underestimating the mean target strength. Robinson also found that within his method modal artifacts are possible due to fitting multiple cubics.

Peterson *et al.*¹¹ introduced parametric method by using the Rayleigh distribution to model the unknown $p(\sigma_{bs})$ pdf. Medwin and Clay¹² extensively discussed numerical solution of parametric methods further simplified by Ehrenberg.¹³ The technique adjusts the unknown parameter in the Rayleigh distribution until this theoretical echo amplitude pdf and measured histogram show closest possible agreement. However, these techniques should only be employed when Rayleigh-distributed on-axis amplitude holds. Ehrenberg derived a general expression for the beam pattern factor pdf, and noted that the Rayleigh model is valid only for fish length to a wavelength ratio greater than 25. Clay and Heist¹⁴ found that a two-parameter Rice pdf was justified and that both fish activity and length condition the Rice parameters. This research also confirmed the importance of the swim bladder in acoustic scattering over a very wide frequency range.

The investigation on parametric methods was extended by Stanton, who formulated so-called “extremal pdf” for the case of overlapping fish echoes that occurs when fish concentration exceeds so-called critical fish density.^{15,12}

Lindem¹⁶ modified Craig and Forbes’ method by setting all negative estimates to zero. He also correlated fish length with modes in the indirect estimates.

Clay^{17,12} formulated Craig and Forbes’ method in terms of deconvolution of pdf’s. The “single-beam integral equation” was first formulated in terms of voltages and converted to a convolution integral by a change of variables. The numerical deconvolution has been implemented using the Z-transform polynomial long division. However, the deconvolution of actual data demonstrates oscillations and drifts in the results, especially for small echo amplitudes and low signal-to-noise ratio.¹⁷

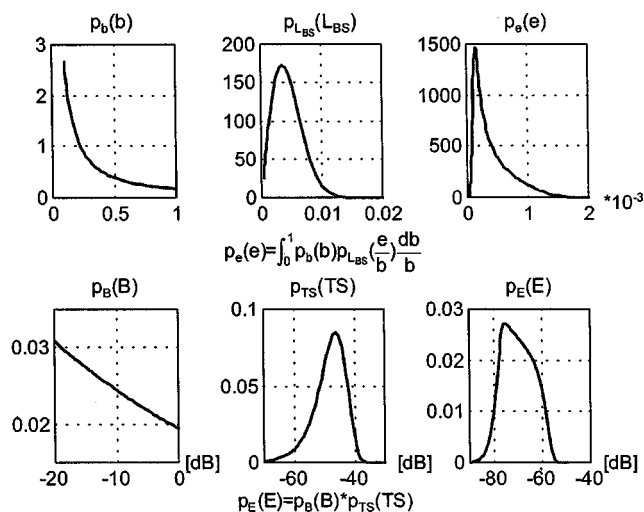


FIG. 2. Typical form of beam pattern pdf, target strength pdf, and fish echo pdf in absolute (a, b, c) and logarithmic (d, e, f) domain.

Miinalainen and Eronen¹⁸ used a least squares method but concluded that the use of nonnegative least squares for deconvolution was too time consuming and subject to noise. Therefore they used a modified singular value decomposition (SVD) routine in which all negative values in the solution were converted to zeros.¹⁸

Rudstam *et al.*¹⁹ used Clay’s deconvolution with Rice pdf’s fitted *a posteriori* to estimate fish target strength and density. The use of Rician pdf is questionable because bumpiness in deconvolved data could be artefactual, but not the result of a combination of Ricians (fitted for fish size groups).⁷ He also confirmed the good fit of acoustical estimates when compared with net studies.¹⁹

These artefactual modes are inherent to all deconvolution techniques and may be treated as a result of data under-sampling. On the other hand, *a posteriori* and parametric techniques presume knowledge of the fish scattering model, which may not be available. Therefore they can be overconstrained by the target strength pdf model and so are not robust. In general, although deconvolution and related conventional techniques can suffer from modal artifacts, if large sample sizes provided they offer better estimates as negative values of estimates are then largely avoided.

Graphical presentation of the relations between pdf’s in question is shown in Fig. 2. In absolute domain the echo amplitude pdf $p_e(e)$ is related with beam pattern $p_b(b)$ and backscattering length $p_L(\mathcal{L}_{BS})$ pdf’s by “single-beam integral equation,”⁷ equivalent to so-called Mellin convolution. In logarithmic domain the measured echo level pdf $p_E(E)$ results from the convolution of $p_B(B)$ and $p_{TS}(TS)$ pdf’s.

II. ANALYSIS OF THE INVERSE TECHNIQUES FOR TS-ESTIMATION

The recently developed direct and iterative inverse techniques^{7,20–26} have been applied by the authors as the novel indirect methods of fish target strength estimation. Although these methods are more complex than deconvolution and other conventional techniques, but by this cost they partly avoid the problems of ill-conditioned equations and

related problems with a matrix inversion. For instance, to avoid ill-conditioning a Singular Value Decomposition (SVD) method replaces inverse matrix by pseudoinverse, which eigenvalues, if too small, are replaced by zeros.¹⁸ In addition, some of these methods, viz.: a subclass of iterative techniques (EMS and MER) allow avoiding negative-valued solutions of a discrete form of a “single-beam integral equation” by superimposing proper constraints on solutions obtained by iterations.⁷ The most promising among these methods which are subject of further analysis are:

- Expectation, Maximization and Smoothing (EMS),⁷
- Tichnov and Maximum Entropy Regularization (MER),^{22,23}
- Windowed Singular Value Decomposition (WSVD),^{20,23}
- Wavelet Decomposition.^{24,27}

A. Expectation maximization and smoothing (EMS)

Estimation of the acoustical backscattering length \mathcal{L}_{BS} or backscattering cross-section σ_{bs} from fish echoes requires solving of the so-called “single-beam integral equation”⁷ which represents the pdf of echo expressed as the product of two random variables $z = xy$:

$$f_z(z) = \int_0^\infty f_x(x) f_y\left(\frac{z}{x}\right) \frac{dx}{x}, \quad (1)$$

where x is the random variable representing backscattering length of fish \mathcal{L}_{BS} ; $\mathcal{L}_{BS} = (\sigma_{bs})^{1/2}$, y is the random variable representing beam pattern of echosounder transducer $b(\theta, \varphi)$, z is the random variable representing echo amplitude (voltage v or acoustic pressure p), and $f(\cdot)$ is the pdf of random variable in question.

The EMS technique is applied to estimate the scaled pdf of the “on-axis” voltages, $f_x(x)$ in Eq. (1) from which fish backscattering length \mathcal{L}_{BS} or backscattering cross-section σ_{bs} is determined. In addition the method constrains estimates to be positive and reduces the time needed to converge by smoothing groups of estimates per iteration. To apply the EMS method, the “single-beam integral equation” (1), is transformed to a convolution:⁷

$$f_E(E) = \int_{-\infty}^\infty f_{TS}(TS) f_B(E - TS) dTS, \quad (2)$$

where the logarithmic variables E , TS , and B represent echo level $E = 20 \log p/p_1$, target strength $TS = 20 \log \mathcal{L}_{BS}$, and beam pattern in decibels $B = 20 \log b(\theta, \varphi)$, and the logarithmic transformation changes the product $z = xy$ of random variables in absolute domain to their sum in logarithmic domain $E = TS + B$.

With this logarithmic transformation the “single-beam integral equation” is “less ill-conditioned” than the Fredholm integral equation in the absolute variable (voltages), and is thus easier to solve. In addition the convolution equation solution (in the logarithmic domain) involves uniform sampling of the histogram. For optimal solution in the absolute domain, in contrast, a nonuniform sampling is required, which is more difficult to achieve in practice. Furthermore

the logarithmic approach gives the estimates obtained directly as TS values, which are most commonly used in fisheries practice.

Equation (2) can be approximated by the convolution sum to numerically obtain the solution:

$$f_E(j) = \sum_{i=1}^n f_{TS}(i) f_B(j-i) \Delta i, \quad (3)$$

where the indices i and j correspond to bins in TS and B space, respectively.

Equation (3) is a linear equation and can be presented in a matrix form:

$$\mathbf{z} = \mathbf{K}\mathbf{x}, \quad (4)$$

where \mathbf{K} is the kernel matrix of the equation corresponding to B space, \mathbf{z} is the observed data matrix in \mathcal{E} space, and \mathbf{x} is the unknown function matrix in TS space.

Every iteration procedure performed during solution when using the EMS method consists of three steps called, respectively: expectation, maximization, and smoothing.

The first step estimates the statistics of $z(x)$ as a conditional expectation:

$$z^{(n)} = E\left(z \left| \sum_{i=1}^N z_{ij}, x_{ij}^{(n)} \right.\right), \quad (5)$$

where (n) denotes the n th iteration.

The second step takes the estimated data to calculate maximum likelihood estimates as a solution of the following equation:

$$E(z_{ij}|x) = z^{(n)}. \quad (6)$$

The last step in every iteration smoothes solution x using Gaussian kernel with locally weighted end points. The smoothing process centers the kernel at each data point:

$$x^{(n0)} = \sum_j \mathbf{S}_{ij} x^{(n)}, \quad (7)$$

where \mathbf{S} is the smoothing matrix.

Finally, assuming that observations z result from a non-homogeneous Poisson process we received equation describing the first two steps in a form:⁷

$$x^{(n)} = \frac{x^{(n-1)}}{\sum_i K_{ij}} \left(\frac{z'}{x^{(n-1)} \mathbf{K}^T \mathbf{K}} \right). \quad (8)$$

B. Regularization methods

Linear inverse problems can be treated as a reconstruction of an unknown function $f(\cdot)$ (target strength pdf) out of the observed function $z(\cdot)$ (echo pdf). Thus the “single-beam integral equation” can be presented as a linear operator equation:

$$z = Kf + n, \quad (9)$$

where K is the linear operator, and n represents error or noise.

One way to solve this problem is to apply square regularization introduced by Tichonov *et al.*,²² according to which the solution estimate can be obtained as

$$\hat{f}_\lambda = (\mathbf{K}^* \mathbf{K} + \lambda \mathbf{I})^{-1} \mathbf{K}^* z, \quad (10)$$

where \mathbf{I} is the identity matrix, \mathbf{K}^* is the conjugate matrix, and λ is the regularization parameter.

Alternatively one can use iterative backprojection,²⁵ which does not require matrices to be inverted, which for ill-conditioned equations constitutes a major problem. The key issue is an optimal selection of regularization parameters in both methods.

A variation of this method, which introduces a tri-diagonal matrix as a $\lambda \mathbf{I}$ stabilizer, was used in the experimental part of this paper (Sec. IV). The advantage of this modification is the possibility of an additional *regulation* of the solution's behavior in boundary areas. The solution for the considered case is obtained as follows:

Regularization of the convolution integral equation of

the first kind in the TS domain [Eq. (2)] leads to the integral equation of the second kind by adding to the right-hand side of the so-called regulator that depends on the parameter λ and the unknown function f_x and its derivative. Introducing a functional that determines the error in space L_2 , one should determine $f_x(x)$ by minimizing this functional that leads to Euler's equation.

The solution of this equation can be determined by substituting the integration with discrete summation and obtaining the following equation:

$$\mathbf{K}^* \mathbf{K} f_x - \mathbf{K}^* f_z = \lambda \mathbf{C} f_x, \quad (11)$$

where K is the matrix of the transformation kernel formed from the convolution vector f_y , and \mathbf{C} , is the stabilizer matrix in the form

$$\begin{bmatrix} c_{11} & c_{12} & & & 0 \\ c_{21} & c_{22} & c_{23} & & \\ & c_{32} & \ddots & \ddots & \\ 0 & & \ddots & \ddots & c_{nn} \end{bmatrix}, \text{ where } c_{i,j} = 1 + \frac{2}{h^2}, \quad i=2, \dots, n-1, \quad h=1 - \text{sampling step};$$

$$c_{i,i-1} = c_{i-1,i} = -\frac{1}{h^2}, \quad i=2, \dots, n$$

$$c_{11} = c_{nn} = 1 + \frac{k}{h^2}, \quad k = \begin{cases} 1 & \text{for } f'_x(0) = f'_x(n) = 0, \\ 2 & \text{for } f_x(0) = f_x(n). \end{cases} \quad (12)$$

The matrix of stabilizer \mathbf{C} is a tri-diagonal matrix because, when computing the derivative, the difference version was used which gives the relation between the following, current, and previous index in vector f_x . At the same time boundary conditions define the value of the first and last element of the main diagonal.

The pdf estimate of unknown target strength \hat{f}_x derived from Eq. (11) in explicit form can be written as

$$\hat{\mathbf{p}}_{\text{TS}} = (\mathbf{P}_B^* \mathbf{P}_B + \lambda \mathbf{C})^{-1} \mathbf{P}_B^* \mathbf{p}_E, \quad (13)$$

where $\hat{\mathbf{p}}_{\text{TS}}$ is the vector of the estimate of target strength pdf, \mathbf{P}_B is the matrix of beam pattern pdf, \mathbf{P}_B^* is the transpose of a matrix \mathbf{P}_B , λ is the regularization parameter, \mathbf{C} is the matrix of regularization stabilizer, and \mathbf{p}_E is the vector of the echo pdf.

The alternative regularization method analyzed is Maximum Entropy Regularization (MER), which falls into iterative techniques category, in which entropy maximization constraints are achieved in every iteration²³ by minimization of the expression:

$$\|z - \mathbf{K}f\|^2 + \lambda \sum_{i=1}^n f_i \log(w_i f_i), \quad (14)$$

where w_i is the weighting coefficient.

C. Windowed singular value decomposition (WSVD)

Singular value decomposition (SVD) which was introduced in spectral analysis²⁶ has become a very useful method of solving ill-conditioned linear equations.

The basic concept of using SVD to the solution of ill-posed problems is the substitution of a simple operator inversion K^{-1} by so called "pseudo-inverse" operator $(K^* K)^{-1} K^*$. It leads to obtaining a pseudo-inverse matrix, which guarantees a solution with a minimum mean-square error:

$$\mathbf{K} = \mathbf{U} \mathbf{S} \mathbf{V}^T = [\mathbf{U}] \text{diag}(s_i) [\mathbf{V}^T], \quad (15)$$

$$\mathbf{K}^\# = \mathbf{U} \mathbf{S}^{-1} \mathbf{V}^T = [\mathbf{U}] \text{diag}(1/s_i) [\mathbf{V}^T],$$

where $\mathbf{K}^\#$ is a pseudo-inverse matrix computed numerically using the SVD algorithm, \mathbf{U} , \mathbf{V} are the orthonormal matrices, and \mathbf{S} is the diagonal matrix representing the singular values of matrix \mathbf{K} .

If it is further assumed that the product $K^* K$ is a linear operator, and if none of its singular values is a zero value, then the unknown function f can be expanded in a series of the eigenfunctions of the self-adjoint operator $K^* K$ as:

$$f = \sum_j \gamma_j^{-1} [\mathbf{K}f, h_j] e_j, \quad (16)$$

where γ_j is the singular value of $K^* K$, e_j is the eigenfunction of $K^* K$, h_j is the normalized transform defined as $h_j = K(e_j) / \|K(e_j)\|$, $\|\cdot\|_2$ represents a norm in space L_2 , and $[\cdot, \cdot]$ represents a scalar product. In the case when singular values of the $K^* K$ operator approximate zero, it is necessary to introduce weights so that dividing by elements close to zero does not impact the stability of the solution. By adequate introduction of the weights, the so-called WSVD

(Windowed SVD) is obtained which gives the reconstruction rule:

$$\hat{f}_w^{\text{SVD}} = \sum_j w_j \gamma_j^{-1} [z, h_j] e_j, \quad (17)$$

where w_j is the weight used for regularization of the solution, and z is the observations. As an example, the simplest selection of weights is to assume $w_j = 1$ for small indexes j and $w_j = 0$ for large j , which reflects the name assigned to this method. Other approaches are possible as well.²³ If one selects $w_j = k_j^2 / (k_j^2 + \lambda)$, then the regularization method of the previous section is obtained in the direct form, and when $w_j = [1 - (1 - \mu k_j^2)^m]$ we obtain the method of iterative reconstruction (m -number of iteration). By using the form of the Penrose–Moore pseudo-inverse matrix $\mathbf{K}^\#$ as an inverse matrix obtained in the SVD technique, a formula for target strength pdf estimate is obtained:

$$\hat{\mathbf{p}}_{\text{TS}} = \mathbf{P}_B^\# \mathbf{p}_E. \quad (18)$$

D. Wavelet decomposition

The imperfection of SVD arises from the fact that the eigenfunctions of linear operators represent sine and cosine functions (Fourier kernel), which induce oscillations in estimates. In many cases the better solution is obtained by use of other orthonormal function sets, which guarantee better approximation with a smaller numbers of nonzero coefficients. Numerous classes of orthogonal functions represents so-called *wavelets*, i.e., functions being dilations and translations of certain functions called the mother wavelet:²⁴

$$\psi_{j,k} = 2^{j/2} \psi(2^j t - k), \quad j, k \in N. \quad (19)$$

Wavelet expansion of Kf can be written as

$$Kf = \sum_j \sum_k d_{j,k} \psi_{j,k} = \sum_j \sum_k [Kf, \psi_{j,k}] \psi_{j,k}, \quad (20)$$

where $d_{j,k}$ represents wavelet coefficients. Then the estimate of function f can be expressed as

$$\hat{f} = \sum_j \sum_k \delta_\lambda([y, \psi_{j,k}]) K^{-1} \psi_{j,k}, \quad (21)$$

where $\delta_\lambda(\cdot)$ is the function removing the noise effect by nonlinear thresholding, at which level λ depends on a number of samples and noise variance. The product $[y, \psi_{j,k}]$ has a sense of estimates $\hat{d}_{j,k}$ of coefficients $d_{j,k}$.

For discrete vector f with a length n , its discrete wavelet transform (DWT) could also be represented as a matrix operation:

$$w = \mathbf{W}f, \quad (22)$$

where \mathbf{W} is orthogonal matrix of $n \times n$ elements being the sampled values of wavelet functions $\psi_{j,k}(t)$ in discrete time moments t/n , $t = 0, 1, \dots, n-1$.

Application of a DWT to solving linear operator equation (4) in the context of decomposition Eq. (20) leads to calculation of wavelet coefficients as a matrix multiplication:

$$d = \mathbf{W}y, \quad (23)$$

which after thresholding gives the solution

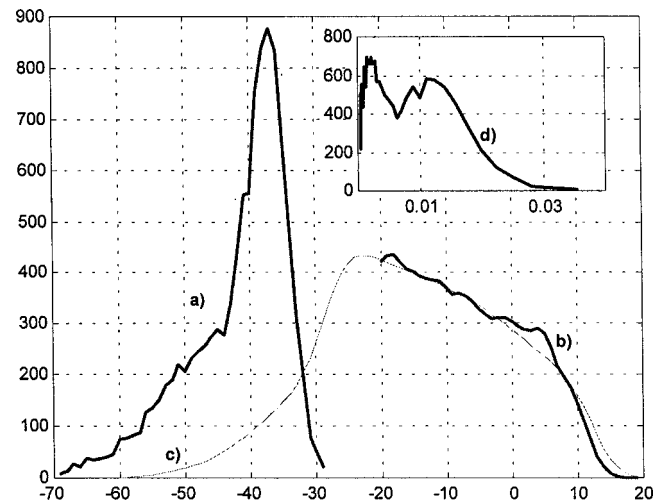


FIG. 3. Illustration of the incompleteness of data in indirect TS-estimation methods; (a) pdf estimate of target strength obtained from dual-beam reference, (b) measured echo level pdf estimate (observable data), (c) result of the convolution of target strength pdf estimate with beam pattern pdf estimate (full range of data covering observable and unobservable data), (d) pdf estimate of backscattering length \mathcal{L}_{BS} obtained from dual-beam reference.

$$f = (\mathbf{W}\mathbf{K})^{-1} \hat{d}. \quad (24)$$

The inverse operator being used on a matrix $\mathbf{W}\mathbf{K}$ represents a wavelet modified by operator \mathbf{K} . This technique is called the Vaguelette–Wavelet Decomposition (VWD)²⁴ which also has a reversed form called Wavelet–Vaguelette Decomposition (WVD).

The VWD technique was implemented in fish target strength pdf estimation from echo level pdf using matrix transforms. The inverse of beam pattern pdf was the kernel \mathbf{K} of the operator equation in the considered case. The Symmlet wavelets described in the Appendix were used in calculations.²⁸

III. RESULTS

In order to analyze and compare the performance of the investigated inverse methods of fish target strength estimation the actual fish echo data were used. The data were acquired from acoustic survey on pelagic fish populations (mostly salmon and trout) in Coeur d'Alene Lake, Idaho (provided by J. B. Hedgepeth, Biosonics, Inc., Seattle and E. Parkinson, University of Vancouver, Canada) using dual-beam digital echosounder of 420-kHz operating frequency and 0.4-ms pulse length. There were processed records of over 6500 pings from which over 10 000 fish echoes were extracted for analysis. The data obtained from the dual-beam *in situ* target strength estimation—which has been further used as a reference for comparison of the analyzed methods performance—were as follows: mean target strength: -41.9 dB modal value -38.0 dB.²⁷ These dual-beam data are presented in Fig. 3 in the form of TS pdf estimate (curve a) and \mathcal{L}_{BS} pdf's estimate (curve d). When treating the latter as an actual pdf of the data presented, the bimodal distribution is easily observed. This is probably due to the presence of two

populations of fish distributed in two depth layers, where smaller ones are in the upper stratum and the larger ones are in deeper water.

The data from the narrow beam of the echosounder—which has been further used for TS-estimation using indirect methods—are characterized by “incompleteness” due to the limited dynamic range of the echosounder in connection with the presence of the sidelobes of its transducer beam pattern. This problem is illustrated in Fig. 3, by using the dual-beam data. The function covering the range -70 dB to -30 dB represents the estimate of target strength pdf (curve a). The convolution of TS pdf (curve a) with the beam pattern pdf (curve b) results in the echo level pdf (curve c), having the range -60 dB to 20 dB. However, the narrow-beam data used for indirect TS-estimation cover only the range -20 dB to 20 dB, which is consistent only with the part of the echo pdf obtained by the convolution. This “incompleteness” of data obtained from indirect TS-estimation methods can be explained by the limited detection capabilities of the echoes received from small fish, when located in the areas of lower directivity of the beam (far off its axis).

Mathematical interpretation of this constraint is explained by Eq. (25) which represents the convolution of the target strength and beam pattern data in the matrix form:

$$\begin{bmatrix} E_1 \\ E_2 \\ E_3 \\ \mathbf{E}_4 \\ \mathbf{E}_5 \\ \mathbf{E}_6 \\ \mathbf{E}_7 \end{bmatrix} = \begin{bmatrix} B_1 & & & & & & \\ B_2 & B_1 & & & & & \\ B_3 & B_2 & B_1 & & & & \\ \mathbf{B}_4 & \mathbf{B}_3 & \mathbf{B}_2 & \mathbf{B}_1 & & & \\ & \mathbf{B}_4 & \mathbf{B}_3 & \mathbf{B}_2 & & & \\ & & \mathbf{B}_4 & \mathbf{B}_3 & & & \\ & & & \mathbf{B}_4 & & & \end{bmatrix} \begin{bmatrix} \mathbf{TS}_1 \\ \mathbf{TS}_2 \\ \mathbf{TS}_3 \\ \mathbf{TS}_4 \end{bmatrix}. \quad (25)$$

The boldfaced elements of the \mathbf{E} vector on the left side of the equation represent the measured data (echo), while the boldfaced elements on the right side of the equation represent those components of target strength and beam pattern pdf's which generate the measurable data.

In analysis of experimental data the approximation of narrow-beam pattern pdf was used as derived in Ref. 20. The 3-dB beamwidth was 6° and the drop off was -11 dB on 6° . The main-lobe of beam pattern was approximated by the expression:

$$b(\theta) = \left(1 - (1 - 2^{-\gamma}) \frac{(1 - \cos \theta)}{1 - \cos \theta_{3 \text{ dB}}} \right)^{1/\gamma}, \quad (26)$$

where the exponential coefficient γ was fitted to be equal to -0.1 .

Equation (26) leads to the following approximation of beam pattern pdf:

$$p_b(b) = \frac{k}{b^{1-\gamma}}, \quad (27)$$

where k is the normalization constant.

Equation (27) allows for construction of matrix being a kernel of linear equation in question.

The results of the target strength pdf's estimation obtained from the experimental data with the use of the ana-

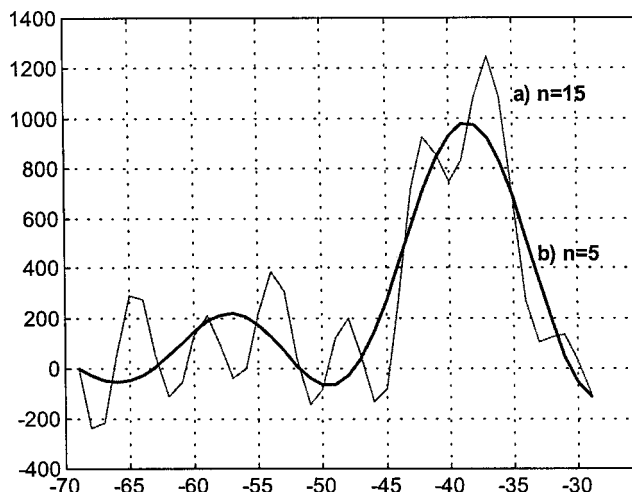


FIG. 4. Target strength pdf estimates obtained by WSVD with different window size n ; (a) $n = 15$, (b) $n = 5$.

lyzed indirect methods are presented below. Figure 4 illustrates the estimates obtained by windowed SVD technique utilizing 5 and 15 first elements (window size) of the diagonal matrix constructed in the SVD procedure (the remaining elements were neglected after inversion). As is seen, using only 5 most significant of all 40 elements is sufficient to extraction of the main modal value of the estimated pdf. The presence of the oscillations which appeared for the small TS-values region of the estimate should be noted.

Figure 5 presents the results of the two regularization techniques: the first one (a) uses the scalar value of the regularization parameter; the second (b) applies the tridiagonal stabilizer matrix, which allows determining the boundary conditions of the solution.

Figure 6 illustrates the target strength pdf estimates obtained by two iterative techniques—EMS and MER. Both estimates demonstrate good smoothness of estimates and, in addition, properly locate the main mode of the pdf.

Figure 7 shows the wavelet decomposition results along with the wavelet coefficients before and after thresholding.

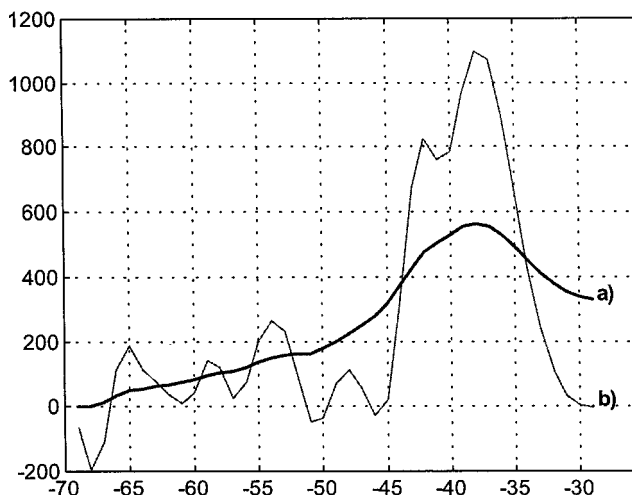


FIG. 5. Target strength pdf estimates obtained by regularization method; (a) optimal regularization parameter, (b) nonoptimal regularization parameter.

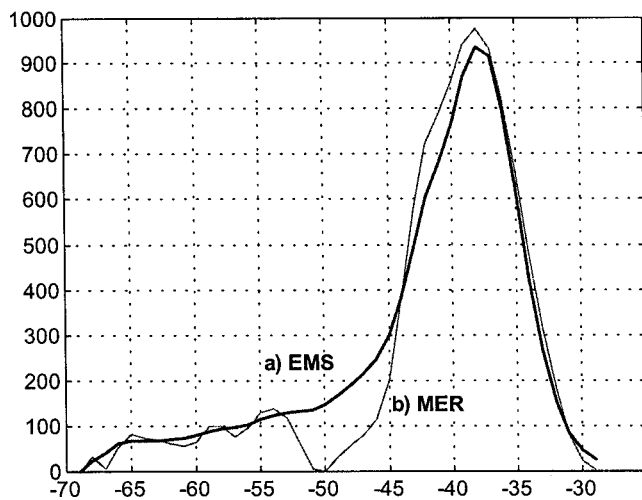


FIG. 6. Target strength pdf estimates obtained by EMS and MER methods; (a) EMS, (b) MER.

IV. CONCLUSIONS

Modern digital data processing techniques of inverse problem solution have been successfully applied to indirect *in situ* estimation of fish target strength. The results obtained by introduced techniques demonstrate their usefulness to the problem in question, as they constitute the tools of reliable and much more accurate TS-estimation than other state-of-the-art methods.

Comparison of the performance of the introduced and analyzed inverse techniques for fish target strength estimation is presented in Fig. 8 and Fig. 9 in the form of TS pdf's and \mathcal{L}_{BS} pdf's estimate plots, respectively. The first TS density estimate plot (a) was obtained from the dual-beam method and is used as a reference for investigated estimates derived from the indirect techniques. The second plot (b) shows the estimate calculated by "conventional" Craig and Forbes method for comparison purposes. Plots (c) through (g) illustrate the estimates obtained by newly developed indirect techniques. As is easily seen, these estimates, as compared to plot (b), demonstrate significantly improved results as well as refer to the shape (smoothness) and accuracy

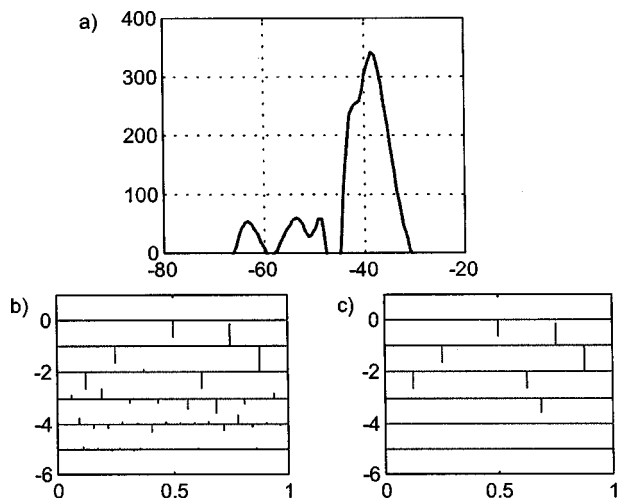


FIG. 7. Wavelet-vaguelette decomposition with thresholding (a), wavelet coefficients before (b), and after thresholding (c).

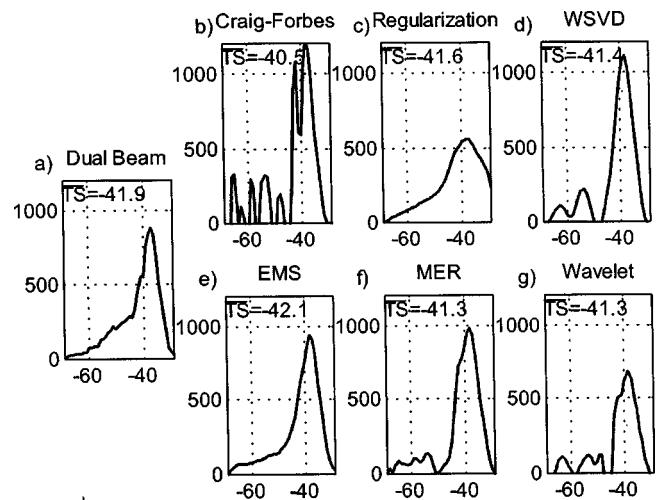


FIG. 8. Comparison of the target strength pdf's estimates obtained by different indirect methods with dual-beam method reference.

(mean value) of the TS pdf's estimates. For instance, the bias introduced by Craig and Forbes method (calculated as a difference of its mean value *re*: dual-beam reference mean value) is equal to 1.4 dB, while the bias introduced by the newly developed methods never exceeds 0.6 dB, and in the best case (EMS estimate) the bias is only 0.2 dB.

Among the introduced methods, two iterative techniques—EMS and MER—seem to be superior to other methods as they give smoothed (by iterations) estimates, mainly similar to the dual-beam reference. WSVD and wavelet methods produce slight artefactual modes and oscillations, which appeared for the small TS-estimate values region, but on the other hand they give much faster numerical solutions, and also correctly locate the main modal value of the estimated pdf's. The Tichonov Regularization [plot (c)] seems to appear the poorest of the analyzed estimates, probably due to its oversmoothing which results from not well-defined value of the regularization parameter.

The \mathcal{L}_{BS} pdf's estimate plots shown in Fig. 9 demonstrate practically similar features and performance as their analogs in the logarithmic domain, presented in Fig. 8. As

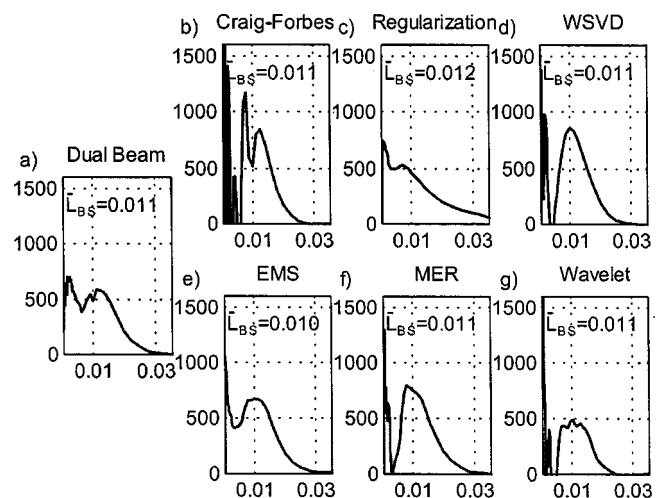


FIG. 9. Comparison of the backscattering length pdf's estimates obtained by different indirect methods with dual-beam method reference.

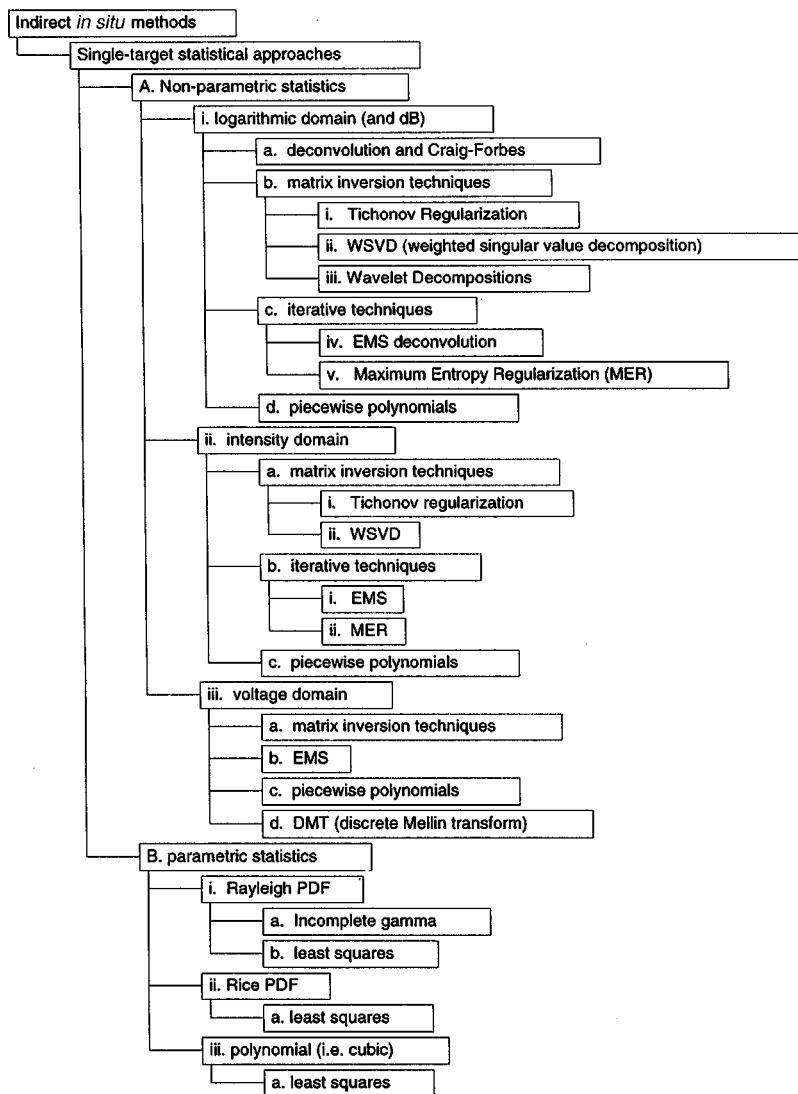


FIG. 10. Revised classification scheme of indirect *in situ* target strength estimation methods.

mentioned in the description of Fig. 3, the absolute domain estimates are superior to their logarithmic domain counterparts as they more clearly reveal the bimodality of pdf plots, which is easily seen from comparison of the dual-beam pdf estimates represented by curve (a) in Fig. 8 and Fig. 9.

However, in general, the \mathcal{L}_{BS} pdf's estimate performance is poorer than TS pdf's estimates as the latter are "less ill-conditioned." This is due to the fact that estimates in absolute domain are derived from the "single-beam integral" (1) which kernel argument and differential have both the rational form.

To conclude this paper, some revision of target strength estimation methods classification scheme as introduced by Foote¹ is proposed in order to include the newly developed methods discussed in this paper. The primary changes included in Fig. 10 refer to:

- (1) placement of MER (maximum entropy regularization) along with the EMS method, labeled by the authors as *iterative methods* in a logarithmic domain;
- (2) addition of the DMT (Discrete Mellin Transform) method to the voltage domain;²¹
- (3) addition of the WSVD (windowed singular value decomposition) method to matrix inversion as a subclass;

- (4) addition of the Wavelet Decomposition method to the same class of matrix inversion.

Other changes follow, in general, Hedgepeth⁷ suggestions and include the addition of the voltage domain, the inclusion of Craig-Forbes and deconvolution in the same sub-section, and the possibility of fitting the Rice pdf in parametric estimation (as opposed to *a posteriori*).

ACKNOWLEDGMENTS

The authors wish to express their thanks to Dr. John Hedgepeth for kindly supplying a set of the echo survey data. The authors would also like to thank Dr. Clarence Clay and Dr. Robert Chivers for valuable comments on the revision of this paper.

APPENDIX: SYMMLET WAVELETS

Symmlet wavelets are extensions of Daubechies wavelets²⁸ with increased symmetry while still greatly retaining their simplicity. In general, the symmlets are compactly supported wavelets with the least asymmetry and the highest number of vanishing moments for a given support

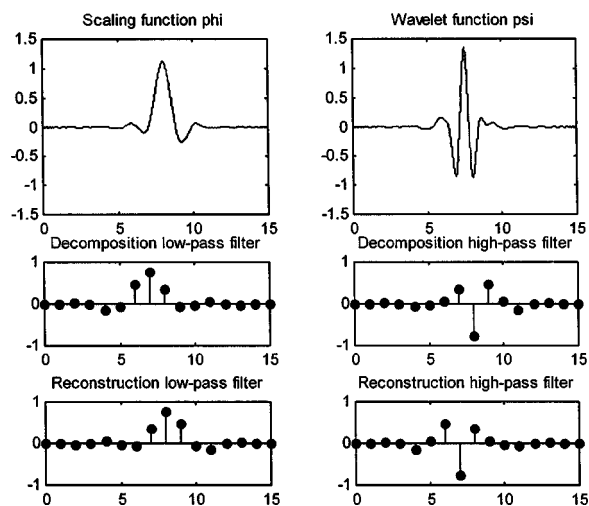


FIG. A1. Symmlet wavelets of order 8; the scaling function and wavelet function are presented along with its associated filters.

width and nearly linear-phase scaling filters. The other properties symmlets inherit from Daubechies wavelets are orthogonality, possibility of realization of Discrete Wavelet Transform, and existence of a fast algorithm. These wavelets have no explicit functional form.

The symmlet wavelet of order 8 is illustrated in Fig. A1 along with its associated scaling function and decomposition/reconstruction filters. The symmlet wavelets were used in wavelet–vaguelette inverse method presented in Sec. IID.

- ¹K. G. Foote, "Fish target strengths for use in echo integrator surveys," J. Acoust. Soc. Am. **82**, 981–987 (1987).
- ²A. Stepnowski, A. C. Gucu, and F. Bingel, "Assessment of the pelagic fish resources in the Southern Black Sea using echo integration and dual-beam processing," Arch. Acoust. **18**, 83–94 (1990).
- ³D. N. MacLennan and E. J. Simmonds, *Fisheries Acoustics* (Chapman and Hall, London, 1992).
- ⁴K. G. Foote, "Summary of methods for determining fish target strength at ultrasonic frequencies," ICES J. Mar. Sci. **48**, 211–217 (1991).
- ⁵G. L. Goddard and V. G. Welsby, "Statistical measurement of the acoustic target strength of live fish," Rapp. P.-V. Reun.-Cons. Int. Explor. Mer **170**, 70–73 (1977).
- ⁶J. E. Ehrenberg, "New methods for indirect measuring the mean acoustic cross-section of fish," FAO Fish. Rep. **300**, 91–97 (1983).
- ⁷J. B. Hedgepeth, "Stock assessment with hydroacoustic estimates of abundance via tuning and smoothed EM estimation," University of Washington Ph. D. Dissertation, Seattle, WA, (1994) [subsequently published in part in ICES J. Mar. Sci. **56**, 36–50 (1999)].

- ⁸A. Stepnowski and R. S. Mitchell, "ECOLOG II: A real-time acoustic signal processing system for fish stock assessment," Ultrasonics **28**, 256–265 (1990).
- ⁹R. E. Craig and S. T. Forbes, "A sonar for fish counting," Fiskeridir. Skr., Ser. Havunders. **15**, 210–219 (1969).
- ¹⁰B. J. Robinson, "In situ measurements of the target strengths of pelagic fish," FAO Fish. Rep. **300**, 99–103 (1983).
- ¹¹M. L. Peterson, C. S. Clay, and S. B. Brandt, "Acoustic estimates of fish density and scattering function," J. Acoust. Soc. Am. **60**, 618–622 (1976).
- ¹²H. Medwin and C. S. Clay, *Fundamentals of Acoustical Oceanography* (Academic New York, 1998).
- ¹³J. E. Ehrenberg, "A review of in situ target strength estimation techniques," ICES/FAO Symposium on Fisheries Acoustics, art. No 104, Bergen (1982).
- ¹⁴C. S. Clay and B. G. Heist, "Acoustic scattering by fish—Acoustic models and a two-parameter fit," J. Acoust. Soc. Am. **75**, 1077–1083 (1985).
- ¹⁵T. K. Stanton, "Volume scattering: Echo peak PDF," J. Acoust. Soc. Am. **77**, 1358–1366 (1985a).
- ¹⁶T. Lindem, "Success with conventional in situ determination of fish target strength," FAO Fish. Rep. **300**, 104–111 (1983).
- ¹⁷C. S. Clay, "Deconvolution of the fish scattering pdf from the echo pdf for a single transducer sonar," J. Acoust. Soc. Am. **73**, 1989–1994 (1983).
- ¹⁸O. H. J. Miinalainen and T. Eronen, "A least-squares method for in situ estimation of fish target strength with a single transducer sonar," Rapp. P.-V. Reun.-Cons. Int. Explor. Mer. **189**, 312–316 (1990).
- ¹⁹L. G. Rudstam, C. S. Clay, and J. J. Magnuson, "Density and size estimates of cisco (*Coregonus artedii*) using analysis of echo peak pdf from a single transducer sonar," Can. J. Fish. Aquat. Sci. **44**, 811–821 (1987).
- ²⁰M. Moszyński, "Inverse filtering of hydroacoustic echo pdf's for target strength estimation," Ph.D. Dissertation (in Polish), Technical University of Gdańsk (1997).
- ²¹M. Moszyński and A. Stepnowski, "Using the discrete Mellin transform (DMT) and singular value decomposition (SVD) for estimation of fish target strength from acoustics echoes," Proceedings of the 3rd European Conference on Underwater Acoustics, Heraklion, Crete, Vol. 2, pp. 733–738 (1996).
- ²²A. N. Tichonov, "Solution of incorrectly formulated problems and the regularization method," Soviet Math. Doklady **4**, 1035–1038 (1963).
- ²³P. C. Hansen, "Numerical tools for analysis and solution of Fredholm integral equations of the first kind," Inverse Probl. **8**, 849–872 (1992).
- ²⁴D. L. Donoho, I. M. Johnstone, G. Kerkycharian, and D. Picard, "Density estimation by wavelet thresholding," The Annals of Statistics, **24**, 508–539 (1996).
- ²⁵O. N. Strand, "Theory and methods related to the singular function expansion and Landweber's iteration for integral equations of the first kind," SIAM (Soc. Ind. Appl. Math.) J. Numer. Anal. **5**, 287–297 (1973).
- ²⁶S. L. Marple, *Digital Spectral Analysis* (Prentice-Hall, New York, 1987).
- ²⁷A. Stepnowski, "Comparison of the novel inverse techniques for fish target strength estimation," Proceedings of the Fourth European Conference on Underwater Acoustics, Rome, Italy, pp. 187–192 (1998).
- ²⁸I. Daubechies, "Ten lectures about wavelets SIAM" (1992).

The matched-phase coherent multi-frequency matched-field processor

Gregory J. Orris, Michael Nicholas, and John S. Perkins
United States Naval Research Laboratory, Washington, D.C. 20375-5200

(Received 26 March 1999; revised 28 January 2000; accepted 1 February 2000)

Coherent multi-frequency matched-field processing is investigated using a matched-phase coherent matched-field processor. Its main difference from previous coherent processors is that the relative phases of the Fourier components contained within the recorded signal are not assumed to be known *a priori*. Rather they are considered free parameters that can be determined using a global functional minimization algorithm. Additionally, this processor uses only the cross-frequency terms, making it less susceptible to the detrimental effects of ambient noise; in one example, this processor shows a five decibel improvement over a similar coherent processor. Along with its increased sensitivity with respect to the broadcast source levels, this coherent processor exhibits superior range resolution as compared with multi-frequency incoherent processors, due to the cross-frequency interference of the vertical eigenmodes. Within this work we explore the efficacy of the algorithms used to determine the relative phases along with the performance of the matched-phase coherent processor itself, performed within the context of data collected during an event from the SWellEx-96 experiment. Performance comparisons between this processor, an incoherent processor, and another coherent processor are demonstrated using this data set. [S0001-4966(00)02705-3]

PACS numbers: 43.60.Gk, 43.30.Wi [DLB]

INTRODUCTION

In this paper we consider broadband, or multi-frequency, coherent matched-field processing for locating an acoustic source. We show that the cross-frequency phase relationships can be considered free parameters that may be determined using global functional minimization methods. This results in a coherent processor with increased range resolution and improved sensitivity to low-level signals even when only a few frequencies are processed coherently.

As the term is most commonly used, matched-field processing (MFP) is a brute force method for passively locating an acoustic source. For many trial source positions (the *fiducial search space*), repeated application of a forward propagation model is used to simulate the received field on a given set of sensors (usually an array with some vertical extent). As data are received, they are compared (matched) to each of the simulated fields (the *replicas*). The matching results are displayed as a spatially dependent function (this is referred to as the “ambiguity function” or “ambiguity surface”) which depicts the degree to which each of the replicas matched the data. If the simulated field is accurate and the signal-to-noise-ratio is not too low, the replica computed for the true source position will correlate well with the data, and there will be a peak in the ambiguity function at that point. Because of the computational expense required to repeatedly and accurately model the field over large search volumes, most MFP algorithms are implemented for a narrow band of frequencies. Bucker is generally credited with introducing the concept of MFP to the underwater acoustics community.¹ For an overview and history of narrow-band matched-field processing see Ref. 2 and the references therein.

More recently, papers on broadband MFP have appeared. These methods should more appropriately be considered multi-frequency algorithms, since usually only a small

number of frequencies and not a continuous band are processed. Baggeroer *et al.*³ recommended a simple incoherent average of single-frequency ambiguity functions. Typically, in a single-frequency ambiguity function, there will be false peaks, referred to as sidelobes—an analogy with plane-wave beamforming. These sidelobes often appear at different positions for different frequencies, so the incoherent average over frequency generally works well to reinforce the main peak at the true source location while reducing the sidelobes. Some recent papers have demonstrated this quite well using at-sea data.^{4,5} However, for many cases and in particular in the analysis of data from the SWellEx-96 exercise, improvements in sidelobe suppression were generally small. For most localization problems it may well turn out that fully coherent broadband approaches will not prove worth the extra effort required.

The first papers on coherent broadband MFP considered time-domain matched-field source localization. Clay⁶ proposed a time-domain method that requires knowledge of the source function. Frazer and Pecholcs⁷ considered the possibility of time-domain source localization using just a single hydrophone and devised methods that did not require knowledge of the source function. However, they assumed the data were limited to a known band of frequencies. They pointed out that there exists a tradeoff between narrow-band/large aperture and wide-band/small aperture. Frazer and Pecholcs went further to discuss how their processors can be improved if the source spectrum (amplitude only or amplitude and phase) is known, and how to extend their ideas to multiple hydrophones. Although Brienzo and Hodgkiss⁸ relied on frequency-domain modeling, they developed a generalized conventional beamformer that used simulated channel im-

pulse responses as matched-filters and successfully applied it to shot data.

Tolstoy⁹ and Westwood¹⁰ also proposed coherent frequency-domain broadband processors. Westwood demonstrated the performance of his processor on at-sea data, and his processor has recently been applied by Knobles and Mitchell¹¹ to data from a large aperture (1-km) three-element array. The Westwood broadband processor does not use the phase of the source spectrum in any way, while the Tolstoy processor tacitly assumes that the phase relationship is obtainable via some other means. Similarly, Yang¹² has proposed a coherent broadband method in the context of matched-mode processing that assumes the original source spectrum is known.

Generally, in passive matched-field processing one does not have exact knowledge of either the amplitude or the phase of the source spectrum: The latter might be particularly difficult to determine. A coherent broadband technique that ignores the phase of the source spectrum in fact makes a tacit assumption about the phase, in essence, that it is trivially constant across all processed frequencies. If that assumption is incorrect then such an algorithm might suffer more degradation in performance than a reliable incoherent procedure. To address this problem Michalopoulou and Porter¹³ removed the phase differences between frequencies by normalizing the data by the phase of the first hydrophone. In Ref. 14 Michalopoulou proposed a more complex phase normalization-estimation scheme to overcome the unknown source spectrum difficulty.

We propose to search for the phase factors as additional unknowns in a focalization type problem. This idea is discussed by Michalopoulou¹⁴ but dismissed as being too computationally burdensome. This may generally be the case, yet significant improvements might be possible by processing in a fully coherent manner, only two or three frequencies. We wish to point out that if the position of the source is known, then replicas can be generated as a function of the other physical parameters, for example, geoacoustic, water column sound speed, etc. Such techniques have been called matched-field inversion.¹⁵ Thus a phase-sensitive coherent broadband (or multi-frequency) matched-field technique might prove useful in these types of problems.

We will presently develop the MFP notation used throughout the remainder paper. Then we briefly describe a common incoherent MFP algorithm to be used as a baseline for further comparison and then transition to the coherent processor as used to analyze the SWellEx-96 data. We end the next section by describing the search technique used to obtain the relative phases required by the coherent algorithm. The SWellEx-96 experiment is described and the MFP techniques to be presently developed are applied to the data. Finally, we conclude with some discussion about the relative effectiveness of the matched-phase coherent algorithm versus other coherent and incoherent methods.

I. MULTI-FREQUENCY MATCHED-FIELD PROCESSING

Within this section the notation used throughout the remainder of the paper is defined. The matched-phase MFP

algorithm is presented as a natural development and extension of the single-frequency Bartlett MFP processor. We start by considering MFP as a generic “inversion” problem: Given a recorded signal at many locations, “What is the most likely configuration of independent variables that is consistent with the data?” We present this problem within the context of a multi-frequency algorithm. Throughout it is assumed that the acoustic field can be accurately modeled with a fluid–fluid acoustic propagation model. This section ends with the presentation of the idea that under some rather broad considerations knowledge of the source spectrum, specifically the inter-frequency complex phase relationships, heretofore regarded as unusable, unknowable, or unreliable, may be regarded as free parameters.

A. The matched-phase MFP coherent processor

The $N+d$ -dimensional search space X^{N+d} , where there are N unknown nonspatial parameters and d spatial dimensions is searched by assigning a functional value to each point in X^{N+d} , through a mapping function called the “processor.” This real-valued function denoted by \mathcal{P} must necessarily be a function of the best estimate of the field at the receiver locations, as well as a function of the data collected at these locations. To assist in the analysis one typically defines the complex-valued “replica” vector $\tilde{\mathbf{p}}$ and the “data” vector $\tilde{\mathbf{d}}$. The “replica” vector is formed from the Fourier transformed pressure field, $\tilde{p}(\omega, \mathbf{x})$ sampled at the M receiver locations $\{\mathbf{x}_i\}_{i=1}^M$, where $\tilde{p}(\omega, \mathbf{x})$ is a modeled solution to the acoustic wave equation for a source at \mathbf{x}_s emitting acoustic energy at radial frequency ω with source strength $\tilde{S}(\omega)$, i.e.,

$$\rho \nabla \cdot \left(\frac{1}{\rho} \nabla \tilde{p} \right) + \frac{\omega^2}{c^2} \tilde{p} = \tilde{S}(\omega) \delta(\mathbf{x} - \mathbf{x}_s). \quad (1)$$

Here ρ and c are the spatially dependent density and sound speed, respectively. Similarly, the data collected at the M receiver locations are processed to form the data vector $\tilde{\mathbf{d}} = \{\tilde{d}_i(\omega)\}_{i=1}^M$ by Fourier transforming the respective time series. The processor then correlates (or matches) the two vectors $\tilde{\mathbf{p}}$ and $\tilde{\mathbf{d}}$, representing the modeled pressure and the data, respectively. It is evident that \mathcal{P} is explicitly a function of the data and model, and implicitly a function of the model parameters and spatial coordinates.

The single-frequency Bartlett processor is a particularly simple example of an MFP processor that has several multi-frequency implementations. It is defined as³

$$\begin{aligned} \mathcal{P}_B(\omega, \mathbf{x}) &= \frac{\langle |\tilde{\mathbf{d}}^\dagger(\omega) \cdot \tilde{\mathbf{p}}(\omega, \mathbf{x})|^2 \rangle_T}{N_B(\omega, \mathbf{x})} \\ &= \tilde{\mathbf{u}}^\dagger(\omega, \mathbf{x}) \cdot \mathbf{K}(\omega) \cdot \tilde{\mathbf{u}}(\omega, \mathbf{x}), \end{aligned} \quad (2)$$

where we denote $\langle \dots \rangle_T$ as the time average and \mathbf{x}^\dagger as the Hermitian transpose of the column vector \mathbf{x} . The normalization factor $N_B(\omega, \mathbf{x})$ that sets $\mathcal{P}_B = 1$ when $\tilde{\mathbf{d}} \propto \tilde{\mathbf{p}}$ has been absorbed into the definition of the normalized replica vector $\tilde{\mathbf{u}}$ and the cross spectral density (CSD) matrix \mathbf{K} . The CSD matrix, $\mathbf{K}(\omega)$, is defined as the normalized, time-averaged outer-product of the data vector with its Hermitian conjugate,

$$\mathbf{K}(\omega) = \frac{\mathbf{D}(\omega)}{\text{tr}[\mathbf{D}(\omega)]}, \quad (3)$$

where the “data matrix,” $\mathbf{D}(\omega)$ is defined as

$$\mathbf{D}(\omega) = \langle \tilde{\mathbf{d}}(\omega) \otimes \tilde{\mathbf{d}}^\dagger(\omega) \rangle_T, \quad (4)$$

and $\text{tr}[\mathbf{D}]$ is the trace of the matrix \mathbf{D} . Here we have introduced the notation $\mathbf{x} \otimes \mathbf{y}^\dagger$ as the outer product of the vectors \mathbf{x} and \mathbf{y}^* , where the resultant $\{i, j\}$ matrix element is $x_i y_j^*$ and y^* is the complex conjugate of y (see, e.g., Ref. 16, p. 73).

A straightforward method for extending $\mathcal{P}_B(\omega, \mathbf{x})$ to the multi-frequency regime is the incoherent average; an average of single-frequency MFP ambiguity functions in decibels,³

$$\begin{aligned} & \frac{1}{L} \sum_{n=1}^L 10 \log_{10} \mathcal{P}_B(\omega_n, \mathbf{x}) \\ &= 10 \log_{10} \left[\left(\prod_{n=1}^L \mathcal{P}_B(\omega_n, \mathbf{x}) \right)^{1/L} \right]. \end{aligned} \quad (5)$$

This multi-frequency incoherent processor is seen to be equivalent to the geometric average of the individual single-frequency processors in linear units. This is to be noted as distinct from the linear average, where the single-frequency results are averaged before converting to decibel units. However, there are strong theoretical arguments for using the former method in incoherent processors.¹⁷

One of the most important features of incoherent processors is that the normalization condition is applied to each frequency independently. This results in the contribution from each processed frequency having an intrinsically equal weight in the sum of Eq. (5). This choice of normalization condition has a highly desirable characteristic: Localization via measurement of the relative field strengths at different receiving locations does not depend on any *a priori* knowledge of the frequency spectrum emitted by the source. The drawback of this normalization is that there are no free parameters upon which enhancement of the algorithm can be based, since all unknown parameters pertaining to the source and receiver other than locations are normalized out of the processor.

The coherent matched-field processor that most naturally follows from Eq. (2) is defined as

$$\mathcal{P}_C(\mathbf{x}) = \left[\left\langle \left| \frac{1}{L} \sum_{n=1}^L \frac{\tilde{\mathbf{d}}^\dagger(\omega_n) \cdot \tilde{\mathbf{p}}(\omega_n, \mathbf{x})}{N_C(\omega_n, \mathbf{x})} e^{i\varphi(\omega_n)} \right|^2 \right\rangle_T \right]. \quad (6)$$

We note that this processor is similar to that presented by Tolstoy.¹⁸ However, we prefer to explicitly include the $\varphi(\omega_n)$'s within the definition. The presence of the φ 's is used to properly account for the nonzero phase difference between frequencies within the propagation model used to create the replica field. The normalization factor $N_C(\omega_n, \mathbf{x})$, in a similar fashion to the Bartlett processor's normalization, consists of two factors that normalize the replicas and the data vector independently:

$$N_C(\omega_n, \mathbf{x}) = \sqrt{\text{tr}[\mathbf{D}(\omega_n)]} \sqrt{\tilde{\mathbf{p}}^\dagger(\omega_n, \mathbf{x}) \cdot \tilde{\mathbf{p}}(\omega_n, \mathbf{x})}. \quad (7)$$

The most natural generalization of the cross-spectral density matrix \mathbf{K} to include cross-frequency components is

$$\mathbf{K}(\omega_i, \omega_j) = \mathbf{K}_{ij} = \frac{\langle \tilde{\mathbf{d}}(\omega_i) \otimes \tilde{\mathbf{d}}^\dagger(\omega_j) \rangle_T}{\sqrt{\text{tr}[\mathbf{D}(\omega_i)]} \sqrt{\text{tr}[\mathbf{D}(\omega_j)]}}. \quad (8)$$

Now Eq. (6) can be written as

$$\mathcal{P}_C(\mathbf{x}) = \left\langle \left| \frac{1}{L} \sum_{n=1}^L e^{i\varphi(\omega_n)} \frac{\tilde{\mathbf{d}}^\dagger(\omega_n) \cdot \tilde{\mathbf{u}}(\omega_n, \mathbf{x})}{\sqrt{\text{tr}[\mathbf{D}(\omega_n)]}} \right|^2 \right\rangle_T \quad (9a)$$

$$= \frac{1}{L^2} \sum_{m,n=1}^L e^{-i(\varphi(\omega_m) - \varphi(\omega_n))} \tilde{\mathbf{u}}_m^\dagger \cdot \mathbf{K}_{m,n} \cdot \tilde{\mathbf{u}}_n \quad (9b)$$

$$= \frac{1}{L^2} \left[\sum_{m=1}^L \tilde{\mathbf{u}}_m^\dagger \cdot \mathbf{K}_{m,m} \cdot \tilde{\mathbf{u}}_m + \sum_{n \neq m} e^{-i(\varphi(\omega_m) - \varphi(\omega_n))} \tilde{\mathbf{u}}_m^\dagger \cdot \mathbf{K}_{m,n} \cdot \tilde{\mathbf{u}}_n \right], \quad (9c)$$

where we again use the normalized replica $\tilde{\mathbf{u}} = \tilde{\mathbf{p}}/|\tilde{\mathbf{p}}|$, and introduce the subscript frequency notation $\tilde{\mathbf{u}}_n = \tilde{\mathbf{u}}(\omega_n, \mathbf{x})$.

Consider the two distinct types of terms in Eq. (9c) separately: In the event that any two processed frequencies ω_m and ω_n are uncorrelated, the time average in Eq. (9) will generally result in $|\mathbf{K}_{m,n}| \ll |\mathbf{K}_{n,n}|$, for $n \neq m$. The coherent processor Eq. (9) then amounts to the linear average of single-frequency Bartlett processors divided by the processed bandwidth. Conversely, in cases where there is a significant cross-frequency correlation, the magnitude of the cross-frequency term in Eq. (9c) can become comparable to that of single-frequency term. Unlike the former term of (9c) whose spatial structure is dependent only upon the propagational characteristics of the ocean waveguide at each separate frequency (i.e., mono-tonal modal interference), the spatial structure of cross-frequency term is dependent upon pairs of discrete frequencies (i.e., cross-frequency modal interference). This results in the spatial structure of the ambiguity function being dominated by the horizontal interference associated between pairs of dominant eigenmodes of the processed frequencies, in addition to the more commonly encountered vertical interference associated with mono-tone propagation. For an appropriate choice of processed frequencies, this interference pattern can have the important effect of putting the range resolution on par with that of the vertical resolution. One can obtain a reasonable estimate of the length scale of this interference pattern from the vertical eigenvalue spectrums $\{k_n(\omega_1)\}_{n=1}^\infty$ and $\{k_m(\omega_2)\}_{m=1}^\infty$ associated with frequencies ω_1 and ω_2 , respectively. The interference pattern will be characterized by the wavelength $\lambda_{1,2}$ roughly corresponding to

$$\lambda_{1,2} \sim \frac{2\pi}{|\Re(k_{n_{\max}}(\omega_1) - k_{m_{\max}}(\omega_2))|}, \quad (10)$$

where $k_{n_{\max}}(\omega_1)$ corresponds to the vertical eigenvalue associated with the eigenmode that has the largest modal coefficient for the given environmental parameters, and $\Re(x)$ signifies the real part of x .

We seek to find an improved coherent algorithm to Eq. (6) for use in low SNR scenarios. We revisit the question of which normalization factor is the appropriate one to apply to the terms of a coherent processor. Consider the normalization factor of Eq. (6) as defined in Eq. (7). Our choice of the normalization factor N_C leads to the desired result when the cross-frequency correlation, $\langle \tilde{\mathbf{d}}(\omega_n) \otimes \tilde{\mathbf{d}}^\dagger(\omega_m) \rangle$ for $n \neq m$, is of comparable magnitude to that of the mono-tone correlation, $\langle \tilde{\mathbf{d}}(\omega_n) \otimes \tilde{\mathbf{d}}^\dagger(\omega_n) \rangle$. This choice of N_C is likely to cause an underestimation of the magnitude of the cross-frequency component in comparison to the mono-tone component from the source under low SNR conditions. This is because typical ambient noise sources have decorrelated cross-frequency components but correlated mono-tone components. Thus if we may assume that the sources we desire to detect are radiating acoustic energy in a manner such that the cross-frequency components are well correlated, then we are naturally lead to the definition of the “matched-phase” coherent processor,

$$\mathcal{P}_M(\mathbf{x}) = \frac{1}{L(L-1)} \sum_{n \neq m} e^{-i(\varphi(\omega_m) - \varphi(\omega_n))} \tilde{\mathbf{u}}_m^\dagger \cdot \mathbf{K}_{m,n} \cdot \tilde{\mathbf{u}}_n. \quad (11)$$

This is essentially Eq. (6) without the mono-tone components and normalized so that a perfect match will result in unity output of the processor. By neglecting the mono-tone components it will be seen in the next section that Eq. (11) can out perform Eq. (9) by as much as five decibels (peak-to-sidelobe ratio) at low SNR, while still performing favorably at high SNR.

B. Determination of relative phases

It is likely that even in the best of circumstances, the exact phase broadcast by the source is not obtainable due to the presence of ambient noise, relative movement of the source and receiver, and nonlinearities in the source itself. Matters are further complicated by the fact that many sources which are encountered, e.g., breaking waves, bubble clouds, shipping, etc., do not radiate with a time-independent coherent phase relationships. It may also be that the data could have been substantially modified by any preprocessing performed before it is stored on permanent media and later analyzed. Thus it is our belief that searching for or requiring knowledge of the *exact* source function may often be of dubious benefit. This would make the onerous requirement that the MFP algorithm be capable of incorporating the distortion of the phase information due to the signal processing of the entire system, including the unknown noise component. Nevertheless, a truly coherent multi-frequency linear matched-field processor must account for the possibility that there exists a nontrivial, nonstationary but relatively slowly varying complex phase structure between the frequencies recorded on the hydrophone array.

In light of this we propose to determine a set of relative phases *a posteriori* satisfying a set of criteria based on the performance of the MFP processor, instead of searching for the exact phases emitted by the source. Typically a function based on Eq. (11) is minimized, subject to some set of con-

straints that are consistent with the physics of the particular data set. Without knowledge of the source characteristics or noise field it is natural to use \mathcal{P}_M or \mathcal{P}_C , as defined above, as the function to be maximized over the variables of the search space. This choice makes the tacit assumption that the peak value of the ambiguity function is the most likely location of the source, all other variables kept constant: This is a questionable assumption under low SNR conditions. If the phases are varied until the peak attains its maximum value, the phases of the modeled replicas are said to be “matched” to the *data*. Note that this does not necessarily imply that we have determined the exact source function.

Given that $\max \mathcal{P}_M = 1$, this problem is equivalent to minimizing $1 - \mathcal{P}_M$, i.e.,

$$\mathcal{P}_M(\mathbf{x}_{\max}) = 1 - \min_{\substack{\{\varphi_i \in [0, 2\pi) \forall i=1, N_f\} \\ \mathbf{x} \in E^d}} [1 - \mathcal{P}_M(\mathbf{x})], \quad (12)$$

where E^d is the d -Euclidean space comprising the fiducial search region (i.e., for the specific cases presented in the next section E^d is the volume of the ocean to be searched). This minimization only makes sense if it is performed simultaneously with respect to the both the φ 's and the peak location \mathbf{x} . In Eq. (12) we have implicitly defined the vector \mathbf{x}_{\max} as the location of the maximum peak of the set over the set of all possible ambiguity surfaces. Similarly, we define the vector $\mathbf{x}_{\max}^{(i)}$ as the maximum of the ambiguity surface for the i th realization of phases.

For only a few frequencies ($N_f \leq 3$) it is possible to sample the search space directly with enough density to provide a reasonable estimate of the ambiguity function's dependence upon the relative phases. Under these circumstances the free parameters in Eq. (12) can be determined via inspection. Unfortunately, higher-dimensional search spaces are too complex to efficiently search with exhaustive techniques and additionally are not convenient to render: They require a global search methods. Further complicating matters, as will be demonstrated in the next section, the parameter landscape corresponding to the search space of the phases may contain many local minima. Thus one would expect that many standard methods of functional minimization would be rendered ineffective. To address these issues we have implemented a variant of the global search method known as simulated annealing.^{16,19,20}

In our simulated annealing implementation the initial and final values of the temperature (T_0 and T_f), and an initial set of the $N_f - 1$ unconstrained relative phases,

$$\varphi_{1,m} = \varphi(\omega_1) - \varphi(\omega_m), \quad 2 \leq m \leq N_f, \quad (13)$$

are given as initial conditions. The remaining $N_f(N_f - 3)/2 + 1$ combinations of phases are constrained, as they are constructed from the various combinations of the first $N_f - 1$. At each iteration of the search algorithm a new set of φ 's is determined. Equation (12) is used as the energy function directly with these phases. It is compared with the currently accepted value of the energy. If the new iteration's phases produce a better solution to Eq. (12), then the new phases are accepted as the global phase solution. If not, then a uniform random number (denoted by R) between zero and one is

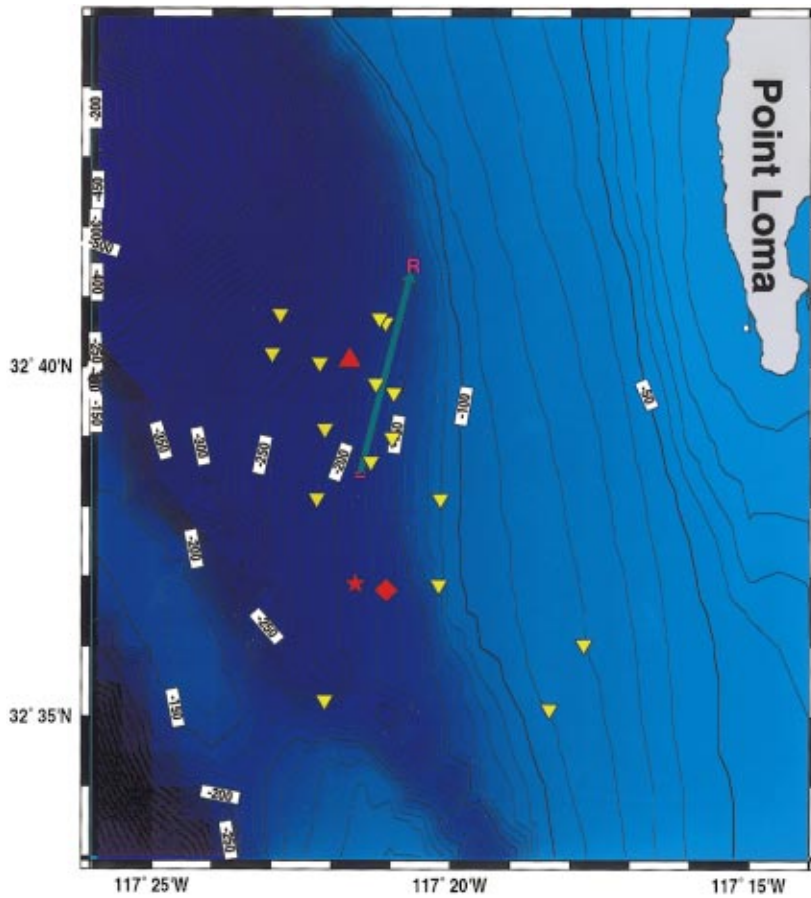


FIG. 1. This bathymetry chart of the experimental region of SWellEx-96 shows the location and relative orientation of the two main experimental arrays, R/V FLIP (large triangle) and NRL's SVLA system. The location of this system is given in two parts: the array (star) and the buoy (diamond) containing the instrumentation. The various way points used throughout the experiment are denoted by their actual call letters on the chart. The small inverted triangles correspond to positions where CTD measurements were conducted. The arrow indicates the track of R/V SPROUL during this event. Contour intervals are in 10-m increments.

generated. If R is less than or equal to the Boltzmann exponential factor, i.e.,

$$R \leq \exp \left[- \frac{(1 - \mathcal{P}_M(\mathbf{x}_{\max}^{(i)}))}{k_B T_i} \right], \quad (14)$$

where T_i is the temperature at iteration i and k_B is a constant relating the expected value of the numerator to the temperature, then the new phases were chosen, despite having a higher overall energy (and thus not being the true solution). Finally, the temperature are lowered via a predefined cooling schedule, and the whole process is repeated until $T_i \leq T_f$.

It is well known that this procedure will correctly locate the minimum of the function [e.g., Eq. (12)] with probability one, provided that the cooling schedule is sufficiently gradual and that all the independent variables are sampled uniformly at each iteration of the algorithm.²⁰ Unfortunately, this schedule is almost always too slow to be of practical use for problems with a large number of unknowns. The search may be sped up using the cooling schedule

$$T_i = \alpha T_{i-1}, \quad (15)$$

where $\alpha < 1 - \epsilon$, where $\epsilon \ll 1$. Using Eq. (15) speeds convergence of simulated annealing, but the price to be paid for that speed is that a much smaller subspace of phases are sampled. Thus we are not guaranteed of determining the absolute minimum in Eq. (12). This results in the algorithm converging to what is termed a "false minimum." To address this at each iteration of the search algorithm our implementation randomly chooses a set perturbations $\delta\varphi$ to the set of φ 's

from a *normal* distribution with zero mean and standard deviation $\sqrt{2\pi}$. This particular distribution samples points far enough from the current location to achieve the required effect of being able to tunnel out of local minima with sufficient probability. Additionally this distribution results in a small enough average displacement that the process more closely resembles a random walk over the parameter landscape rather than a scattered sampling of the environment. This adjustment to the simulated annealing algorithm usually equates to an exponential increase in convergence time as compared to determining the φ 's via a uniform distribution.

II. APPLICATIONS TO THE SWellEx-96 EXPERIMENT

In this section we present, through direct application to data taken during the SWellEx-96 experiment, analysis demonstrating the efficacy of the coherent matched-phase MFP algorithm. We begin this section by describing the experiment, data collection, and the specific source track used for the analysis that follows. Then we apply the incoherent multi-frequency processor, *a la* Eq. (2) and the coherent processor Eq. (9), to several of the tones broadcast by the source during the processed event. This is done to supply the reader with a baseline to which the matched phased algorithm Eq. (11) is to be compared. For both the coherent algorithm and the matched-phase algorithm, simulated annealing is used to obtain the relative phases when the number of frequencies processed exceeds three. Comparisons to the matched-phase MFP algorithm will be made by presenting all relevant ambiguity functions in decibels. Each has been normalized to

TABLE I. Listing of the way points used to identify the ship track corresponding to event S19, along with the locations of R/V FLIP, the SVLA buoy and SVLA array. See Fig. 1 for geographical orientation.

MFP way point	Latitude N	Longitude W
R	32° 41' 27"	117° 20' 37.8"
=	32° 38' 29.4"	117° 21' 31.2"
SVLA Array	32° 36' 54"	117° 21' 36"
SVLA Buoy	32° 36' 20"	117° 21' 5"
R/V FLIP	32° 40' 15"	117° 21' 38"

the peak value in the figure to represent the same dynamic range. Finally, we discuss the performance of the search algorithms used to estimate the relative phases of the source.

A. The SWellEx-96 experiment

The SWellEx-96 matched-field experiment was a shallow water experiment conducted off the coast of San Diego, California during the month of May 1996. Participants included elements from the Naval Research and Development Warfare Center (now SPAWAR Systems Center, San Diego), the Marine Physics Laboratory (MPL) of the University of California's Scripps Institute of Oceanography (SIO), and the Naval Research Laboratory (NRL). Figure 1 shows a chart of the fiducial region of SWellEx-96 with the location and relative orientation of the two main experimental arrays of the experiment: R/V FLIP and NRL's Satellite-linked Vertical Line Array (SVLA) system, along with the way points used for the event corresponding to the data analyzed. Throughout the exercise the array component of the SVLA was moored at 32°36'54" N Latitude, 117°21'36" W Longitude, in approximately 193 m of water. The SVLA system consisted of an autonomous array of 32 hydrophones with 5-m spacings with the first phone being 5 m from the ocean floor.

SIO's R/V SPROUL was used to tow two sources at different depths along various tracks designed to experimentally investigate the effects of a shallow water spatially dependent waveguide on MFP. The ship track corresponding to the data used in this paper was designated event S19 and commenced at 0802GMT on May 11, 1996. This track started at way point "=" (approximately 2650 m due North of the SVLA array component) and proceeded directly to way point "R" (see Fig. 1 and Table I). This track as can be seen from Fig. 1, lies fortuitously close to an iso-bathymetric contour line, from the SVLA's array component.

The depth of the source that emitted the signal analyzed here was to have nominally remained at 55 m throughout the event. As can be seen from Fig. 2 some deviation from this depth was encountered during the period of time R/V SPROUL was beginning to get under way. The source was calibrated and monitored continuously throughout the experiment via a monitoring hydrophone placed on the tow cable approximately 1 m from the source. During the experiment several sets of frequencies were broadcast. The tones broadcast during event S19 consisted only of the five groups of tones associated with what were called the "pilot tones" whose frequencies were at 112 Hz, 130 Hz, 148 Hz, 166 Hz, and 201 Hz (see Table II). Each pilot tone was simulta-

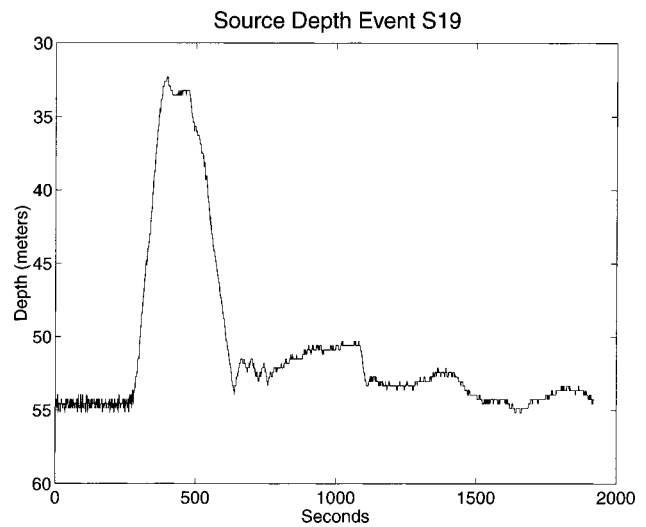


FIG. 2. Source depth as a function of time during event S19. The time axis has been set relative to 08:02:00 GMT which corresponds to the zero of time shown in Fig. 3(b).

neously broadcast with three higher-frequency subtones at 3-Hz intervals. The first nonpilot tone was 3 Hz higher in frequency and broadcast at a source level 25 dB lower than the pilot tone; the next two were at source levels 28 dB and 31 dB lower than the pilot tones and 6 Hz and 9 Hz higher in frequency, respectively. Due to *in situ* tuning, levels between the pilot tones varied slightly throughout SWellEx-96, although not substantially within any given event as can be seen from the stability of lines in Fig. 3(b).

The data collected on the SVLA were immediately base banded and filtered to a 256-Hz window with an adjustable center frequency. Throughout the experiment the center frequency was maintained at 128 Hz, giving the operational band of frequencies from 0 Hz to 255 Hz. After the filtering, the data were transmitted through a fiber-optic cable to the buoy component of the SVLA system where it was temporarily stored on hard disk. At that point the data were both saved to 4-mm DAT tape and telemetered via satellite link and/or line-of-sight rf link to a shore station present at Point Loma 10 km away. This system achieved data transfer rates routinely approaching one Megabit per second, which proved fast enough to allow for extensive *in situ* data pro-

TABLE II. Frequencies of broadcast tones used throughout event S19 of SWellEx-96 along with the source level as measured at the monitoring hydrophone and the estimated SNR levels on the SVLA array.

Broadcast tones for SWellEx-96					
Frequency (Hz)	112	130	148	166	201
Source level (dB)	155	156	147	150	138
SNR at array (dB)	15.9	15.7	12.9	12.5	13.8
Frequency (Hz)	115	133	151	169	204
Source level (dB)	130	131	122	125	115
SNR at array (dB)	-9.1	-9.3	-12.1	-12.5	-11.2
Frequency (Hz)	118	136	154	172	207
Source level (dB)	127	128	119	122	112
SNR at array (dB)	-12.1	-12.3	-15.1	-15.5	-14.2
Frequency (Hz)	121	139	157	175	210
Source level (dB)	124	125	116	119	109
SNR at array (dB)	-15.1	-15.3	-18.1	-18.5	-17.2

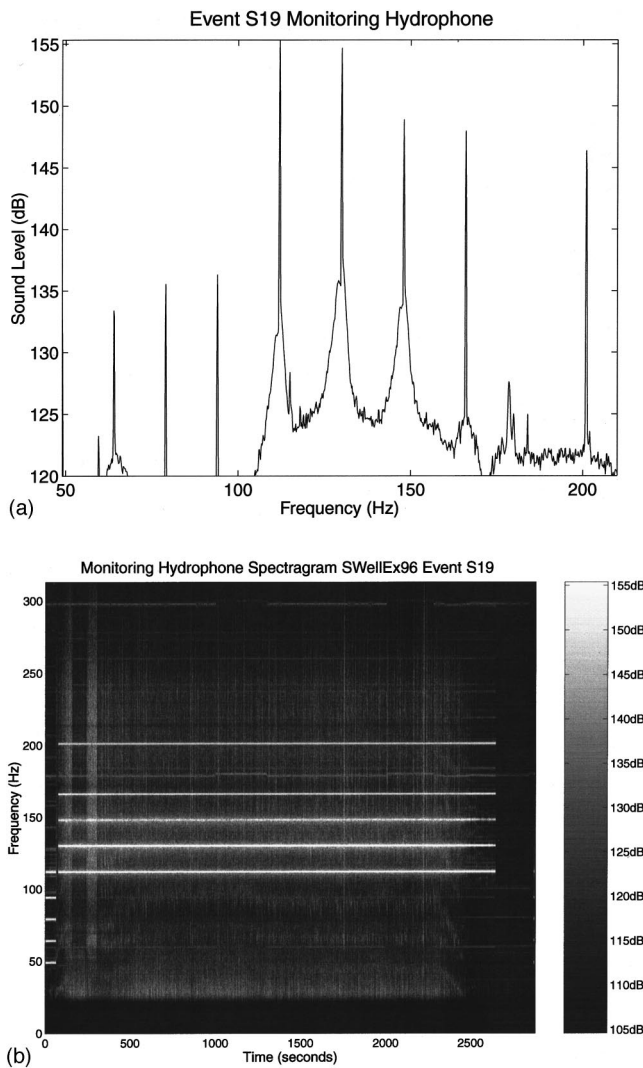


FIG. 3. In (a) we show a calibrated spectrum computed from data collected on the monitoring hydrophone for the duration of event S19. (b) is the spectrogram of the same event. The event immediately preceding this was a stationary event where the source was maintained at a constant range and depth, and with a different set of tones being broadcast. The event immediately following this was a stationary event with no source in order to estimate the ambient noise levels present during the experiment. This is evident as the pilot tones disappear after approximately 2600 s.

cessing. The data were then divided into 2-s records, windowed with a standard Hanning window, whereupon a single FFT of length 512 of the data was taken and the results averaged over 15 equal contiguous nonoverlapping intervals, to produce the time-averaged outer product of the data vectors. This provided the output FFT with enough resolution (0.5 Hz) to give a reasonable signal-to-noise ratio (SNR), yet

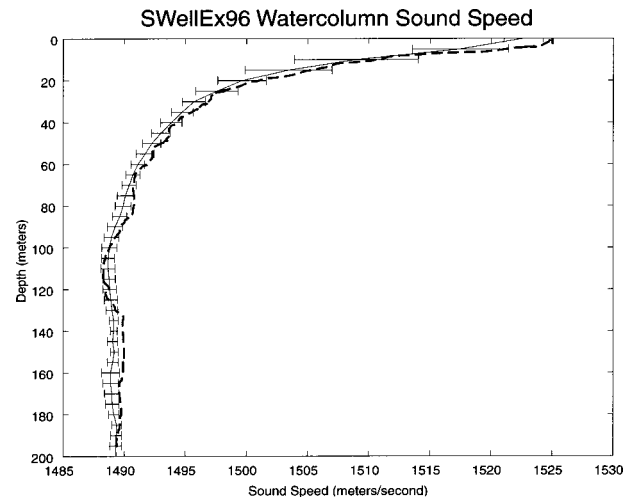


FIG. 4. Mean sound speed as a function of depth from the 18 measurements taken during the course of the exercise. Also plotted are the error bars corresponding to one standard deviation. Larger error bars near the surface (≤ 30 m) indicate some diurnal heating effects. The dashed line is the specific soundspeed profile used for simulations.

large enough to be free of worry from Doppler shifts while R/V SPROUL was under way.

Estimates of the SNR were performed in a manner similar to that of Booth *et al.*⁵ An average of the power over six frequency bins at both higher and lower frequencies than the pilot tones not associated with the known emissions from the source, yet close to each pilot tone was taken (in this case 30 in total). This result was then averaged over the entire array, along with a similar average at each of the pilot tone frequencies:⁵

$$\text{SNR}(\omega_{\text{pilot}}) \approx 10 \log_{10} \left[\left(\frac{\sum_{m=1}^M P_m(\omega_{\text{pilot}})}{(1/N_n) \sum_{k=1}^{N_0} \sum_{m=1}^M P_m(\omega_k)} \right) - 1 \right], \quad (16)$$

where $P_m(\omega)$ is the power at frequency ω on hydrophone m , N_n the number of noise frequencies sampled, and M the number of receivers. The estimates so calculated for the pilot tones are also listed in Table II, and range from +12.5 dB to +15.9 dB. Thus we expect that tones broadcast with source levels 25 dB lower, i.e., the first sub-tones, should have resulted in SNR's at the array of from -12.5 dB to -9 dB. We note that archival data suggest that the ambient noise in this region at 80 dB to 85 dB. Given the time of day and the calm seas during this event it is likely to have been closer to the lower limit.

TABLE III. Environmental parameters used for the sediment within the forward propagation model. These parameters were found using environmental inversion techniques with *in situ* data over a small region round the SVLA array component.

Sediment layer	Environmental inversion sediment parameters					
	Material	Depth (m)	Top sound speed (m/s)	Sound speed gradient (m/s/m)	Attenuation (dB/ λ)	Density (gr/cm ³)
Top	Sand	30	1572	0.6883	0.3	1.76
Secondary	Sandstone	800	1881	1.7	0.09	2.06
Basement	Basalt	N/A	5200	0	0.03	2.66

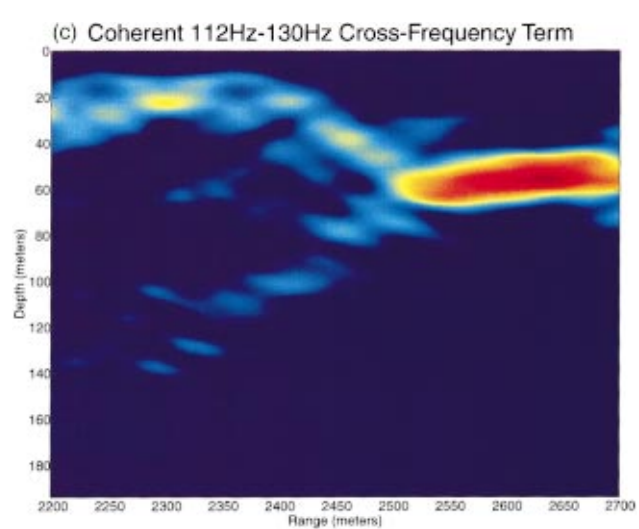
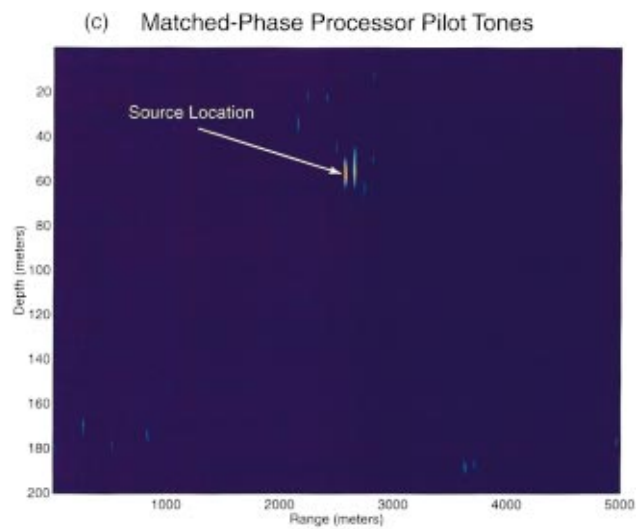
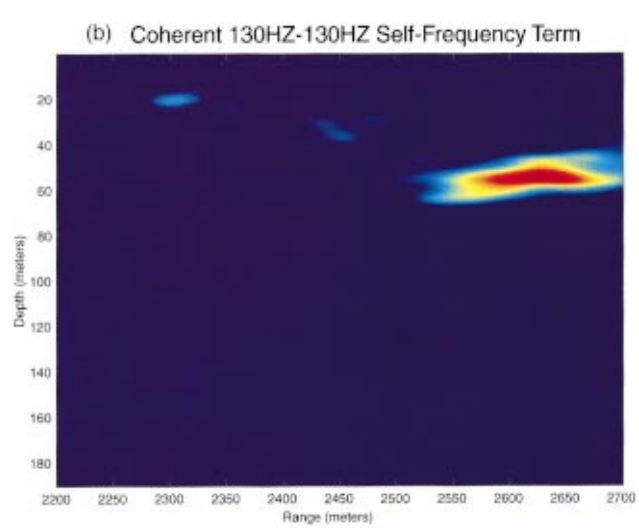
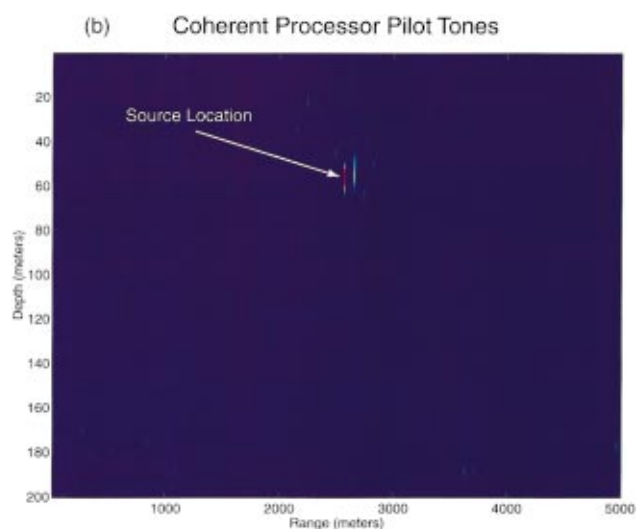
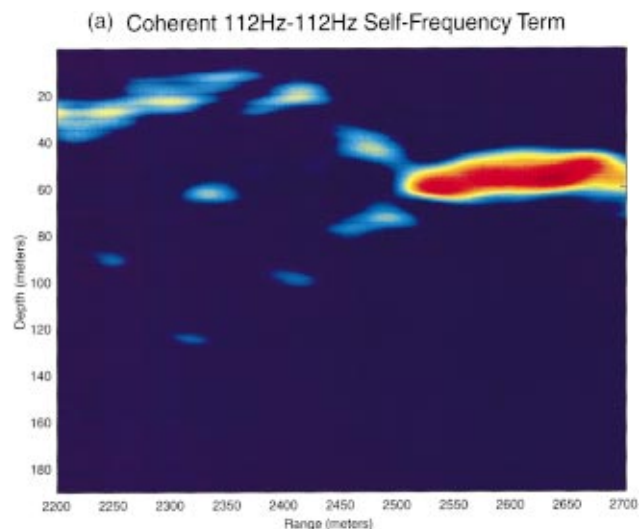
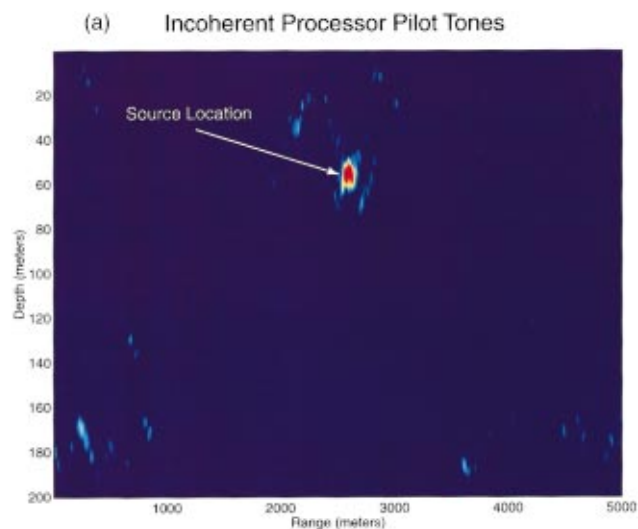


FIG. 5. In this set of figures we present results obtained by applying the incoherent Bartlett, coherent, and matched-phase multi-frequency processors to the pilot tones. The dynamic range in all figures is 6 dB from the peak value. Use of the coherent algorithms has reduced the small sidelobes of the incoherent processor to be nonexistent. The range resolution is 10 m and the depth resolution is 1 m. Note that the coherent algorithms are at the limit of their range resolution.

FIG. 6. An enlargement of the ambiguity surface close to the true source position for 112-Hz and 130-Hz pilot tones. Range resolution has been increased to 1 m. In (a) and (b) we have plotted the squared term from Eq. (9c), while in (c) the term corresponding to cross-frequency component of Eq. (9c) is plotted. These plots are in decibels and the dynamic range of the color table has been increased to the top 10 dB to enhance the sidelobes on either side of the main source peak.

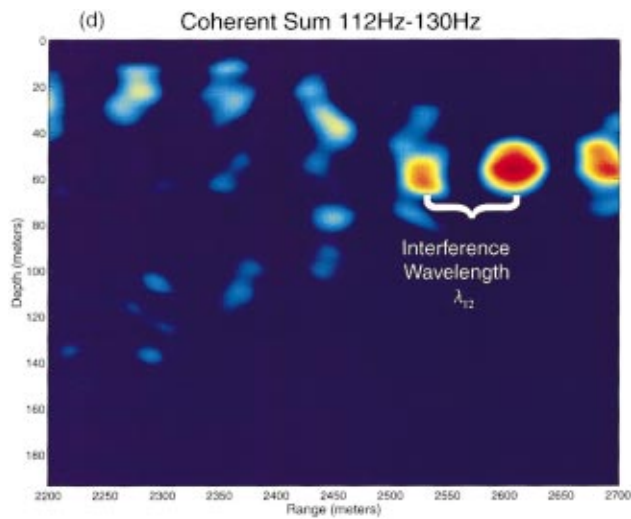


FIG. 6. (Continued.)

Figure 1 also shows the 18 locations where CTD measurements were carried out during the course of the experiment (these are denoted by the inverted triangles). The mean calculated sound speed profile from these measurements are plotted in Fig. 4 along with the specific profile used in the simulations (dashed line). Also plotted in this figure are error bars corresponding to a single standard deviation for the isobaric data. As can be seen from Fig. 4 the profile was downward refracting and had slight variations in the top 30 m of the watercolumn, most likely caused by diurnal heating of the top layers of the ocean.

The sediment in the region was determined via environmental inversion techniques currently under development at NRL²¹ and correlates well with archival data obtained from MPL.²² Near the SVLA array it was determined that there were two main layers of sedimentation. The top layer was 30 m thick and composed of sand with an approximate density of 1.76 gr/cm³. A second layer of a denser substrate being 800 m thick and with a density of approximately 2.06 gr/cm³ was also present. (See Table III for the exact values used in the data analysis.) Also, these environmental inversions reproduced the array tilt that was periodically sampled during the course of the experiment: Array tilts were generally observed to be less than one degree from vertical. For simplicity our analysis presented below assumed the array to be vertical.

TABLE IV. First six eigenmode phase speeds for frequencies 112 Hz and 130 Hz. Environmental parameters were derived from Table III and the soundspeed profile shown in Fig. 4. Eigenvalues were determined by the normal mode model KRAKEN (Ref. 23).

Eigenvalue	Phase speed (m/s)	
	$\nu_1 = 112$ Hz	$\nu_2 = 130$ Hz
1	1489.699 064	1489.456 379
2	1492.988 618	1492.071 674
3	1497.771 821	1495.883 917
4	1503.922 759	1500.639 184
5	1511.589 362	1506.428 694
6	1521.148 584	1513.557 404

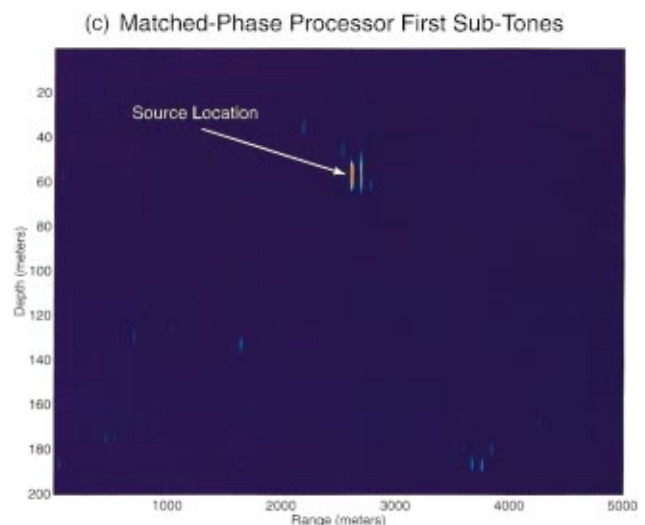
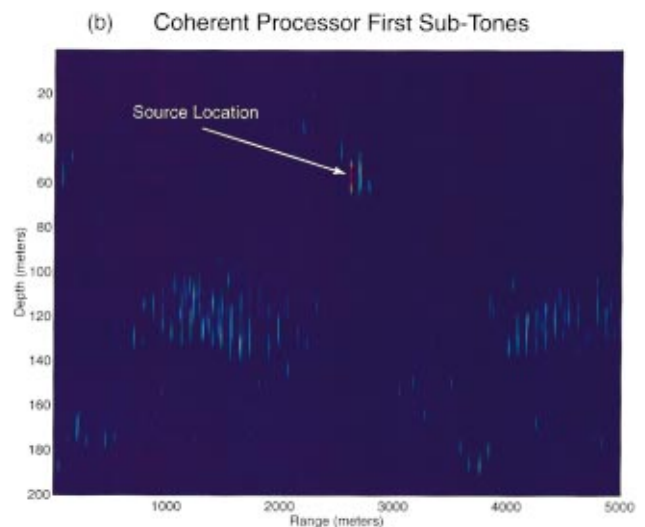
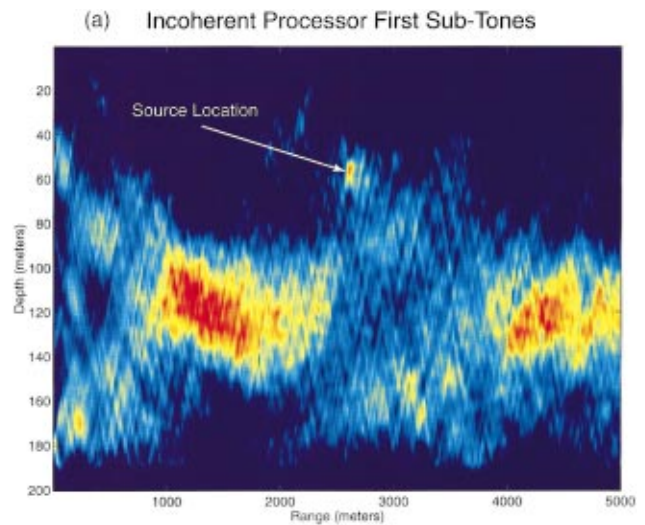


FIG. 7. Comparison of the three algorithms to the first set of nonpilot tones broadcast during event S19. All figures are plotted with 6 dB of dynamic range. Actual source location is at a range of 2600 m and depth of 55 m. Note the effect of noise creeping into the output of coherent algorithm in (b). The dynamic range in these figures has been set to 6 dB.

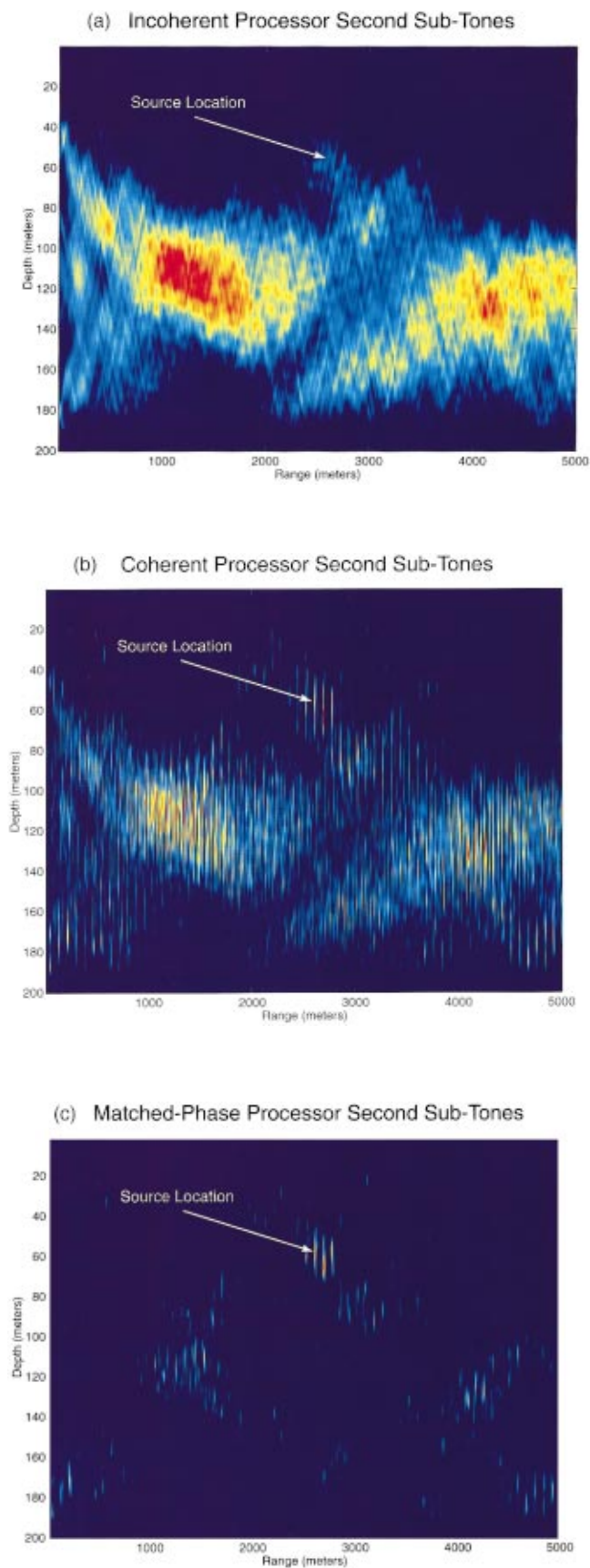


FIG. 8. At the lowest SNR analyzed in this paper, both the incoherent and coherent algorithms suffered from the high level of self-frequency correlation of the noise. Only the matched-phase processor was relatively unaffected. Its peak-to-sidelobe ratio being approximately 5 dB higher than the coherent algorithm of Eq. (9). The total dynamic range of the colormap is 6 dB.

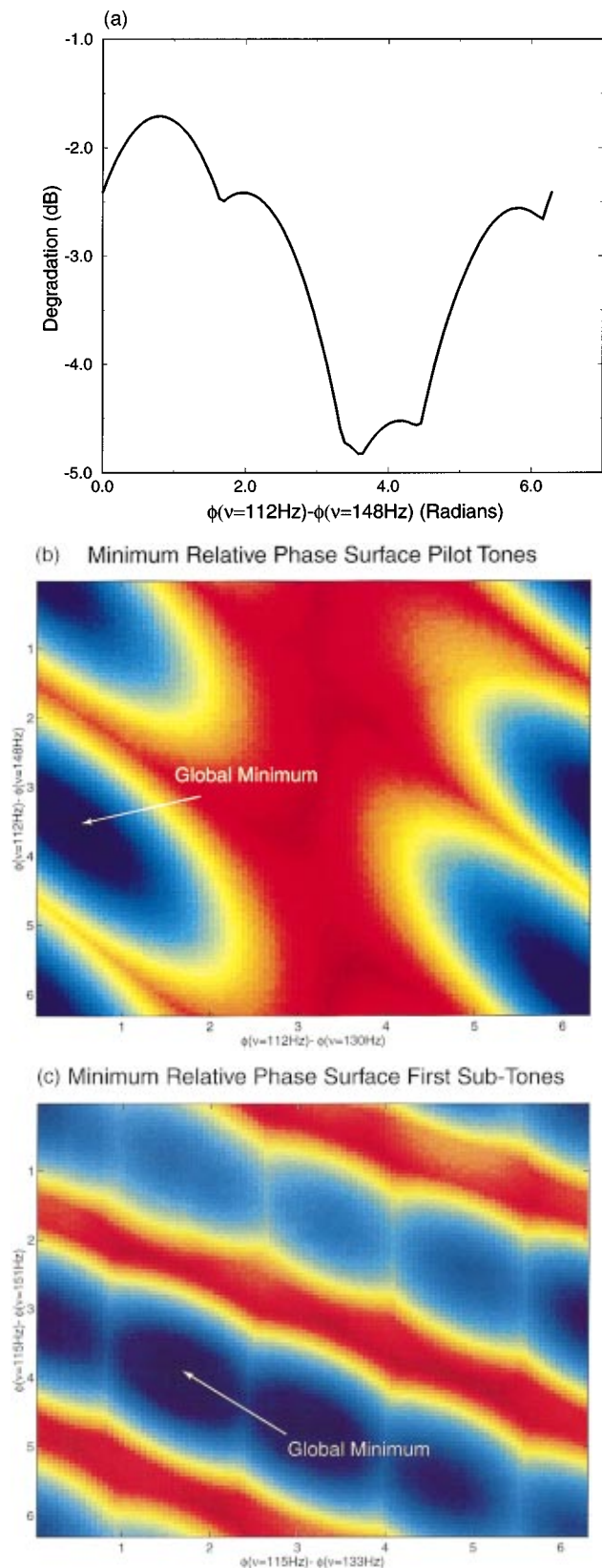


FIG. 9. This figure shows the maximum of the matched-phase coherent multi-frequency processor as a function of the relative phase relationship between processed frequencies. In (a) we show the processed frequencies 148 Hz and 112 Hz. Adding the additional frequency of 130 Hz, (b) shows the level surface of solutions to Eq. (12). Note that the minimum value corresponds to nontrivial relative phases. In (c) the strongest nonpilot tones (115 Hz, 133 Hz, and 151 Hz) were processed in the same fashion with an increase in the number of local extrema. In these figures the total dynamic range of the colormap is 6 dB.

B. MFP analysis

Since both source and receiver were along a roughly range-independent North–South iso-bathymetry contour line (see Fig. 1) at the beginning of event S19 (R/V SPROUL was at way point “=” at this time), range-independent normal mode methods were employed to calculate the replicas in the range-depth plane with an initial resolution of 10 m in the range direction and 1 m in the depth direction. The replicas were calculated using the principle of reciprocity: The 32 hydrophones were treated as sources and the acoustic field from each was propagated out to a distance of 5 km, each replica vector was then normalized for each point in the range-depth plane. Consistently, the propagation model predicted the transmission loss to the region of the source to be between 65 dB and 70 dB, using the input parameters from Table III and the sound speed profile from Fig. 4. These values are within reasonable limits given the source levels, the historical noise levels, and the estimated SNR on the array.

In Fig. 5(a) we show the resultant ambiguity functions where we have used the pilot tone frequencies (see Table II) in a Bartlett multi-frequency processor via Eq. (5), plotted with a total dynamic range of 6 dB, normalized to the peak value. The high source level relative to the background of these frequencies, evident from Table II, resulted in a relatively high SNR at each receiver and thus the performance of the Bartlett multi-frequency processor is exceptional. We note that the ability of the processor to resolve the range of the source is entirely a function of the highest frequency used in the algorithm. In applying Eqs. (9) and (11) to the same frequencies, Fig. 5(b) and (c) show the added resolution afforded the coherent algorithms. By coherently processing the five pilot tones the depth resolution and the range resolution are now nearly equal. (Range resolution in these images is 10 m per pixel, while depth resolution is 1 m per pixel.)

To demonstrate further the coherent algorithm's resolution enhancement, Fig. 6(a)–(b) show a blow-up of the region around the source location, of the three different terms present for the two frequency coherent algorithms, with frequencies $\nu_1 = 112$ Hz and $\nu_2 = 130$ Hz. The range resolution for these plots has been increased to 1 m per pixel and the dynamic of the colortable has been increased to 10 dB to highlight the sidelobes. Using the sediment parameters from Table III with the sound speed profile from Fig. 4, we used KRAKEN²³ to calculate the vertical eigenvalue spectrum. Table IV lists the six largest vertical eigenvalues at each frequency. Given the experimental setup and the fact that the third eigenmode has a node at approximately 55 m for both the 112-Hz and 130-Hz frequencies, using Eq. (10) we find the interference wavelength to be $\lambda_{1,2} = 82.45$ m. This value is very close to that estimated by directly measuring the peak to peak range in Fig. 6(d).

Figure 7(a) shows how the incoherent processor Eq. (5) suffers when the processed frequencies were switched to the first set of sub-tones at 115 Hz, 133 Hz, 151 Hz, 169 Hz, and 204 Hz with estimated SNRs of below -9 dB (see Table II). These figures have been adjusted to represent a total dynamic range of 6 dB. Not only have the sidelobes overtaken the true

location as the peak of the ambiguity function, but the main sidelobe has increased in size dramatically. Processing additional frequencies is unlikely to improve this behavior since all of the low source-level, single-frequency Bartlett ambiguity functions have sidelobes in the same general vicinity and often at a higher level than the peak of the true source location. In contrast, the matched-phase algorithm applied to the same lower level sub-tones shown in Fig. 7(c) demonstrate the ability of the matched-phase MFP processor to determine the source location more precisely than incoherent methods. Regions of large ambiguities corresponding to incorrect matches are broken up, even with the addition of only one or two frequencies, by the interference between modes of different frequencies. As more frequencies are added the effect becomes more pronounced. Also we show the results from applying Eq. (9) to these tones [Fig. 7(b)]. There the beginnings of the effects of noise are seen by the slight detection of sidelobes in the same locations as seen in Fig. 7(a).

In the final analysis of these data we have used the incoherent, coherent, and matched-phase algorithms on the second set of broadcast sub-tones (118 Hz, 136 Hz, 154 Hz, 169 Hz, and 207 Hz). The estimated SNR at the array for these tones ranges from -15 dB to -12 dB. At these low levels and with the ambient noise having a strong self-frequency correlation, as is evident from the previous incoherent results of Fig. 7(a), it is not surprising that Fig. 8(a) shows that the incoherent method loses the source entirely. Figure 8(b) shows that the normalization factor of Eq. (9) clearly overestimates the level of correlation between the data and the replicas of the like-frequency terms of Eq. (9c). This is seen as the increase in the size and intensity of the sidelobes as opposed to the matched-phase MFP processor using the same frequencies, shown in Fig. 8(c). By processing only the cross-frequency terms *a la* Eq. (11) we have eliminated much of the effects associated with the self-frequency correlation of the ambient noise field and have achieved approximately a five decibel peak-to-sidelobe ratio improvement in performance.

C. Global search for relative phases

As previously stated, for a small number of processed frequencies, it is possible to directly sample the range of relative phases densely enough to determine the extrema of the processor. We found that a discretized phase interval of $\pi/50$ proved to be sufficient to determine the relative phases for a small number of processed frequencies ($N_f \leq 3$). A two-frequency example is shown in Fig. 9(a). Here the maximum value in decibels of Eq. (6) is plotted as a function of the relative phase difference between frequencies $\nu_1 = 112$ Hz and $\nu_2 = 148$ Hz: Note that global maximum corresponds to a nonzero relative phase for these two frequencies. Including one more frequency ($\nu = 130$ Hz) a two-dimensional level surface was generated by exhaustively searching for the relative phases corresponding to the solution to Eq. (12). This result, shown in Fig. 9(b), demonstrates that the minimum value corresponds to the nontrivial relative phases of $\phi(\nu = 112 \text{ Hz}) - \phi(\nu = 130 \text{ Hz}) = 3.6$ radians and

$\phi(\nu=112\text{ Hz}) - \phi(\nu=148\text{ Hz}) = 0.5$ radians, with the third phase being constrained to $\phi(\nu=130\text{ Hz}) - \phi(\nu=148\text{ Hz}) = -3.1$ radians.

Continuing our analysis on to the tones with the next lowest source levels, in Fig. 9(c) we show the level surface produced by exhaustively searching for the solution to Eq. (12) for the emitted tones at frequencies of 115 Hz, 133 Hz, and 151 Hz. The minimum value corresponds to the phases $\phi(\nu=115\text{ Hz}) - \phi(\nu=133\text{ Hz}) = 3.99$ radians and $\phi(\nu=115\text{ Hz}) - \phi(\nu=151\text{ Hz}) = 1.72$ radians, with the third phase being constrained to $\phi(\nu=133\text{ Hz}) - \phi(\nu=151\text{ Hz}) = 2.27$ radians. The main new feature encountered in processing these lower SNR frequencies is the addition of several new local minima. It was generally found that as the SNR dropped across the array, the number of local minima increased.

As a benchmark the simulated annealing algorithm was employed on the two cases of three frequencies on which the exhaustive search technique was used [Fig. 9(b) and (c)]. In both cases, the extrema were determined to within 1% of their actual values. Further successful tests of the robustness of the algorithm were performed by starting the phases at different locations within the search space, and by adjusting the rate of cooling of the temperature. The algorithm was extended to handle five frequencies and subsequently employed to determine the optimal relative phases used to create the ambiguity functions described above. Convergence times generally ran several seconds to a few minutes, depending on the initial and final temperatures chosen.

III. DISCUSSION AND CONCLUSION

It should be emphasized that incoherent methods can work well in many situations involving sources with moderate to high signal-to-noise-ratio (SNR). Yet for MFP to really attain its full potential, it will be necessary to move to realistic three-dimensional dynamic (i.e., time-dependent) search spaces and to situations with SNRs much less than zero. As we have shown, if a source is weak and/or is only broadcasting a few frequencies (or only a few frequencies are to be processed), then the performance of the incoherent processors can be degraded significantly. Fortunately, incoherent processing techniques do not come close to using all of the available information contained in the data. Specifically the source spectrum levels and phase relationships between the various broadcast frequencies are assumed unrecoverable unknowns and either tacitly assumed to be trivial or normalized out somewhere in the incoherent processor.

We have explored several MFP processors in the literature that use information pertaining to the broadcast spectrum levels, which for brevity have been excluded from this paper. Our implementation of these processors on data collected during the SWellEx-96 experiment met with only mixed success. Specifically, when several well separated tones were broadcast with large linear differences in source levels, the output tended to be dominated by the frequency with the best degradation, which in this case directly corresponded to the highest broadcast source level at high SNR

and to the ambient noise field at low SNR. We have demonstrated in this paper that under some circumstances the information content of the relative phases can be exploited and might be a more promising means for increasing the resolution of MFP while simultaneously ameliorating some of the effects of noise and low SNR.

Our proposed processor's (Eq. (11)) chief difference from previous multi-frequency MFP processors is the method used to determine the relative phase information contained within the recorded data. Additionally, in an effort to improve this processor's performance in the presence of ambient noise, we solely use the cross-frequency terms in a coherent sum thereby disregarding the mono-tone terms. While the advantages of this new processor depend on the experimental situation, under conditions present in the SWellEx-96 experiment we have shown it to be superior in performance to the incoherent multi-frequency Bartlett MFP processor at reducing sidelobe intensities relative to the source peak. In general, we also found its performance superior to another coherent processor previously seen in the literature, showing as much as a five decibel enhancement in peak-to-sidelobe ratio at low SNR.

For a small number of relative phases a solution could be determined via inspection. This implementation assumed that the maximum peak of an ambiguity surface should correspond roughly to some source's position, albeit not necessarily to that of the desired source. This was a valid assumption for the tones used in the SWellEx-96 experiment with the highest source levels, where the array experienced a high SNR, but might be questionable for the frequencies broadcast at lower source levels, where the SNRs were significantly lower. One of the advantages of using global search algorithms is that more information about the system may be included in a natural way within the cost function to be minimized.

Finally, we wish to point out that the results presented represent only the beginning of the SWellEx-96 Event S19, an admittedly small amount of data. However, there is nothing to suggest that these data are unique with respect to the success of the algorithm. This work was intended only to show the efficacy of global minimization routines to ascertain the appropriate relative phases to be used within a coherent processor. Since the original writing of this manuscript this algorithm has been more extensively tested on more data from SWellEx-96, as well as other more recent experimental data sets. These newer results do not substantially differ from those presented here and it is in the interest of brevity that they have not been included. They are certain to be the topic of other forthcoming papers.

ACKNOWLEDGMENTS

The authors would like to thank William A. Kuperman of Scripps Institute of Oceanography's Marine Physics Laboratory at the University of California, San Diego for helpful comments and discussions. This work was supported by the Office of Naval Research.

- ¹H. 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).
- ²A. 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).
- ³A. Baggeroer, W. A. Kuperman, and H. Schmidt, "Matched-field processing: Source localization in correlated noise as an optimum parameter estimation problem," *J. Acoust. Soc. Am.* **83**, 571–587 (1988).
- ⁴S. M. Jesus, "Broadband matched-field processing of transient signals in shallow water," *J. Acoust. Soc. Am.* **93**, 1841–1850 (1993).
- ⁵N. 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).
- ⁶C. S. Clay, "Optimum time-domain signal transmission and source location in a waveguide," *J. Acoust. Soc. Am.* **81**, 660–664 (1987).
- ⁷L. N. Frazer and P. I. Pecholcs, "Single-hydrophone localization," *J. Acoust. Soc. Am.* **88**, 995–1002 (1990).
- ⁸R. K. Brienzo and W. S. Hodgkiss, "Broadband matched-field processing," *J. Acoust. Soc. Am.* **94**, 2821–2831 (1993).
- ⁹A. Tolstoy, "Computational aspects of matched-field processing in underwater acoustics," in *Computational Acoustics*, edited by D. Lee, A. Cakmak, and R. Vichnevetsky (North Holland, Amsterdam, 1990).
- ¹⁰E. K. Westwood, "Broadband matched-field source localization," *J. Acoust. Soc. Am.* **91**, 2777–2789 (1992).
- ¹¹D. P. Knobles and S. K. Mitchell, "Broadband localization by matched fields in range and bearing in shallow water," *J. Acoust. Soc. Am.* **96**, 1813–1820 (1994).
- ¹²T. Yang, "Broadband source localization and signature estimation," *J. Acoust. Soc. Am.* **93**, 1797–1806 (1993).
- ¹³Z. H. Michalopoulou and M. B. Porter, "Matched-field processing for broadband source localization," *IEEE J. Ocean Eng.* **21**, 384–392 (1996).
- ¹⁴Z.-H. Michalopoulou, "Robust multi-tone matched-field inversion: A coherent approach," *J. Acoust. Soc. Am.* **104**, 163–170 (1998).
- ¹⁵In Special Issue on Inversion Techniques and the Variability of Sound Propagation in Shallow Water, *IEEE J. Ocean Eng.* **21**, 324–485 (1996). There are several papers on this topic in this special issue.
- ¹⁶W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes in C, The Art of Scientific Computing*, 2nd ed. (Cambridge, New York, 1992).
- ¹⁷N. C. Makris, "A foundation for logarithmic measures of fluctuating intensity in pattern recognition," *Opt. Lett.* **20**, 2012–2014 (1995).
- ¹⁸A. Tolstoy, *Matched Field Processing* (World Scientific, New York, 1993).
- ¹⁹N. Metropolis, A. Rosenbluth, M. Rosenbluth, A. Teller, and E. Teller, "Equations of state calculations by fast computing machines," *J. Chem. Phys.* **21**, 1087–1092 (1953).
- ²⁰H. Szu and R. Hartley, "Fast simulated annealing," *Phys. Lett. A* **122**, 157–162 (1987).
- ²¹L. T. Fialkowski, J. S. Perkins, M. D. Collins, M. Michael, and J. A. Fawcett, "Matched-field source tracking by ambiguity surface averaging," *J. Acoust. Soc. Am.* (submitted).
- ²²R. T. Bachman, P. W. Schey, N. O. Booth, and F. J. Ryan, "Geoacoustic databases for matched-field processing: Preliminary results in shallow water of San Diego, California," *J. Acoust. Soc. Am.* **99**, 2077–2085 (1996).
- ²³M. B. Porter, "The KRAKEN Normal Mode Program," Tech. Rep. NRL/MR/5120-92-6920, Naval Research Laboratory (1992).

Use of higher order statistics in source signature estimation

Michael K. Broadhead, Lisa A. Pflug, and Robert L. Field

Naval Research Laboratory, Code 7173, Stennis Space Center, Mississippi 39529

(Received 15 October 1997; revised 30 July 1999; accepted 27 December 1999)

Higher order statistical blind deconvolution methods are implemented for use in removing multipath distortion from passively received underwater acoustic transient signals. Using single channel data and simulations, it is demonstrated that a fourth order method based on cumulant maximization can work well if the associated multipath Green's function is sufficiently "sparse." The iterative method is parameterized by filter length, and while there is a range of values at which the best solutions are obtained with conventional convergence criteria, useful solutions exist across a much broader range of filter lengths if the iterations are not always allowed to proceed to convergence. The fourth order objective functional is generalized to arbitrary order, and the method is shown to also produce good results for the third order objective functional. [S0001-4966(90)01604-6]

PACS numbers: 43.60.Cg [JCB]

INTRODUCTION

The passive sonar classification problem can be decomposed into two stages: (1) recovering the source time signature of a transient event from a set of received signals by accounting for environmental distortion effects; and (2) applying a pattern recognition algorithm to the estimated source signature, for final classification. By environmental distortion, we refer to effects present in the received data at the sensor array that are not present in the source signature, introduced by propagation through the ocean environment. In our case, environmental effects consist primarily of multipath effects. For a separable point source, if we incorporate the environmental effects into a Green's function and assume time-invariance, the received pressure time series at a desired location can be modeled as the convolution of the transient source signature with the Green's function (Burdic, 1991; Frisk, 1994). The Green's function, of course, depends upon the environmental acoustic parameters and the source and receiver locations.

If the Green's function is known, it can be used to deconvolve the measured or received time series to obtain an estimate of the source signature. This is referred to as the deterministic deconvolution approach. Broadhead *et al.* (1993) and Broadhead (1995) reviewed this approach [see, also, Mignerey and Finette (1992) and Finette *et al.* (1993)], and performed a study in a bottom-limited propagation environment, showing that there is significant sensitivity to inaccuracy in the calculated Green's function (at least in the single channel case). Error in the Green's function has many possible sources, the principal ones in shallow water probably being inaccuracy in the source location and in the seabed geoacoustic parameters.

Broadhead (1995) used a statistical source estimation approach to address the problem of recovering a source signature without specific knowledge of its location, or the environmental parameters necessary to accurately compute the Green's functions. The method used, called the Minimum Entropy Deconvolution (MED) method, was introduced by Wiggins (1978) and is an example of blind deconvolution, i.e., deconvolution using only the output signal of a convo-

lutional process (the filter is unknown). The goal of the MED method is to produce a filter that drives the output of the system to lower entropy (greater order). This was motivated by the "sparse" nature of the underlying Green's function in the original application (i.e., hydrocarbon induced amplitude anomalies in reflection seismograms).

In this paper the work begun in Broadhead (1995) is advanced in several ways. First, a broader interpretation of the Wiggins algorithm in terms of more recent higher order statistics research is presented. This broader interpretation leads to a generalized objective functional. Second, a more thorough examination of the solution space provided by the MED method is presented by: (a) introducing a new degree of freedom (iteration number); (b) including more data and simulation examples; and (c) presenting detailed performance results. Third, a system of equations is derived for the generalized objective function, and the third order case is tested. The conclusion reached is that such statistical methods show promise for application in transient sonar signal classification in the presence of multipath distortion.

I. CONVOLUTIONAL MODEL AND MULTIPATH

In an ocean environment, the propagation tends to be ducted, i.e., after a while, the propagation becomes confined to a finite vertical region. For the purpose of illustration, consider the homogeneous ocean represented in Fig. 1. Shown here are three paths that an omni-directional signature $s(t)$ could take: (1) direct path; (2) first bottom bounce; (3) first surface-bottom bounce. Treating the bottom as perfectly rigid, and with no volume attenuation effects, we could represent the arrival time and amplitude structure of a small amplitude (linear) pressure wave by the quantity labeled $g(t)$. The signal at the receiver, $r(t)$, is represented by summing up the various arrivals of $s(t)$ via the different paths, appropriately scaled and time-delayed according to the prescribed $g(t)$. The formal mathematical operation for representing the above construction is convolution, written

$$x(t) = s(t) * g(t) = \int_{-\infty}^{\infty} g(\tau) s(t - \tau) d\tau. \quad (1)$$

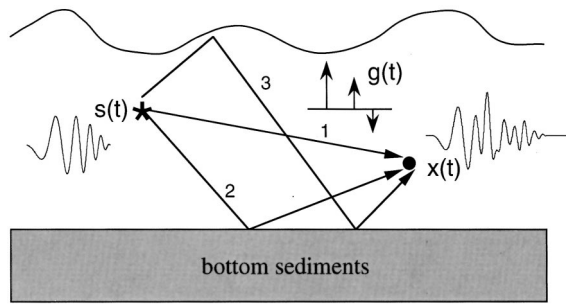


FIG. 1. Transient source with multipath distortion at the receiver due to a Green's function representing three transmission paths in the environment.

Here, $g(t)$ represents the Green's function, which is the point source solution to the acoustic wave equation. The Green's function represents the physical processes that distort $s(t)$ as it travels to the receiver. When multipath effects dominate, we assume we can represent the discrete-time Green's function by

$$g(t) = \sum_{i=0}^{N-1} g_i \delta(t - t_i), \quad (2)$$

where t_0, t_1, \dots, t_{N-1} represent the arrival times, and g_0, g_1, \dots, g_{N-1} represent pulse amplitudes. Combining Eqs. (1) and (2), we obtain

$$x(t) = \sum_{i=0}^{N-1} g_i s(t - t_i). \quad (3)$$

Hence, convolution is the appropriate mathematical operation to represent our intuitive description of the multipath effects above (scaling, shifting, and summing the multiple source pulse arrivals).

As a more concrete illustration, Fig. 2 depicts three Green's functions calculated from the numerical solution to the wave equation for a real ocean environment (Field and Leclerc, 1993). The location of the receiver in the experiment for which these Green's functions are modeled is the Atlantic Ocean at the southern end of the Blake Plateau

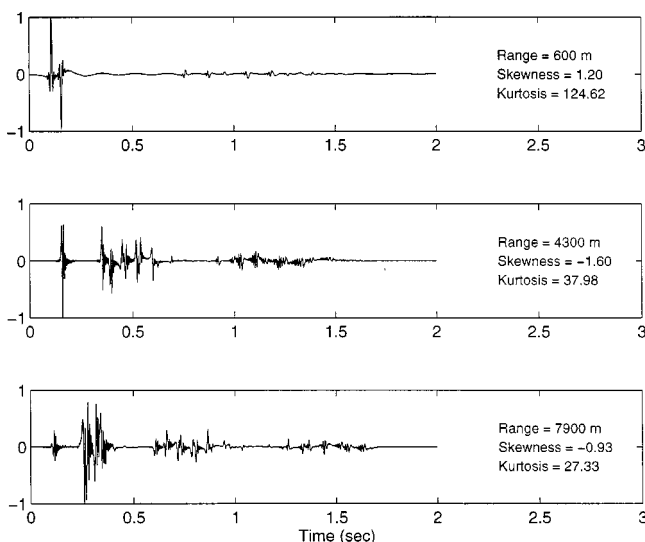


FIG. 2. Modeled Green's functions at three ranges for a depth of 250 m.

(27.5° N, 78.3° W) at a depth of 250 m. The bottom depth at the receiver location is 915 m and the three ranges at which the Green's functions are calculated are upslope from the receiver. The three Green's functions in Fig. 2 are used in the simulations presented later. Whereas in the physics of linear wave phenomena, the term Green's function is generally used, the corresponding linear systems theory term is impulse response function (IRF).

II. HIGHER ORDER STATISTICS OF MULTIPATH

In Wiggins' paper (1978), the Bright Spot seismic amplitude problem for a gas/water sand hydrocarbon reservoir was considered. This system can often be characterized by an IRF that is made up of a few, well-spaced, relatively large reflection coefficients. Wiggins wanted to avoid the stationary, white noise assumptions typically adopted in blind deconvolution algorithms at that time ("blind" because the source signature is rarely measured for land seismic exploration), since the IRF could be characterized as sparse, highly ordered, and from the point of view of information theory, of relatively low entropy [see Walden (1985) for a more in detailed discussion of this latter point of view].

To develop a statistical deconvolution method (blind, but making assumptions of a statistical nature, usually about the IRF) for Bright Spot analysis, Wiggins needed a measure of "sparseness," and chose the Varimax norm, which, for a single channel signal x of length N , is given by

$$V = \sum_{i=0}^{N-1} x_i^4 / \left(\sum_{i=0}^{N-1} x_i^2 \right)^2. \quad (4)$$

He then proceeded to consider the stationary points of the functional, V , in such a way as to converge to at least a local maximum. Eventually, the seismic deconvolution community recognized that the Varimax norm is fundamentally a normalized fourth order statistic (Walden, 1985; Tugnait, 1990; Cadzow, 1996). Consider, for example, the standard sample estimator for kurtosis

$$\hat{K} = \left[\frac{1}{N} \sum_{i=0}^{N-1} \left(\frac{x_i - \bar{x}}{\sigma_x} \right)^4 \right] - 3, \quad (5)$$

where \bar{x} is the mean of x , and σ_x is the standard deviation of x . Kurtosis is very closely related to the expression for V , as shown in detail in the following section.

To illustrate that kurtosis, and hence the Varimax norm, are measures of sparseness, consider Fig. 3, which shows how kurtosis decreases as the number of spikes in a Green's function increases. Beginning with an 1100-point function of zero amplitudes, increasing numbers of Gaussian-distributed spikes (nonzero amplitudes) are added to create simulated representatives for Green's functions (ordering is ignored because it is irrelevant for kurtosis). For each case, 100 realizations are used to calculate the average kurtosis. Kurtosis decreases rapidly as the number of spikes increases to about 100 out of 1100 points, and decreases steadily with additional spikes. Thus even though the Gaussian-distributed nonzero samples within the function have essentially zero

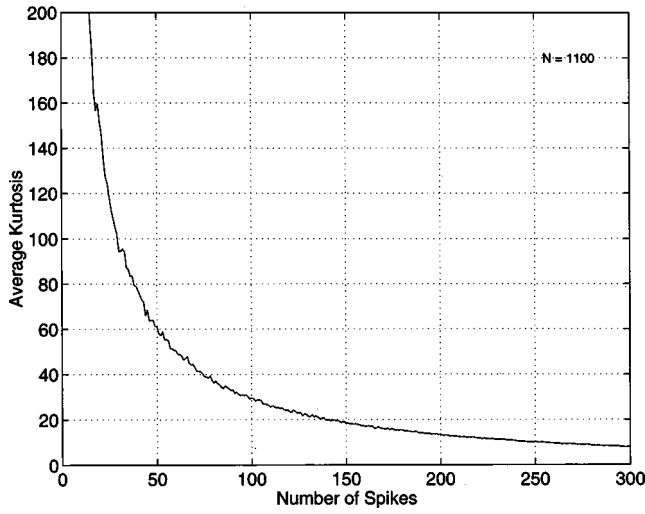


FIG. 3. Kurtosis versus number of nonzero spikes in a 1100-point sequence.

kurtosis, the additional zero values in the function make the distribution more narrow and peaked, hence, higher kurtosis. This feature of sparseness is typical of many underwater acoustic Green's functions. For example, the 600-m Green's functions in Fig. 2(a), with essentially two prominent spikes and many near-zero values, has a high kurtosis of 124.62. The 4300-m and 7900-m functions [Fig. 2(b),(c)] also have a significant number of near-zero values, although these larger ranges, with more severe multipath and spikes in the Green's functions, have lower (but significant) kurtosis values of 37.98 and 27.33.

Unfortunately, the MED method did not seem very applicable to reflection seismology data, probably due to the very low kurtosis values of typical seismic reflection sequences. However, there are signal types in other application areas where the Wiggins' approach, and its more modern successors, can be applied. As shown by Figs. 2 and 3, the functions that represent multipath distortion in underwater acoustics is one such area.

III. CUMULANT BASED FUNCTIONALS

To further explain the relationship between the Varimax norm and kurtosis, a discussion of cumulants and moments is warranted. Cumulants of a stationary process, x , are defined formally in the following way (Papoulis, 1984; Nikias and Raghuveer, 1987; Cadzow, 1996)

$$c_r^x = (-i)^r \frac{d^r \Psi_x(\omega)}{d(\omega)^r} \bigg|_{\omega=0}, \quad r=1,2,3,\dots, \quad (6)$$

where r is the cumulant order, $\Psi_x(\omega) = \ln[\phi_x(\omega)]$ is the cumulant generating function, and

$$\phi_x(\omega) = \int_{-\infty}^{\infty} e^{i\omega x} f(x) dx$$

is the moment generating function. It is possible to express the cumulants as functions of the moments of the random variable under analysis. The relationship between the first four cumulants and moments are (Nikias and Petropulu, 1993)

$$\begin{aligned} c_1 &= m_1, & c_2 &= m_2 - m_1^2, \\ c_3 &= m_3 - 3m_2m_1 + 2m_1^3, \\ c_4 &= m_4 - 4m_3m_1 - 3m_2^2 + 12m_2m_1^2 - 6m_1^4, \end{aligned} \quad (7)$$

where the mean of x is given by $m_1 = E\{x\}$. If the process is zero-mean, then the relationships between the first four cumulants and moments are much simpler

$$\begin{aligned} c_1 &= m_1, & c_2 &= m_2, \\ c_3 &= m_3, & c_4 &= m_4 - 3m_2^2. \end{aligned} \quad (8)$$

Since the pdf is often not available, estimates of the cumulants are generally used.

To avoid the property that cumulants depend on signal size, we consider the class of normalized cumulants (Cadzow, 1996)

$$c^x(p, q) = \frac{c_p^x}{|c_q^x|^{p/q}}, \quad (9)$$

where p and q are integers. The definition of normalized kurtosis ($p=4$, $q=2$) is generally referred to as simply "kurtosis"

$$K = c^x(4, 2) = \frac{c_4^x}{(c_2^x)^2} = \left[\frac{m_4^x}{(m_2^x)^2} \right] - 3. \quad (10)$$

For comparison, consider the sample kurtosis estimator given in Eq. (5) and assume that x is zero-mean (a natural assumption for Green's functions). If $\sigma_x = \sqrt{m_2}$, then

$$\begin{aligned} \hat{K} &= \left[\frac{1}{N} \sum_{i=0}^{N-1} \left(\frac{x_i}{\sqrt{m_2^x}} \right)^4 \right] - 3 \\ &= \left[\frac{1}{N} \sum_{i=0}^{N-1} (x_i)^4 / (m_2^x)^2 \right] - 3 = \left[\frac{m_4^x}{(m_2^x)^2} \right] - 3, \end{aligned} \quad (11)$$

which is the same as the expression for K in Eq. (10). Expanding for m_2 on the left-hand side of Eq. (11) gives

$$\hat{K} = \frac{1}{N} \left[\sum_{i=0}^{N-1} (x_i)^4 \right] / \left[\left(\sum_{i=0}^{N-1} (x_i)^2 \right)^2 \right] - 3, \quad (12)$$

which, except for the scale factor involving N and the subtraction of 3, is the Varimax norm given in Eq. (4), i.e.,

$$\hat{K} = \frac{1}{N} V - 3. \quad (13)$$

Hence we can expect that kurtosis well represents the signal properties exploited by the MED method. This also confirms Wiggins' intuition in choosing the Varimax norm for a sparseness measure. However, the more generalized framework of higher order statistics allows greater flexibility in the generation of alternative functionals [e.g., other (p, q) choices in Eq. (9)].

Although kurtosis is in some manner a measure of sparseness, the more common signal characteristic associated with kurtosis is that of non-Gaussianity. For a Gaussian process, all odd order moments are theoretically zero; and the even moments of orders four and higher are functions of the first and second order moments only. It can also be shown

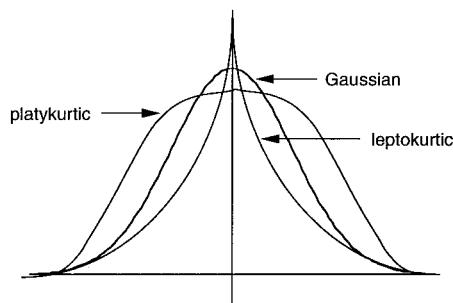


FIG. 4. Probability density functions for processes with negative (platykurtic), positive (leptokurtic), and zero (Gaussian) values of kurtosis.

that all odd and even order cumulants of orders above two are zero for a zero-mean Gaussian process. Hence, a nonzero kurtosis is an indication of departure from Gaussianity. A negative value for kurtosis indicates a platykurtic distribution for the process, and a positive value for kurtosis indicates a leptokurtic distribution, as shown by the sample pdf's in Fig. 4. From this figure, kurtosis can be seen to be a measure of the "peakedness" of a distribution. However, it is more accurately a measure of the tails of a distribution (Walden, 1985; Ali, 1974), with leptokurtic distributions having longer tails than platykurtic distributions. As the examples in Fig. 2 indicate, underwater acoustic Green's functions are often leptokurtic for tactical ranges.

To illustrate that convolution with a source signal does indeed produce a result of lower kurtosis than the original Green's function, a sample Green's function, with a kurtosis of 208, is convolved with six different source signal types. The results in Table I show that, for each source type, the kurtosis of the convolution with the Green's function, K_{s*g} , is significantly less than that of the original Green's function. Donoho (1981) proved that every filtered version of a white noise process is more Gaussian than the white noise itself. Empirically based observations indicate that the whiteness assumption can be relaxed, as can be noted in the numerical examples to be presented. We do point out, however, that our examples include source signatures that are relatively smooth and compact compared with the Green's function.

IV. WIGGINS' ALGORITHM

The MED algorithm has been thoroughly described in the literature, and will not be repeated here, but a minimum

TABLE I. Kurtosis of six signal types before and after convolution with a Green's function.

Source type, s	K_s	K_g	K_{s*g}
Gaussian modulated cosinusoid	8	208	18
Exponentially damped sinusoid	21	208	34
Broadband, symmetric, linear phase	818	208	160
Tonal	-4	208	3
Ping	9	208	5
Frequency modulated sweep	3	208	24

of terminology must be defined. The objective of the MED method is to produce a filter, f , of length l that is a stationary point of the functional

$$\hat{V} = \sum_{j=0}^{N-1} \hat{g}_j^4 / \left[\sum_{j=0}^{N-1} \hat{g}_j^2 \right]^2, \quad (14)$$

where

$$\hat{g}_j = \sum_{i=1}^L f_i x_{j-i}, \quad (15)$$

\hat{g} is the Green's function estimate, x is the input signal defined in Eq. (3), and V is the Varimax norm (which is essentially kurtosis). Maximization of V occurs by taking

$$\frac{\partial V}{\partial f_l} = 0, \quad (16)$$

which results in a nonlinear system of equations that is solved iteratively. A starting solution is obtained by taking f as a delta function. After iterating to some stopping criterion (to be discussed), V is assumed to have reached a maximum, and the filter f and the Green's function estimate \hat{g} have been found. The Green's function estimate can be written using Eq. (15), and a simple notation for the convolution process, as

$$\hat{g} = f * x = f * s * g. \quad (17)$$

If \hat{g} is an accurate estimate for g , then

$$\hat{g}^{-1} * g = \delta, \quad (18)$$

and

$$f * s = \delta. \quad (19)$$

Thus the source signature, \hat{s} , is finally obtained by calculating the inverse of f .

While the issue of sensitivity to noise is not addressed in these results, both Cadzow (1996) and White (1988) conclude that cumulant maximization methods withstand moderate amounts of noise. In contrast, Longbottom (1988) shows a seismic example in which the Wiggins' method accomplishes noise reduction in part by bandpass filtering the input data, but fails because the resulting passband is too narrow for the signal. The mechanism by which additive noise degrades the performance of the Wiggins algorithm has not yet been explained. While underwater transient classifiers are only expected to work at high signal-to-noise ratios and allow the majority of noise outside the signal passband to be filtered beforehand, the problem of noise must still be considered in this application, and is an issue of ongoing research.

V. SIGNAL DESCRIPTION

Two types of source signatures are used here for testing the performance of the MED algorithm. The first is a short-time pulse, estimated from a source array mounted hydrophone during the experiment in the vicinity of the Blake Plateau described earlier. The second is an exponentially damped sinusoid signal, called the long pulse, which was

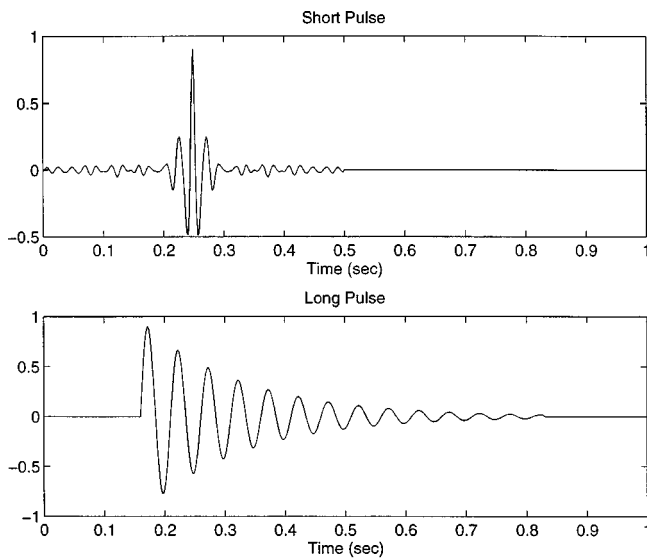


FIG. 5. Two source signatures: (a) the short pulse, and (b) the long pulse.

created to test the algorithm on a received signal with more overlap between the various arrivals. The two source signatures are shown in Fig. 5.

Three types of received signals are used as input to the MED algorithm: two are generated from the short pulse source, and one from the long pulse. The two source signals in Fig. 5 are convolved with the three modeled Green's functions given in Fig. 2 to generate simulated received signals for testing, and are referred to as SPSIMUL and LPSIMUL. Measured received signals for the propagated short pulse were recorded during the experiment and are also used for testing. This type of signal is referred to as SPMEAS. The MED input signals (received signals) at the propagation ranges of 600 m, 4300 m, and 7900 m are shown in Figs. 6, 7, and 8, for SPSIMUL, LPSIMUL, and SPMEAS. The SPMEAS input signals look slightly different from the SPSIMUL input signals, in part because the former contains additional noise contamination and, in the 7900-m case es-

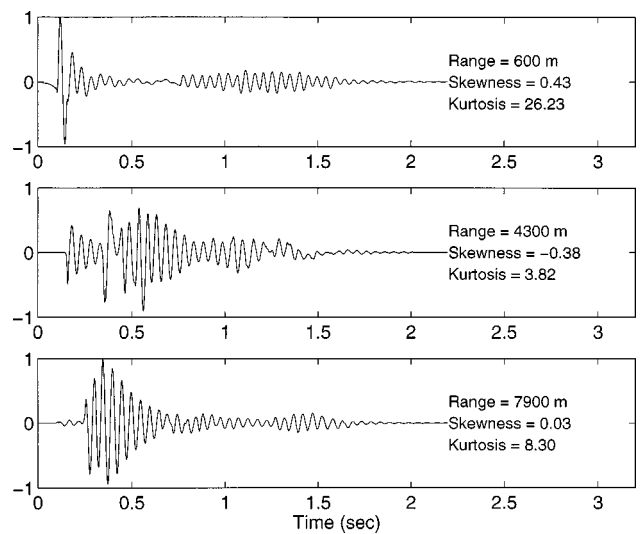


FIG. 7. Simulated input signals for the long pulse at the three propagation ranges.

pecially, because of inaccuracies in the modeled Green's function.

VI. PROCESSING METHODOLOGY

Two basic processing methodologies are examined. Method 1 uses a conventional convergence criterion for maximization of V , and the output is the absolute value of the correlation coefficient between the true source and the source estimate produced by the algorithm at filter lengths from 1 to 50. While the true source is not known in practice, it is used here to quantify the similarity between the true and estimated sources and evaluate performance of the algorithm. The convergence criterion is as follows: The correlation coefficient is calculated between the current MED filter iterate and the previous iterate. When this value exceeds a specified tolerance, the algorithm ceases. A tolerance level of

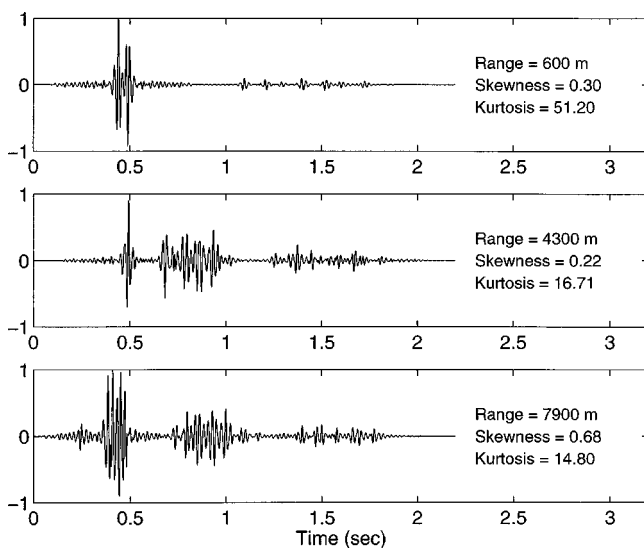


FIG. 6. Simulated input signals for the short pulse at the three propagation ranges.

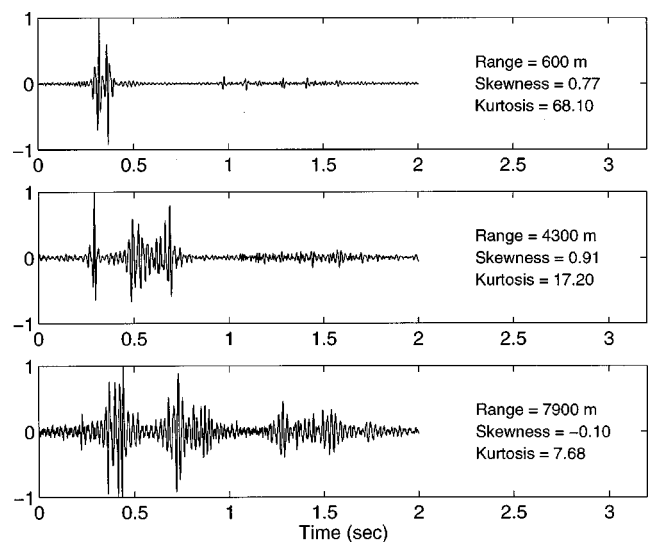


FIG. 8. Measured input signals for the short pulse at the three propagation ranges.

TABLE II. Highest correlation coefficients for Method 1.

Range	600 m	4300 m	7900 m
SPSIMUL	0.975	0.946	0.864
LPSIMUL	0.995	0.963	0.992
SPMEAS	0.896	0.802	0.822

0.9999 is used for the results presented in this paper. The results of this first methodology will be represented by solid curves in the following figures.

The output of method 2 is the absolute value of the correlation coefficient between the known source signature and the best possible source estimate achieved during the iterations to convergence, for each filter length. Whereas the output in the first methodology gives the source estimate associated with the maximization of V , the output in the second methodology gives the best source estimate generated out of all the iterations, without regard to actually trying to maximize V . In practice, this second method introduces iteration number as an additional free parameter in the solution space. The number of iterations required to reach the convergence criterion is generally between 30 and 40 iterations. Results from the second methodology will be represented by curves with long dashes in the following figures.

For comparison, the correlation coefficient of the known source signature with the MED input signal is also given. This value gives a measure of how much distortion is introduced by the multipath, and will be indicated by short dashes in the following figures.

There are two stages in each method where some regularization could be required. On a given iteration, the Toeplitz coefficient matrix may become ill-conditioned. Also, during the calculation of the inverse of the MED filter, there may be a problem with spectral zeros (a frequency domain method was used). For the SPSIMUL and SPMEAS cases, a pre-whitening value of 0.01% is used to avoid an ill-conditioned Toeplitz matrix. This value is also used for the 4300-m range of LPSIMUL. No pre-whitening is used for the other two ranges for LPSIMUL. In no case is pre-whitening used to calculate the inverse filter. A definite sensitivity to the amount of pre-whitening was noticed and would require further investigation for use in practice. A bandpass filter is used in obtaining the source estimate from the MED filter. This filter provides SNR enhancement in the cases with noise present (data), and additional stability in general.

VII. MED RESULTS

Tables II and III contain the correlation coefficients between the known sources and the source estimates generated by the MED algorithm for methods 1 and 2. For the different

TABLE III. Highest correlation coefficients for Method 2.

Range	600 m	4300 m	7900 m
SPSIMUL	0.975	0.958	0.893
LPSIMUL	0.995	0.965	0.993
SPMEAS	0.902	0.865	0.858

TABLE IV. Correlation coefficients with no preprocessing.

Range	600 m	4300 m	7900 m
SPSIMUL	0.682	0.539	0.589
LPSIMUL	0.546	0.647	0.848
SPMEAS	0.639	0.447	0.450

source types and processing methodologies, only the highest correlation coefficients are included in these tables, i.e., the filter lengths required to obtain these results are not taken into consideration. In most cases the highest coefficients obtained for the two methods are comparable. However, they are significantly different at other filter lengths, for some cases, as discussed later. In all cases, the best results were a significant improvement over doing no preprocessing at all, where the γ 's are, for all but one case, less than 0.7, as shown in Table IV. As expected, results are better for the short pulse simulations than for the real measurements. Not only did the data have some noise, but the "true" source is not known with complete accuracy.

In Fig. 9, we display results as a function of filter length for the LPSIMUL, at 600 m. This figure demonstrates that application of a conventional convergence criterion can work very well. This also happens to be a case where the associated Green's function kurtosis is very high (124.6). The simulated input signal (after convolution with the long pulse) has a kurtosis of 26.2. Figure 9 illustrates a case where allowing the iterations to proceed to convergence, and maximizing V , essentially gives the best source estimates for all filter lengths tested.

Figure 10 depicts the source signature estimates for the highest and lowest correlation coefficients for SPSIMUL with method 1. At the 600-m range, $\gamma=0.975$, and at the 7900-m range, $\gamma=0.864$ (compare the source estimates in Fig. 10 to the short pulse source shown in Fig. 5). These estimates are shown to provide an idea of the visual quality of match expected for the range of correlation coefficients relevant to this study.

Next, consider the contrasting MED results for

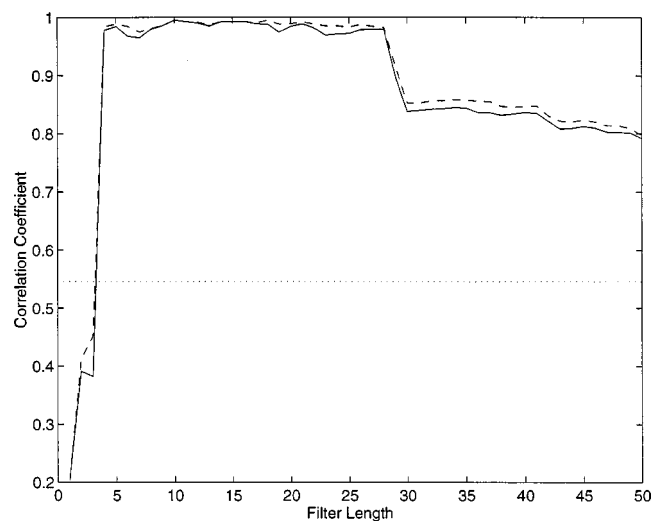


FIG. 9. Correlation coefficient versus filter length resulting from the two MED methods for the LPSIMUL, 600-m case.

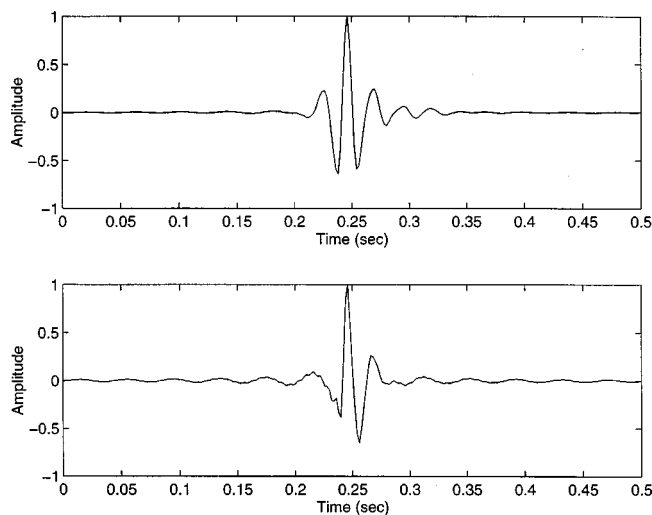


FIG. 10. Source estimates for SPSIMUL at (a) the 600-m range, and (b) the 7900-m range.

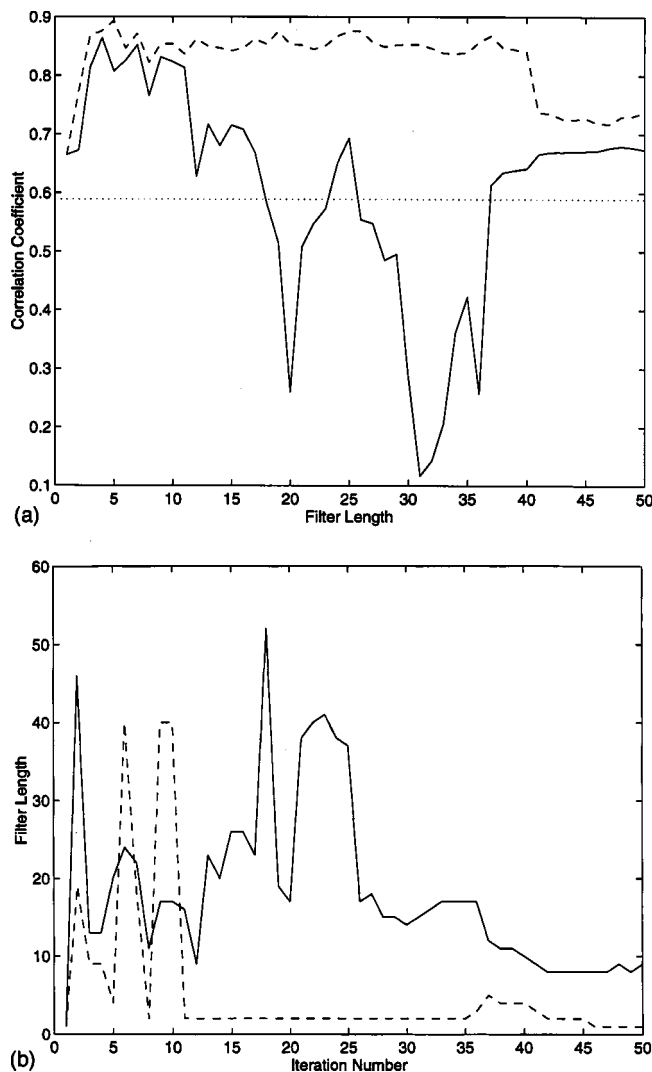


FIG. 11. (a) Correlation coefficient versus filter length and (b) filter length versus iteration number resulting from the two MED methods for the SPSIMUL, 7900-m case.

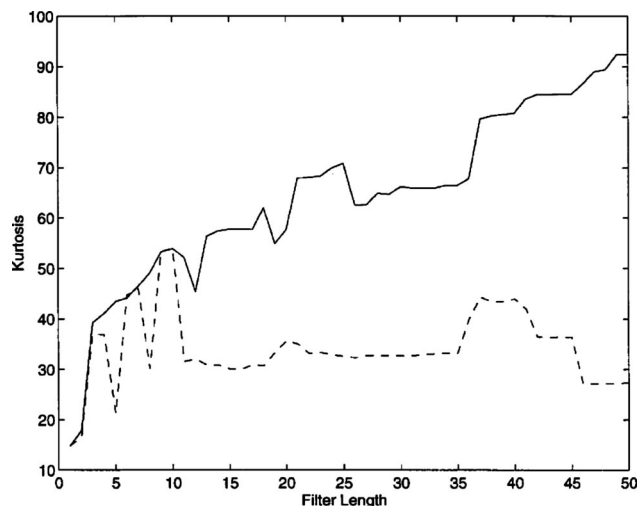


FIG. 12. Kurtosis of the estimated Green's functions over the range of filter lengths for SPSIMUL, 7900-m case.

SPSIMUL at the 7900-m range, shown in Fig. 11. Figure 11(a) shows that, for some filter lengths, allowing the V norm to be maximized, as in method 1, generates some source estimates that are worse than the input signal. However, the dashed curve for method 2 shows that good solutions are available at these filter lengths, only at many less iterations (often, only 2) than needed to maximize the V norm. This is confirmed by the results presented in Fig. 11(b), where the solid curve represents the number of iterations required to maximize the V norm (method 1), and the dashed curve represents the number of iterations associated with the best possible source estimate as the iterations proceed toward maximization of V (method 2). In contrast to the previous case (LPSIMUL), the initial Green's function kurtosis in the second case (SPSIMUL) is much lower (27.3), and the change in kurtosis after convolution with the short pulse is also less (14.8). Note that the starting kurtosis for this Green's function is comparable to the input (received) signal kurtosis in the example in Fig. 9. These factors are probably significant for determining when conventional convergence criteria will, or will not, work well.

In Fig. 12, we display the kurtosis of the MED Green's function estimates produced by both methods for the LPSIMUL, 7900-m case, as a function of filter length. This figure is fairly typical of the results. It shows a steady increase in the final (maximized) kurtosis value as the iterations proceed to convergence, but a level average value for the Green's function kurtosis, indicating that the most accurate estimate of the Green's function is not necessarily the estimate with the highest kurtosis.

VIII. GENERAL CUMULANT OBJECTIVE FUNCTIONAL

The Varimax norm is only one of a broader class of objective functions that can be used in the Wiggins' approach. The r th order objective function for this application, in the form of Eq. (14), is

$$\hat{V}^r = \sum_{j=0}^{N-1} \hat{g}_i^r / \left[\sum_{j=0}^{N-1} \hat{g}_j^2 \right]^{r/2}. \quad (20)$$

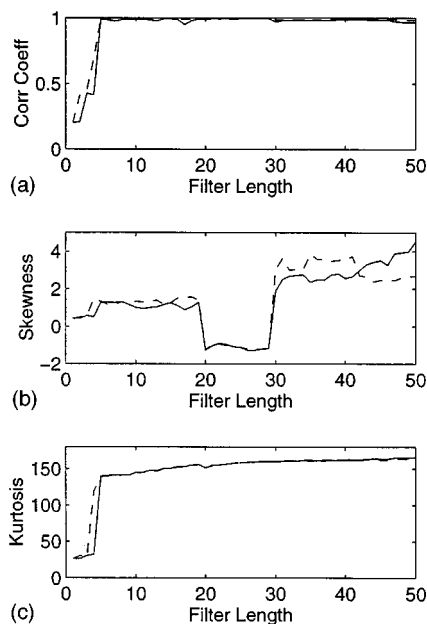


FIG. 13. (a) Correlation coefficients of source estimates, (b) skewness of the estimated Green's functions, and (c) kurtosis of the estimated Green's functions vs filter length for the LPSIMUL, 600-m case using V^3 .

For a zero-mean process, this is equal to Eq. (9) with $q=2$. Finding the stationary points of V^r results in a set of nonlinear equations which can be solved iteratively, as shown in the Appendix. Unlike the case for Wiggins' V norm (or, in this notation, V^4), the V^3 norm is not maximized by Wiggins' iterative method. However, good results can still be achieved with V^3 , as shown by the following two examples.

As is the case for kurtosis, the skewness of the multipath propagated signals are of lower magnitude than the skewness of the original Green's functions, and skewness maximization (or minimization) to achieve a source signal estimate is a natural idea to test. Results of the Wiggins' method with

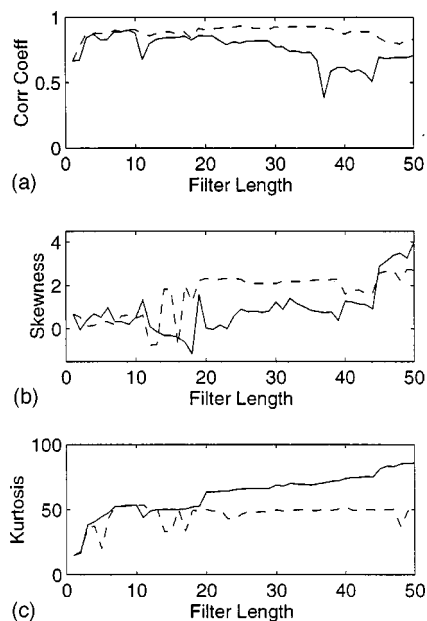


FIG. 14. (a) Correlation coefficients of source estimates, (b) skewness of the estimated Green's functions, and (c) kurtosis of the estimated Green's functions vs filter length for the SPSIMUL, 7900-m case using V^3 .

TABLE V. Highest correlation coefficients for Method 1 with the V^3 norm.

Range	600 m	4300 m	7900 m
SPSIMUL	0.972	0.948	0.911
LPSIMUL	0.997	0.965	0.992
SPMEAS	0.880	0.829	0.820

the V^3 norm (which is related to skewness), for the LPSIMUL case at 600-m range are shown in Fig. 13. The source estimates produced by the algorithm are quite good at filter lengths greater than 5, as shown in Fig. 13(a), and at filter lengths greater than 31, are better than those produced using the V^4 norm (compare to Fig. 9). The resulting Green's function skewness produced by the two methods can be either positive or negative at a given filter length, as indicated by the curves in Fig. 13(b), but is generally increased in magnitude over that of the input signal. The quality of the results may depend more on the fact that the resulting Green's function kurtosis appears to be increased by the algorithm [Fig. 13(b)], much as it is when using the V^4 norm.

The second example is for the SPSIMUL case at 7900 m shown in Fig. 11 for the V^4 norm maximization. The corresponding results for the V^3 norm are in Fig. 14. Although the source estimates from method 2 are slightly worse when using the V^3 norm, the estimates from method 1 are generally better for many filter lengths as shown in Fig. 14(a). As in the first example, the skewness changes in sign over filter lengths, and at some filter lengths, the two methods, in this example, produce values of skewness with opposite signs [Fig. 14(b)]. Again, the kurtosis is increased along with the skewness magnitude [Fig. 14(c)].

Comparing the full set of V^3 norm correlation coefficients for methods 1 and 2 (Tables V and VI) to the correlation coefficients for the V^4 norm (Tables II and III) indicate that, overall, performance is similar. The most notable difference is that use of the V^3 norm produces a 0.047 increase in the correlation coefficient at convergence (method 1) for SPSIMUL at the 7900-m range. However, if the redundancy of good solutions across filter length is taken into consideration, there can be significant differences (e.g., Figs. 11 and 14) and quantifiable comparisons would depend somewhat on how the algorithm results are used in the classifier.

Although use of the V^3 norm in the Wiggins' algorithm produces comparable, and sometimes superior results to those produced by the V^4 norm, the underlying mechanism and sensitivities to input signals and parameters are not yet fully understood. Modifications to the Wiggins' iterative algorithm to address the issue of V^3 maximization/minimization are currently being tested. These sample results, however, show that further investigation is warranted.

TABLE VI. Highest correlation coefficients for Method 2 with the V^3 norm.

Range	600 m	4300 m	7900 m
SPSIMUL	0.975	0.950	0.920
LPSIMUL	0.997	0.965	0.992
SPMEAS	0.909	0.846	0.888

IX. DISCUSSION AND CONCLUSIONS

For the cases studied, the MED method as a preprocessor for classification is always capable of providing better results than doing no preprocessing at all. In most cases, the best results were very good ($\gamma > 0.9$). Even when the processor gave good results only over a narrow range of filter lengths, many good solutions were still available at other filter lengths, only at a smaller number of iterations than maximizing the V norm would require. Since the goal is to produce a class of candidate solution signals for classification, this is useful in that it may be exploitable in increasing the probability of having a “good” solution in the class of signal candidates, albeit at the expense of adding a dimension to the search space.

The degree of kurtosis possessed by the Green’s function representing the multipath distortion appears to be an important factor in the quality of the results. The final kurtosis value of the convolved signal may also be important, and remains to be determined.

In general, maximization of the Green’s function kurtosis is more likely to give good source estimates when the Green’s function kurtosis is high. When it is low, the maximization technique is more likely to achieve the best source estimate before the method converges. Using a third order functional in the Wiggins’ approach also produces good results, although the underlying mechanism and sensitivities driving the algorithm are more difficult to ascertain.

ACKNOWLEDGMENTS

This work was funded by the Office of Naval Research and the Naval Research Laboratory, NRL contribution number NRL/JA/7173-96-0009.

APPENDIX. EXTENSION OF THE WIGGINS APPROACH TO THE r th ORDER OBJECTIVE FUNCTION

Let the r th order normalized cumulant objective, V^r , be represented by

$$\hat{V}^r = \sum_{j=0}^{N-1} y_j^r / \left[\sum_{j=0}^{N-1} y_j^2 \right]^{r/2}, \quad (\text{A1})$$

where

$$y_j = \sum_i f_i x_{j-i}. \quad (\text{A2})$$

y is assumed to have zero mean and x is the input function. While V represents the Varimax norm in Wiggins’ approach, the notation V^r is here used to represent the arbitrary order functional. To find the stationary points of V^r with respect to f_l , the filter of length l , find

$$\frac{\partial V^r}{\partial f_l} = 0. \quad (\text{A3})$$

Expanding Eq. (A3) gives

$$\frac{\partial V^r}{\partial f_l} = \frac{[\sum_j y_j^r]' [\sum_j y_j^2]^{r/2} - [\sum_j y_j^r] [(\sum_j y_j^2)^{r/2}]'}{[\sum_j y_j^2]^r}. \quad (\text{A4})$$

Evaluating the two derivatives in Eq. (A4) yields

$$\begin{aligned} \left(\sum_j y_j^r \right)' &= r \sum_j y_j^{r-1} \frac{\partial y_j}{\partial f_l} \\ &= r \sum_j y_j^{r-1} \frac{\partial}{\partial f_l} \left(\sum_l f_l x_{j-l} \right) \\ &= r \sum_j y_j^{r-1} x_{j-l}, \end{aligned} \quad (\text{A5})$$

and

$$\begin{aligned} \left[\left(\sum_j y_j^2 \right)^{r/2} \right]' &= \frac{r}{2} \left(\sum_j y_j^2 \right)^{r/2-1} \left(\sum_j y_j^2 \right)' \\ &= \frac{r}{2} \left(\sum_j y_j^2 \right)^{r/2-1} \left(2 \sum_j y_j x_{j-l} \right) \\ &= r \left(\sum_j y_j^2 \right)^{r/2-1} \sum_j y_j x_{j-l}. \end{aligned} \quad (\text{A6})$$

Substituting the derivatives from Eqs. (A5) and (A6) into Eq. (A4) gives

$$\frac{\partial V^r}{\partial f_l} = \frac{[r \sum_j y_j^{r-1} x_{j-l}] [\sum_j y_j^2]^{r/2} - [\sum_j y_j^r] [r (\sum_j y_j^2)^{r/2-1} \sum_j y_j x_{j-l}]}{[\sum_j y_j^2]^r}. \quad (\text{A7})$$

For simplicity, let $u = \sum_j y_j^2$, and set Eq. (A7) equal to zero [Eq. (A3)]

$$r u^{-r/2} \sum_j y_j^{r-1} x_{j-l} - r u^{-(r/2+1)} \sum_j y_j^r \sum_j y_j x_{j-l} = 0. \quad (\text{A8})$$

With this substitution, $V^r = \sum_j y_j^r / u^{r/2}$, and Eq. (A8) can be rewritten as

$$u^{-r/2} \sum_j y_j^{r-1} x_{j-l} - u^{-1} V^r \sum_j y_j x_{j-l} = 0. \quad (\text{A9})$$

Now, using Eq. (A2),

$$u^{-1} V^r \sum_j \sum_k f_k x_{k-l} x_{j-l} = u^{-r/2} \sum_j y_j^{r-1} x_{j-l}. \quad (\text{A10})$$

For $r=4$, this is Eq. (4) in Wiggins’ paper. The system of equations given by Eq. (A10) can be solved iteratively for the r th order objective function.

- Ali, M. M. (1974). “Stochastic ordering and kurtosis measure,” *J. Am. Stat. Assoc.* **69**, 543–545.
- Broadhead, M. K. (1995). “Broadband source signature extraction from underwater acoustics data with sparse environmental information,” *J. Acoust. Soc. Am.* **97**, 1322–1325.
- Broadhead, M. K., Field, R. L., and Leclerc, J. H. (1993). “Sensitivity of the deconvolution of acoustic transients to Green’s function mismatch,” *J. Acoust. Soc. Am.* **94**, 994–1002.
- Burdic, William S. (1991). *Underwater Acoustic System Analysis* (Prentice-Hall, New Jersey).
- Cadzow, J. A. (1996). “Blind deconvolution via cumulant extrema,” *IEEE Signal Process. Mag.* **13**, 24–42.
- Donoho, D. (1981). “On minimum entropy deconvolution,” in *Applied Time Series Analysis II*, edited by D. Findley (Academic, New York), pp. 565–608.

- Field, R. L., and Leclere, J. H. (1993). "Measurements of bottom-limited ocean impulse responses and comparisons with the time domain parabolic equation," *J. Acoust. Soc. Am.* **93**, 2599–2616.
- Finette, S., Mignerey, P. C., Smith III, J. F., and Richmond, C. D. (1993). "Broadband source signature extraction using a vertical array," *J. Acoust. Soc. Am.* **94**, 309–318.
- Frisk, George V. (1994). *Ocean and Seabed Acoustics* (Prentice-Hall, New Jersey).
- Longbottom, J., Walden, A. T., and White, R. E. (1988). "Principles and application of maximum kurtosis phase estimation," *Geophysical Prospecting* **36**, 115–138.
- Mignerey, P. C., and Finette, S. (1992). "Multichannel deconvolution of an acoustic transient in an oceanic waveguide," *J. Acoust. Soc. Am.* **92**, 351–364.
- Nikias, C. L., and Petropulu, A. P. (1993). *Higher-Order Spectra Analysis* (Prentice-Hall, New Jersey).
- Nikias, C. L., and Raghuveer, M. R. (1987). "Bispectrum estimation: A digital signal processing framework," *Proc. IEEE* **75**, 869–891.
- Papoulis, A. (1984). *Probability, Random Variables, and Stochastic Processes* (McGraw-Hill, New York).
- Tugnait, J. K. (1990). "Approaches to FIR system identification with noisy data using higher order statistics," *IEEE Trans. Acoust., Speech, Signal Process.* **38**, 1307–1317.
- Walden, A. T. (1985). "Non-Gaussian reflectivity, entropy, and deconvolution," *Geophysics* **50**, 2862–2888.
- White, R. E. (1988). "Maximum kurtosis phase correction," *Geophys. J.* **95**, 371–389.
- Wiggins, R. A. (1978). "Minimum entropy deconvolution," *Geoprospection* **16**, 21–35.

Frequency responses of two- and three-tone distortion product otoacoustic emissions in Mongolian gerbils

David M. Mills^{a)}

Virginia Merrill Bloedel Hearing Research Center, Department of Otolaryngology, Head and Neck Surgery, University of Washington, Seattle, Washington 98195

(Received 24 September 1999; accepted for publication 5 January 2000)

The frequency responses of distortion product otoacoustic emission (DPOAEs) were investigated in adult Mongolian gerbils. The main goal was to investigate in this species the extent to which DPOAE measurements might be useful in estimating cochlear frequency-tuning characteristics. Specifically, this study investigated the parameter space for generation of DPOAEs to determine those regions, if any, where the emission responses gave “simple” frequency responses, i.e., responses similar in form to typical neural responses. At the same time, it was desired to determine in this species the existence, extent, and nature of the more complex three-tone emission frequency responses as observed in some other species [e.g., Martin *et al.*, Hearing Res. **136**, 105–123 (1999)]. In the present work, two-tone frequency response curves (f_2/f_1 ratio functions) were obtained by varying the lower frequency, f_1 , while holding the f_2 frequency and both amplitudes (L_1, L_2) constant. Only for frequencies, f_2 , near 8 kHz did the response at the emission frequency, $2f_1 - f_2$, form a simple, relatively broad peak. At all lower frequencies, the two-tone frequency response curve was typically complex and composed of multiple peaks. In comparison, three-tone frequency responses were constructed by fixing the primary stimulus pair (f_1, f_2) and varying a third tone widely in frequency (f_3) and intensity (L_3). Points in f_3 and L_3 which caused a criterion reduction in primary emission amplitude (at $2f_1 - f_2$) were used to construct emission suppression tuning curves (STCs). Only for primary frequencies, f_2 , at 8 kHz and above were the emission STCs found to be simple, with shapes similar to neural frequency-tuning curves. At lower primary frequencies, particularly for relatively low primary frequency ratios (low f_2/f_1), three-tone responses were very complex. This complex response usually included a region of anomalous suppression in which very low suppression levels (L_3) could result in significant decreases in the primary emission amplitude, often exceeding 12 dB. Regions of such anomalous suppression were typically observed under the following conditions: (1) for all f_2 frequencies from 0.5 to 4 kHz; (2) for f_3 frequencies between 1.4 and 8 kHz; (3) i.e., for f_3 frequencies 1–3 octaves above the primary frequency, f_2 ; (4) at L_3 levels often 10 dB lower or more than the usual “best frequency” threshold, i.e., even lower than the relative minimum threshold found near the primary stimulus frequencies; (5) exhibiting sharp amplitude decreases often accompanied by emission phase shifts of about 180 deg; (6) present in both cubic emissions ($2f_1 - f_2$ and $2f_2 - f_1$); (7) to be less extreme at larger primary stimulus frequency ratios (larger f_2/f_1); and (8) less extreme at larger intensity ratios (larger L_1/L_2). Because of the anomalous behavior at f_2 frequencies below 8 kHz, “simple” emission STCs were typically only obtainable, if at all, near the extreme boundaries of the parameter space giving measurable emission amplitudes. © 2000 Acoustical Society of America. [S0001-4966(00)04104-7]

PACS numbers: 43.64.Jb, 43.64.Kc [BLM]

INTRODUCTION

One of the fundamental tasks of auditory system evaluation is measurement of the frequency response at each stage of the system. In the auditory periphery, responses of individual eighth-nerve fibers provide one fundamental measure of cochlear response (e.g., Dallos *et al.*, 1978; Liberman and Kiang, 1978; Schmiedt and Zwislocki, 1978; Schmiedt, 1989). The stimulus level required for a threshold neural response as a function of stimulus frequency defines a frequency-tuning curve (FTC). A typical neural FTC has a sharp, sensitive tip which defines its “best frequency.”

There is a rapid rise in threshold above the best frequency, and a slower rise at lower frequencies which often forms an extended low-frequency “tail.” The width of the FTC at a criterion level above its best frequency response has been traditionally used as a measure of frequency selectivity.

Direct neural measurements are inherently invasive, of course, and secondary methods have been sought which are less invasive but which might give equivalent information (e.g., Dallos and Cheatham, 1976; Eggermont, 1976; Dallos *et al.*, 1978; Dolan *et al.*, 1985). Since their discovery, otoacoustic emissions have held promise to fulfill such a role (Kemp, 1978).

For two-tone distortion product otoacoustic emissions (DPOAE), it has been suggested that the response of the

^{a)}Electronic mail: dmmills@u.washington.edu

emission amplitude as a function of the relative frequencies of the two primaries (i.e., the f_2/f_1 ratio function) might provide a useful measurement of the frequency selectivity of the cochlea (Gaskill and Brown, 1990). The reasoning is straightforward. Two-tone distortion product emissions are presumed to be generated in the region in the cochlea where there is a maximal "overlap" between the traveling waves associated with the two stimulus frequencies (Martin *et al.*, 1987). If the two stimulus amplitudes are held at constant, low levels, and the difference in frequency between them is increased, the emission amplitude should decrease as the overlap between the two traveling wave patterns decreases. The width of the resulting "frequency response curve" (i.e., the f_2/f_1 ratio function) should be a measure of the relative width of the peak of a typical traveling wave, and hence of the cochlear frequency selectivity (e.g., Kössl and Vater, 1996). In many cases, measurements have been obtained which are consistent with this interpretation (Brown and Gaskill, 1990a; Brown *et al.*, 1992, 1993a; Kössl, 1993).

However, this simple interpretation is complicated by the existence of multiple sharp peaks in the frequency response curve. In mammals, multiple peaks have been particularly noted in mole rats (Kössl *et al.*, 1996) and in gerbils (Mills and Rubel, 1997). Similar, but usually less extreme, multiple peaks can also be seen in published human responses (Gaskill and Brown, 1990; Brown *et al.*, 1993a). A detailed comparison of results from different species is included in the Discussion. The cause of the multiple peaks remains unknown, and consequently also unknown is the extent to which the mechanism(s) responsible might be an underlying problem. In other words, the same unknown mechanism may be responsible for the actual overall shape of frequency curves observed in similar measurements, or this mechanism might be significantly distorting the shape, even when multiple peaks are not obvious. This implies that, even when emission frequency responses are obtained which look like reasonable indicators of cochlear tuning, the observed behavior could be a result of the underlying unknown mechanism(s), and any agreement fortuitous.

An alternative to the two-tone DPOAE frequency response is an emission suppression paradigm, in which three tones are employed (Brown and Kemp, 1984; Martin *et al.*, 1987; Harris *et al.*, 1992; Köppl and Manley, 1993). In the typical measurement paradigm, the frequencies and intensities of two primary tones (at f_1 and f_2) are first chosen to give a measurable emission amplitude at the frequency $2f_1 - f_2$, i.e., at the frequency of the cubic distortion tone (CDT) emission. The primary tones are fixed and the frequency and intensity of a third tone are varied widely. A "suppression tuning curve" (STC) is constructed by connecting the points (frequency f_3 , intensity L_3) which result in a criterion decrease in the CDT amplitude. Initial results for this approach were promising (e.g., Harris *et al.*, 1992; Cianfrone *et al.*, 1994; Frank and Kössl, 1995; Kummer *et al.*, 1995; Abdala *et al.*, 1996; Abdala, 1998; Mills, 1998). That is, the emission STCs so obtained were generally found to have characteristics similar to FTCs. There was a most sensitive tip, usually found near the frequencies of the two primaries, at least for an appropriate choice of primary

stimulus parameters (Mills, 1998). Above this best frequency, the STC rose rapidly as frequency (f_3) was increased. Below the tip, the contour for the criterion response increased more slowly, usually showing an extended low frequency "tail." Such STC responses are broadly similar to published neural-tuning curves (e.g., Ohlemiller and Echteler, 1990).

Finding simple emission STC responses which are similar in form to neural FTCs, it is tempting to interpret them similarly. For example, measurements of the width of the STC at a criterion level above its best frequency "threshold" have been interpreted as an estimate of frequency selectivity, e.g., in human neonates (Abdala and Sininger, 1996). Caution is indicated in such interpretations, however. There is no necessary straightforward connection between *any* kind of suppression tuning curve and a neural frequency-tuning curve (Mills, 1998). This fact may be one reason that the tip-to-tail ratios found with emission STCs are typically much smaller than those found for FTCs of individual neurons (e.g., for gerbil, compare Mills, 1998, with Ohlemiller and Echteler, 1990, and Hellstrom and Schmiedt, 1996).

Recently, additional complications have arisen for the emission STC paradigm. In measurements in rabbits, loop diuretics were reported to change emission STCs differently than neural FTCs would be expected to change under similar circumstances (Martin *et al.*, 1998). In other STC measurements in the same species, the third tone caused changes in the primary emission amplitude at frequencies well away from those of the two primaries and for quite low stimulus levels (Martin *et al.*, 1999). The stimulus levels for the third tone required for criterion suppression were sometimes found *below* levels required for suppressor frequencies in the region of the primaries. Both increases and decreases in the emission amplitude were found, that is, "enhancement" as well as "suppression." The effects were extensive enough to cause considerable doubt as to the mechanisms of emission generation and to the proper interpretation of such STC measurements (Martin *et al.*, 1998, 1999).

The presence of complex behavior of unknown physical origin in both two- and three-tone responses complicates the attempt to measure cochlear frequency tuning using otoacoustic emission paradigms. This complex behavior appears to be present to some degree in all species. As will be shown, in the gerbil these complex responses in their extreme form extend over more than half of the animal's usable frequency range. The strength and persistence of the complex behavior in the gerbil makes this an excellent species in which to investigate this behavior.

There are other reasons to attempt to understand the complex behavior in the Mongolian gerbil in particular. The auditory function of this animal has been well studied, for good reasons. The animal is useful for developmental studies and has, for a rodent, excellent low-frequency hearing (Ryan, 1976). Given the relatively immature neural responses in the young animals, otoacoustic emissions have been essential to recent progress in studying the development of cochlear function in this animal (Norton *et al.*, 1991; Mills and Rubel, 1996). Further progress, however, depends on understanding the cause and extent of the complex frequency behavior. The

development of frequency resolution might still be studied by three-tone emissions if a range of primary stimulus parameters was found for which complex suppression behavior was absent or at least less extreme.

Even so, unless the causes and effects of such complex behavior are understood in detail, the correct interpretation of emission measurements in other situations is necessarily less certain. That is to say, the physical processes responsible for these complex responses are presently unknown. As noted above, these unknown physical causes could actually underlie the observed behavior even when such physical causes produce only one peak, not two, in the two-tone frequency-response curve. Similar arguments apply to the interpretation of three-tone measurements. The present report is not intended to provide a final resolution of these concerns; rather, it is intended to provide detailed results of observations in Mongolian gerbil as a basis for further investigation of these issues.

This report details two- and three-tone responses obtained in adult Mongolian gerbils as a function of the primary stimulus parameters. The primary stimulus frequencies were independently varied across the available range, for a relatively restricted set of primary stimulus levels. The stimulus levels were generally set as low as possible while still obtaining adequate emission amplitudes, with the goal of focusing the investigation on the vulnerable, sensitive aspects of cochlear function. In all cases, the animals investigated were healthy adult gerbils with normal emissions.

I. METHODS

Equipment and procedures were generally the same as for previous studies of two- and three-tone distortion product emissions (Mills, 1997, 1998). In the present study, 30 young adult gerbils of 2–4 months of age were the subjects. After anesthesia, the outer portion of the ear canal was removed on one side and tissue removed over the dorsolateral bulla. In most animals, the bulla was then opened widely by drilling a 1.5-mm-diameter hole in it. This was done to provide both pressure relief and improved middle-ear transmission at low frequencies, i.e., after the closed sound system was sealed to the ear canal. In some cases, when a measurement using the animal's normal middle-ear characteristics was desired, the bulla was instead *vented*. A smaller hole was drilled, and a 1.0-mm i.d. by 30-cm-long tube was force-fit into the hole. The function of the vent tube was to provide static pressure relief without significantly changing the auditory characteristics of the middle ear.

At the start of emission measurements, the emission amplitude resulting from a two-tone stimulus (frequencies f_1 and f_2) with $f_2 = 8$ kHz, $f_2/f_1 = 1.25$, and stimulus levels $L_1 \times L_2 = 50 \times 30$ dB SPL was measured, and monitored at about 1-h intervals throughout the experiment. This emission amplitude was typically in the range of 10–20 dB SPL, and if it was found below this range at any time the experiment was terminated and data discarded back to the last passing measurement.

For most animals, emission studies began with a two-tone frequency response function constructed as follows. The f_2 frequency to be investigated was fixed, as were the stimu-

lus levels (L_1 and L_2) for both primaries. The frequency ratio f_2/f_1 was varied in approximately 1/36-octave steps from 1.02 to 1.90. In this sequence, the frequency f_1 went from approximately equal to f_2 down to a frequency equal to about one half of f_2 , while the emission at $2f_1 - f_2$ went from a frequency close to f_2 to a frequency near zero (see Fig. 1). In this way, the entire “available range” of the frequency ratio (f_2/f_1) was investigated at each primary frequency (f_2). Typical results are shown in Fig. 1. Frequency-response curves were typically constructed for several different sets of stimulus amplitudes ($L_1 \times L_2$), including equal level ($L_1 = L_2$) and unequal levels with L_1 20 dB higher than L_2 (denoted $L_1/L_2 = 20$ dB). As noted above, stimulus levels were generally chosen to be as low as possible while still giving emission amplitudes sufficiently above the noise floor for three-tone measurements, over at least part of the f_2/f_1 range. For most animals, stimulus amplitudes chosen included a set of four levels: $L_1 \times L_2 = 50 \times 30$, 50×50 , 60×40 , and 60×60 dB SPL.

Following analysis of the two-tone emission frequency-response curves, several different primary frequency ratios were chosen for three-tone investigation. These were chosen to sample different regions of the frequency response curve, i.e., to sample different “peaks.” Two-tone input–output functions for each chosen primary frequency and intensity ratio were then measured to be sure that the chosen primary stimulus level was *below* the notch typically found in the input–output function. Figure 2 shows examples of such notches and of the definition of the “notch level,” L_N , i.e., the level where $L_1 = L_N$. Previous results at $f_2 = 8$ kHz had suggested that optimal results for STCs would be found for cases where the L_1 stimulus level was at least 25 dB below L_N (Mills, 1998). This requirement was relaxed for purposes of the present study, because it would have been impossible to measure emissions over a wide range in frequency ratio with the more stringent requirement. In any case, the preferred values for L_1/L_N for STCs for f_2 frequencies other than 8 kHz have not yet been established. For the data presented here, the L_1 levels were always below the notch level, usually by 10–15 dB or more.

For construction of a suppression tuning curve (STC), the primary stimulus pair was fixed. The third tone frequency (f_3) was temporarily fixed, and the stimulus amplitude (L_3) brought up in steps of 3–5 dB starting at very low levels. The maximum level for L_3 was usually 85 dB SPL. The resulting function, of primary emission amplitude versus “suppressor” level (L_3) for fixed f_3 has been termed a “suppression growth function” (e.g., Abdala *et al.*, 1996). Examples of suppression growth functions are shown in Figs. 3 and 5.

Note that, in addition to monitoring the 8-kHz stimulus, the “unsuppressed” primary emission was always observed at the start of each suppression growth function. This provided a natural, additional way to provide a frequent monitoring of cochlear health. Typically, the unsuppressed primary emission amplitude did not vary more than 1–2 dB throughout the time required to construct a complete set of three-tone responses, about 10–15 min. If it did so, the data were not used in the analysis.

Two-tone frequency response curves for f_1 varied, f_2 fixed at:

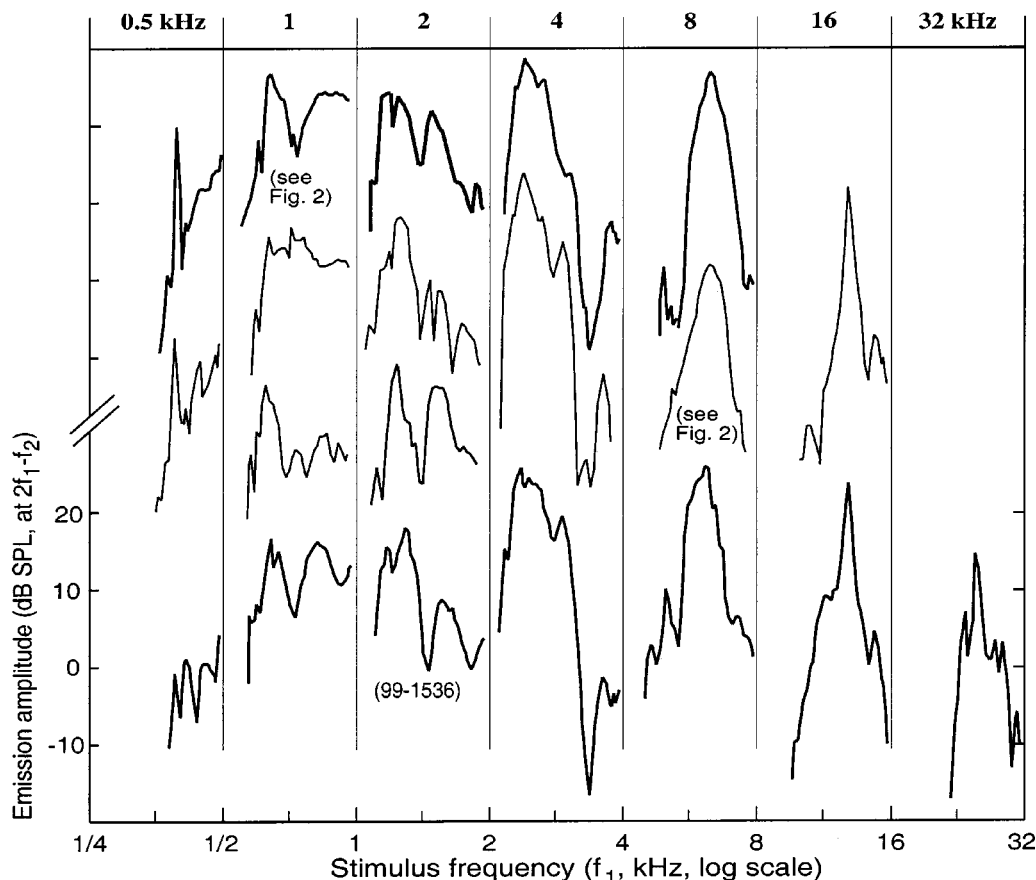


FIG. 1. Emission amplitude response (at $2f_1-f_2$, vertical axis) with fixed upper stimulus frequency, f_2 , while the lower frequency, f_1 , is varied (horizontal axis). The f_2 values were chosen at one-octave intervals, and are noted at the top of the figure. The lowermost curves are for one individual for the case where the stimulus amplitudes were $L_1 \times L_2 = 60 \times 40$ dB SPL, and the resulting emission amplitude is noted on the left axis. Other curves are from other subjects in the study, most of which were studied at only one f_2 frequency. The other curves are shifted vertically for clarity. Typical input-output functions, for the two curves noted, are presented in Fig. 2. Typical phase variations for some of these emission frequency-response functions are shown later. The variations in the curves shown here were not due to noise, as all emission amplitudes plotted were well above the noise floor. As noted in Methods, emission responses which were near or below the estimated noise floor were not included in any data presented in this report.

Third-tone frequencies (f_3) were chosen ranging from two to four octaves below f_2 to two to four octaves above, depending on the f_2 frequency, with typically three to five frequency steps per octave. The frequencies for the third tone were chosen carefully so that the third tone (f_3) was never equal to f_1 or f_2 or to their common distortion products, such as $2f_1-f_2$ and $2f_2-f_1$; and so that the distortion products such as $2f_3-f_1$ or $2f_2-f_3$ were never equal to the primary frequencies or to their common distortion product frequencies. This last constraint required particular attention because the third-tone stimulus level (L_3) was allowed to reach high levels in these measurements, creating a large possible number of measurable distortion products of this type. In order to facilitate the process of choosing useful f_3 frequencies, only a limited number of f_2/f_1 ratios were routinely used for three-tone studies. These were $f_2/f_1 = 1.05, 1.11, 1.25, 1.33, 1.44, 1.60$, and 1.70 . Further, f_2 frequencies were chosen in octave intervals from 0.5 to 32 kHz, and specific sets of f_3 frequencies were calculated, tested, and used consistently for each f_2 frequency. If it was suspected that a previously unknown distortion product of f_3 with f_1 or f_2 was interfering with the response at $2f_1-f_2$, the chosen third-tone frequency (f_3) was changed very slightly. If the suspicious be-

havior disappeared, the interference was considered confirmed and the new f_3 frequency employed in subsequent data collection. For example, this was the conclusion in some cases of suppression growth functions in which the emission amplitude first decreased, at the usual third-tone levels, but then increased again as the third-tone level became large.

Each three-tone response set consisted of 200–300 data points, consisting of primary emission amplitudes as a function of the third-tone frequency and intensity. Changes by a criterion amount in such emission amplitudes define a suppression tuning curve (STC). Off-line, preliminary contour maps representing the three-tone responses were constructed using the SYSTAT (Evanston, IL) program with the “Nexpo” curve-fitting algorithm. Such a curve-fitting program, designed to represent potentially noisy data, is generally very useful for constructing emission contour maps but suffers from an intrinsic problem. Any such program will necessarily tend to average over regions which actually contain sharp changes in emission amplitudes in two dimensions, e.g., the region containing the sharp, sensitive “tip” of the suppression tuning curve. This averaging has the effect of decreasing the depth of the tip. Therefore, all preliminary contour maps were corrected by direct consideration of the measured sup-

pression growth functions in those regions where the initial SYSTAT-generated map indicated there were rapid changes in emission amplitudes (i.e., in two dimensions). Finally, for clarity all emission responses which were less than 5 dB above the noise floor were removed from the data presented in this report. The noise floor was conservatively estimated by an analysis of the seven frequency bins on each side of the emission frequency ($2f_1 - f_2$), and the estimate automatically recorded with each emission measurement.

Animal use and procedures for this study were approved by the Animal Care Committee at the University of Washington.

II. RESULTS

Figure 1 presents a summary of the results of the two-tone frequency-response curves observed in this study. Each curve is the emission ($2f_1 - f_2$) amplitude as a function of f_1 frequency for a fixed f_2 frequency. As noted above, the emission frequency is zero when f_1 is equal to half f_2 ; therefore, plotted in this manner each curve covers just one octave. Choosing f_2 frequencies at octave intervals then allows all such curves to be plotted on one graph without overlap. The f_2 frequencies are noted at the top of Fig. 1. In all cases shown in this figure, stimulus amplitude ratios were $L_1/L_2 = 20$ dB. In most cases, the stimulus levels were $L_1 \times L_2 = 60 \times 40$ dB SPL, but the upper two 8-kHz curves were for $L_1 \times L_2 = 50 \times 30$ dB SPL. The bottom-most row of curves gives the emission response functions for one animal, with stimulus levels $L_1 \times L_2 = 60 \times 40$ dB SPL for all f_2 frequencies. The resulting emission amplitude (dB SPL) is noted on the left axis. All other curves are shifted vertically for clarity, but presented at the same relative scale.

At a given f_2 , the frequency-response curves between animals showed some variation, but there appeared to be some general trends. For low f_2 frequencies, there were typically two or more prominent peaks in the frequency-response function. Only at 8 kHz was there a single, prominent peak which had a reasonable width to represent a frequency-tuning curve, i.e., a width of about a half octave at 10–20 dB below the peak amplitude. At frequencies above 8 kHz, the peaks were very sharp, and sometimes multiple.

In Fig. 1, two tuning curves, one for $f = 1$ kHz and one for $f_2 = 8$ kHz, are marked as referring to Fig. 2. Figure 2 shows input–output functions for these two cases over the range of stimulus frequency ratios. Each input–output, or “growth,” function is measured for constant stimulus frequencies, with the ratio (f_2/f_1) the parameter noted on the figure, and with an amplitude ratio of $L_1/L_2 = 20$ dB. Note that there was a “notch” in each growth function. The notch was at the level $L_1 = 70$ dB SPL for the 1-kHz growth functions, and at $L_1 = 65$ dB SPL for the 8-kHz growth functions. As noted in Methods, similar growth functions were obtained in every case before three-tone responses were measured, to be sure that the primary stimulus amplitudes chosen for three-tone measurements were 10 dB or more below this notch level, L_N . The vertical arrows in Fig. 2 indicate the primary stimulus amplitudes that were chosen for subsequent three-tone measurements in these two cases. Note that immediately below the growth function curves is shown the varia-

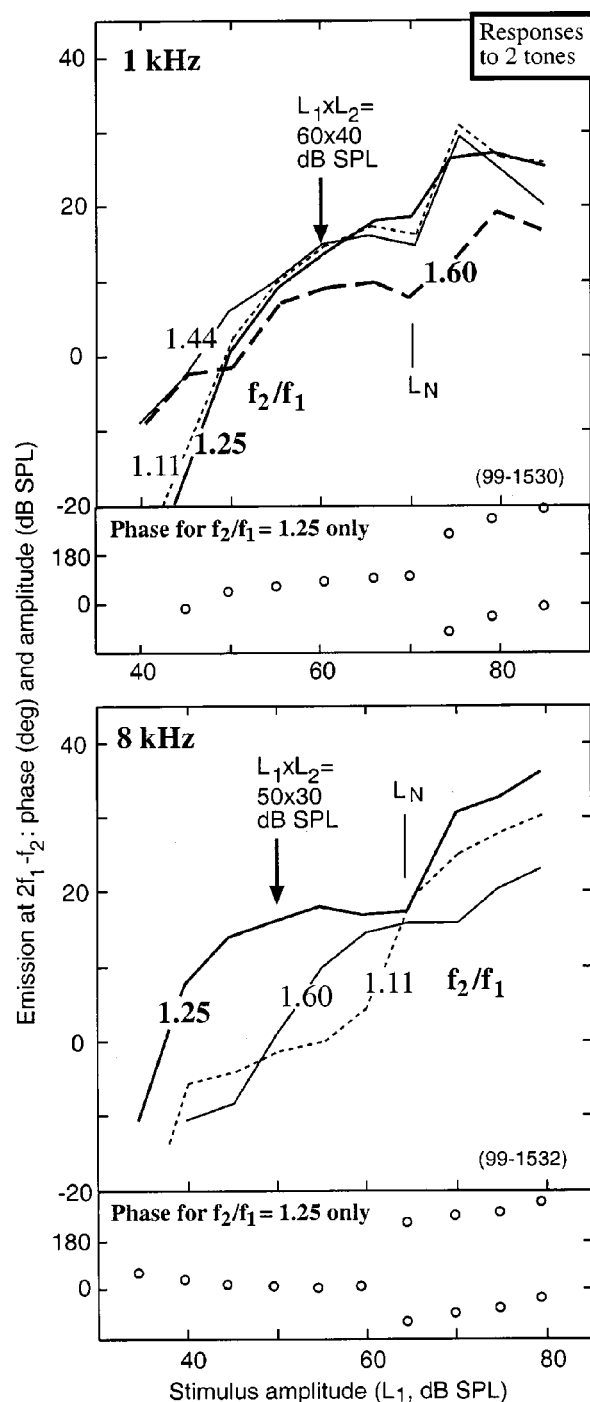


FIG. 2. Typical input–output, or “growth” functions related to two of the frequency-response curves noted in Fig. 1. The growth functions were obtained by fixing the stimulus frequency ratio (f_2/f_1 , parameter noted) and the stimulus amplitude ratio, which for these curves was $L_1/L_2 = 20$ dB SPL. The resulting emission response (at frequency $2f_1 - f_2$) is shown on the vertical axis with the stimulus level (L_1) on the horizontal axis. In each pair of plots, the emission amplitude responses are shown in the upper panel (lines, dB SPL) and a typical phase response is depicted immediately below (open circles, deg). Only the phase response for the case $f_2/f_1 = 1.25$ is shown. The phase response is displayed for a total of $1\frac{1}{2}$ turns, so that ambiguities in angle can be displayed. Only when there was ambiguity are two points shown, separated by 360 deg. The zero-phase reference is arbitrary. Note that there was a relative minimum in the amplitude, a “notch,” at the L_1 level noted by the symbol, L_N . Typically, associated with the notch in amplitude was an abrupt phase shift of about 180 deg, as illustrated. The vertical arrows indicate the stimulus amplitude pairs ($L_1 \times L_2$) for which the frequency-response curves in Fig. 1 were obtained.

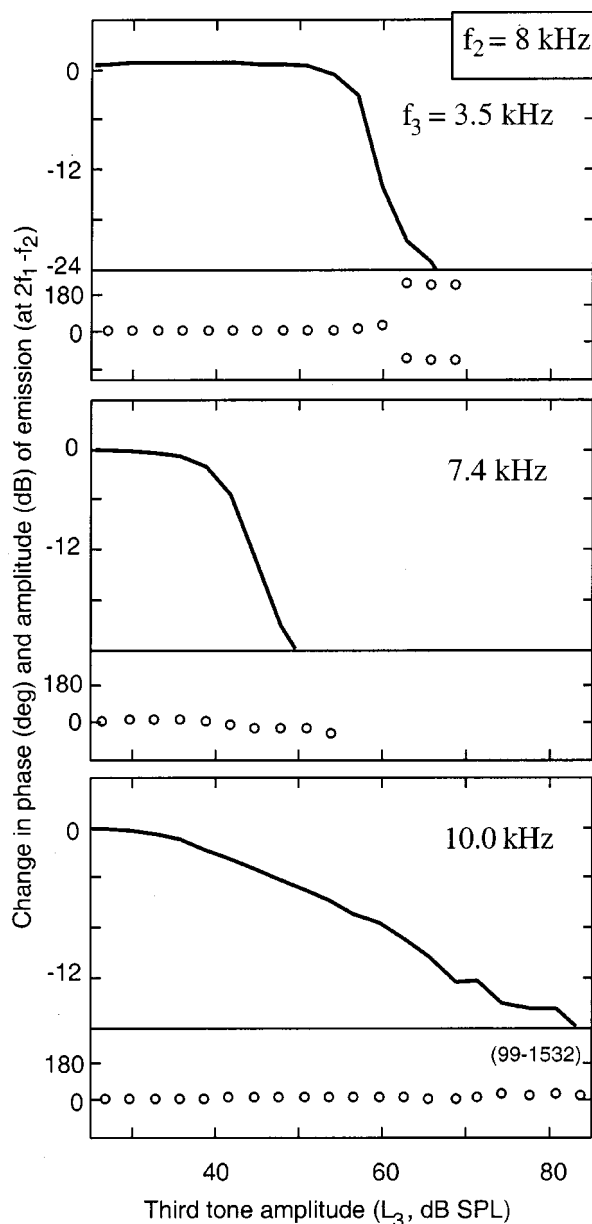


FIG. 3. Emission amplitude (at $2f_1 - f_2$) for a fixed primary stimulus in response to a third tone varied in amplitude and frequency. In this case, the primary stimulus parameters were $f_2 = 8$ kHz, $f_2/f_1 = 1.25$, and $L_1 \times L_2 = 50 \times 30$ dB SPL. Typically, the third-tone frequency was fixed momentarily, as noted by the values for f_3 listed in each panel. The third-tone amplitude (L_3) was increased in steps of 3 or 5 dB, usually over the range from about 20 dB SPL to a maximum of 85 dB SPL. Note there is a pair of panels for each case. The amplitude response of the emission is shown in the upper panel of each pair, as the change from the undisturbed primary emission amplitude. The phase response of the emission is shown immediately below, also as the change with the third tone on as compared to off. The phase responses are shown with a vertical extent of $1\frac{1}{2}$ turns, i.e., 540 deg, to accommodate those situations in which there is a 360-deg ambiguity in the phase angle, as in the top case. For clarity, *only* when there is a possible ambiguity were there two equivalent points shown in the phase data; otherwise, one of the two points was omitted.

tion in emission phase angle, by the open circles, for the case $f_2/f_1 = 1.25$ only. As usual (e.g., Mills, 1997) there was a shift of about 180 deg in the emission phase angle when the notch level was traversed.

Once the primary stimulus parameters were chosen, three-tone “growth of suppression” functions were obtained

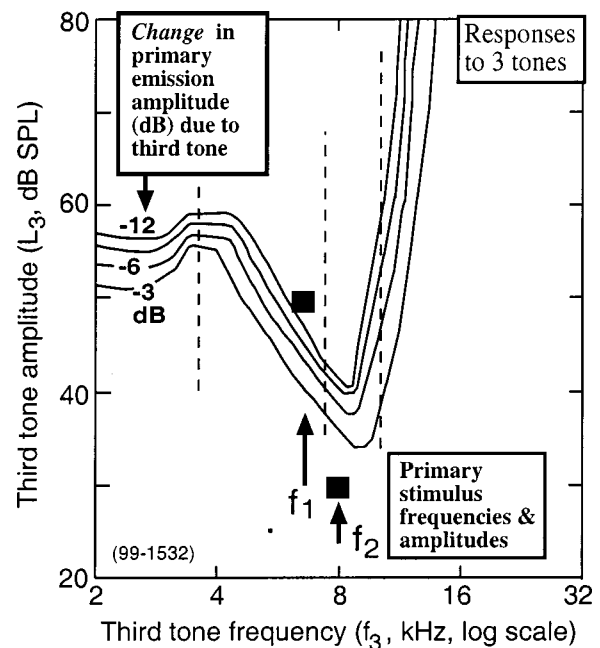


FIG. 4. Three-tone emission response map for $f_2 = 8$ kHz, representing the change in the emission (at $2f_1 - f_2$) with fixed primaries (f_1, f_2) resulting from the variation of a third tone in frequency (f_3 , horizontal axis) and intensity (L_3 , vertical axis). Same case as Fig. 3; the vertical dashed lines indicate the location of the growth of suppression functions shown in Fig. 3. The solid lines represent the loci of points which resulted in a criterion change in the emission amplitude. These contour lines start at a decrease of 3 dB down to 12 dB in 3-dB steps. If there were to be an *increase* in the emission amplitude, as there is in some cases shown later, contours would start at +3 dB and step upward in 3-dB steps. For reference, the frequencies and levels of the fixed primary stimulus pair are indicated by the two filled squares. Note that all three axes are logarithmic.

for 20–30 different f_3 frequencies. Three typical growth of suppression functions are illustrated in Fig. 3. This shows the case for $f_2 = 8$ kHz, known from previous work in the gerbil to result in STCs which are similar in form to neural-tuning curves, i.e., that are not “complex” in any important way (Mills, 1998). This is the same case that was shown in the lower panel of Fig. 2. The primary stimulus frequency ratio chosen for Fig. 3 was $f_2/f_1 = 1.25$, and the primary stimulus amplitudes were $L_1 \times L_2 = 50 \times 30$ dB SPL. From top to bottom, this figure shows growth of suppression functions for the suppressor frequency, f_3 , about an octave below f_2 , in the next panel for f_3 about equal to f_2 , and in the lowest panel for f_3 a half octave above f_2 . Consider first the middle panel of Fig. 2, for which the third-tone frequency (f_3) was between the primary stimulus frequencies f_1 and f_2 . The resulting growth of suppression curve had a steep falloff occurring for a relatively low value of L_3 . For f_3 frequencies well below the primary frequencies (upper plot), the falloff was equally steep but required much higher third-tone stimulus levels, L_3 , before it began the falloff. This defines the low-frequency “tail” of the tuning curve. Note the shift in phase angle of about 180 deg at high L_3 levels. Such phase shifts at similar third-tone frequencies (f_3) were a frequent, but not invariant, finding.

The lowest panel in Fig. 3 illustrates the “simple” response of a primary stimulus to a third tone with a frequency about a half octave above the primary stimulus frequencies.

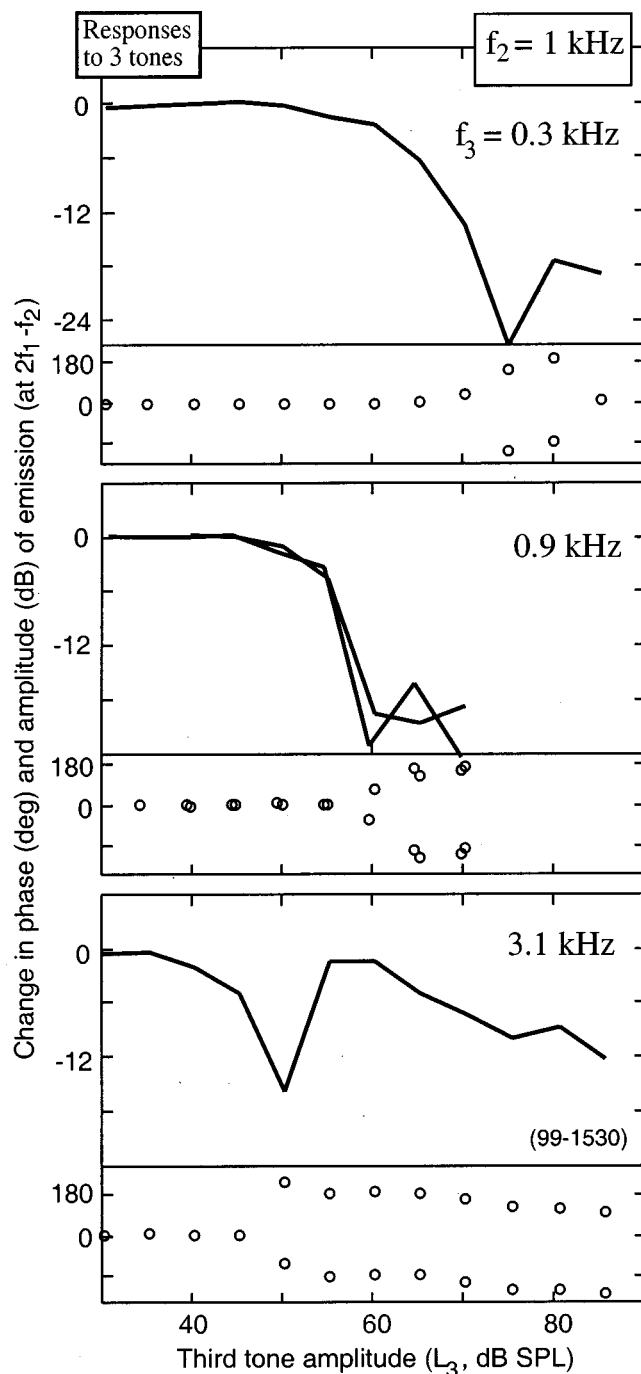


FIG. 5. Growth of suppression functions which showed complex behavior. For primary stimulus parameters $f_2 = 1$ kHz, $f_2/f_1 = 1.25$, and levels $L_1 \times L_2 = 60 \times 40$ dB SPL. Same conventions as Fig. 3. Note that, for $f_3 = 0.9$ kHz (middle panel), growth of suppression functions were measured twice, once at the beginning and once in the middle of the data set, and both functions are shown.

The emission amplitude fell off slowly and monotonically with L_3 level, with the decrease starting at a relatively high level. There was little associated change in the emission phase angle. At even higher third-tone frequencies (not shown), there was no significant effect on the primary emission for any permitted third-tone amplitudes, i.e., up to $L_3 = 90$ dB SPL. In sum, Fig. 3 exemplifies the simple behavior expected for the case that an emission STC can be considered similar to a neural FTC.

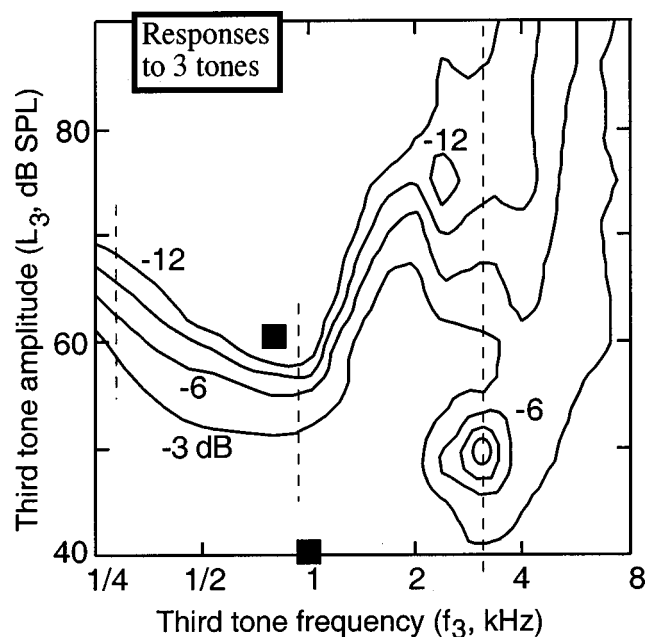


FIG. 6. Three-tone emission response map for $f_2 = 1$ kHz, same animal as Fig. 5. Same conventions as Fig. 4. In addition, note that whenever the slope of the map was potentially ambiguous, as it would be for the closed contours in the lower right of the map, the direction of the change was indicated by labeling one or more of the contours. Any reverses in slope are also noted. Since no other contours are labeled for the closed contours at lower right, one can assume that these closed contours represent a small "hole," i.e., with contour -6 dB as noted, followed by the -9 -dB contour, with -12 dB at the center. The dashed vertical lines represent the fixed frequencies of the "growth of suppression" functions displayed in Fig. 5. These functions represent cross sections of the map in Fig. 6 at those frequencies. The meaning of the holes seen in Fig. 6 may be seen more clearly by inspecting the lower panels in Fig. 5.

As detailed in Methods, a complete set of 20–30 suppression functions was used to construct a contour map which represented the response of the primary emission amplitude to the third-tone frequency and intensity. Figure 4 shows the result for the case in Fig. 3. Note, in Fig. 4, the line marked " -3 dB" represents the locus of all points (in f_3 and L_3 coordinates) in which the third tone was estimated to cause a decrease of primary emission amplitude by 3 dB. In all contour maps, the filled squares indicate the corresponding frequencies and levels of the primary stimulus pair, for reference. The location of the three growth of suppression functions shown in Fig. 3 are indicated in Fig. 4 by the dashed vertical lines. The simple emission contour map exemplified by Fig. 4 is similar in shape to neural frequency-tuning curves in the gerbil (Ohlemiller and Echteler, 1990). That is, there is a single, most sensitive tip, located near the primary stimulus frequencies, with a relatively steep high-frequency rise above the tip and a relatively shallow, extended tail on the low-frequency side. As noted in the Introduction, even when the shapes are similar, of course, there may be considerable quantitative differences between these two types of tuning curves, STC and FTC.

In contrast to the case for $f_2 = 8$ kHz presented in Figs. 3 and 4, Fig. 5 shows typical growth of suppression functions for the case with $f_2 = 1$ kHz. This is the same case shown in the upper panel of Fig. 2. The upper panel in Fig. 5 shows the growth of suppression for a third-tone frequency well

below the primary stimulus frequencies. Note that it was very similar to that in Fig. 3 at an equivalent frequency, i.e., at an octave below the primary frequency. The middle panel is also similar to Fig. 3, except that in this case a 180-deg phase shift was seen. The “threshold” for the third tone was about 50 dB SPL in this case. The lowest panel in Fig. 5, however, demonstrates a clearly complex response. The third-tone stimulus frequency, f_3 , in this case was more than one and a half octaves above the primary stimulus frequency, f_2 . This should be contrasted to the lower panel in Fig. 4, where f_3 was only a half octave above f_2 . For the complex response seen in the lower panel of Fig. 5, there was a sharp decrease in the primary emission amplitude that occurred when the third tone was still quite weak, only $L_3 = 40$ dB SPL. The amplitude of the emission was actually reduced by a factor of 4 (to -12 dB) when the third tone was only 50 dB SPL. In this case a 180-deg phase shift in the primary emission was associated with the complex suppression. Such a phase shift was frequently, but not invariably, observed with such complex three-tone behavior. Note that the observed phase shifts occurred in the primary emission, without any changes in the two stimuli producing this emission, by variation only of the *amplitude* of the third tone.

Following the “hole” in the emission amplitude seen near $L_3 = 50$ dB SPL, the emission amplitude returned to near its unsuppressed level as the third-tone amplitude, L_3 , increased. Observations at nearby f_3 frequencies confirmed that, in fact, the complex behavior did form a hole in the contour map, as shown in Fig. 6.

Comparing Fig. 6 to Fig. 4, it can be seen that the 1-kHz map was not strikingly different from the 8-kHz map for third-tone frequencies (f_3) near and below the primary stimulus frequencies. Starting about one octave above the primary frequencies, however, and continuing until nearly three octaves above them, the behavior in Fig. 6 was clearly unusual. There is no obvious reason why the relatively weak, 40-dB SPL tone should have caused a change in the emission coming from primaries which were *two octaves lower* in frequency. The behavior at the higher f_3 frequencies in Fig. 6 typifies the kind of behavior in this report that is termed “anomalous” or “complex,” in contrast to the simple three-tone map of the form illustrated in Fig. 4.

An overview of the complex behavior found in the three-tone maps is presented in Fig. 7. Conventions are the same as in preceding maps. Contours go from -3 dB in the lower part of the maps to -12 dB at the upper side. The f_2 primary frequencies are noted in the upper left of each map, and proceed in octave steps from 0.5 kHz in the top panel to 8 kHz for the bottom panel. It is clear that the frequency (f_3) region in which the complex emission is found is only weakly dependent on the primary stimulus frequencies. The shaded vertical lines indicate that the majority of the complex behavior is found for f_3 values between 1.4 and 8 kHz.

There is clearly complex behavior observed at all primary frequencies except $f_2 = 8$ kHz. Note that this does not imply that the 8-kHz maps are *not* affected by such processes. The fact that the low-frequency growth of suppression functions (Fig. 3) often did show a characteristic 180-deg phase shift indicates that the effects responsible for the

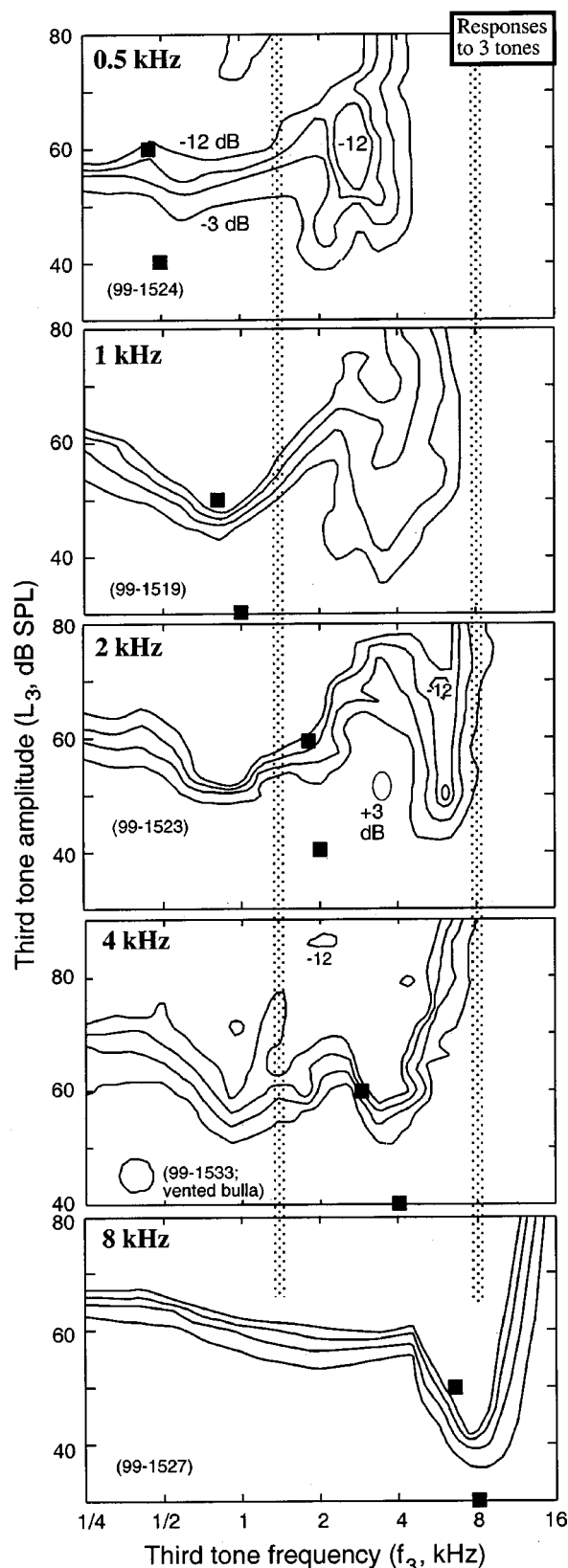


FIG. 7. Three-tone emission response maps over the range of f_2 frequencies considered in detail for this report. Same conventions as Figs. 4 and 6. The f_2 frequency is listed in the upper corner of each map. The primary frequency ratio was usually $f_2/f_1 = 1.11$ or 1.25, and the primary amplitude ratio was always $L_1/L_2 = 20$ dB, absolute values are indicated by the two filled squares in each map. There was only one region where there was an increase in emission amplitude in these five maps; this was in the center map and is indicated by the $+3$ -dB contour.

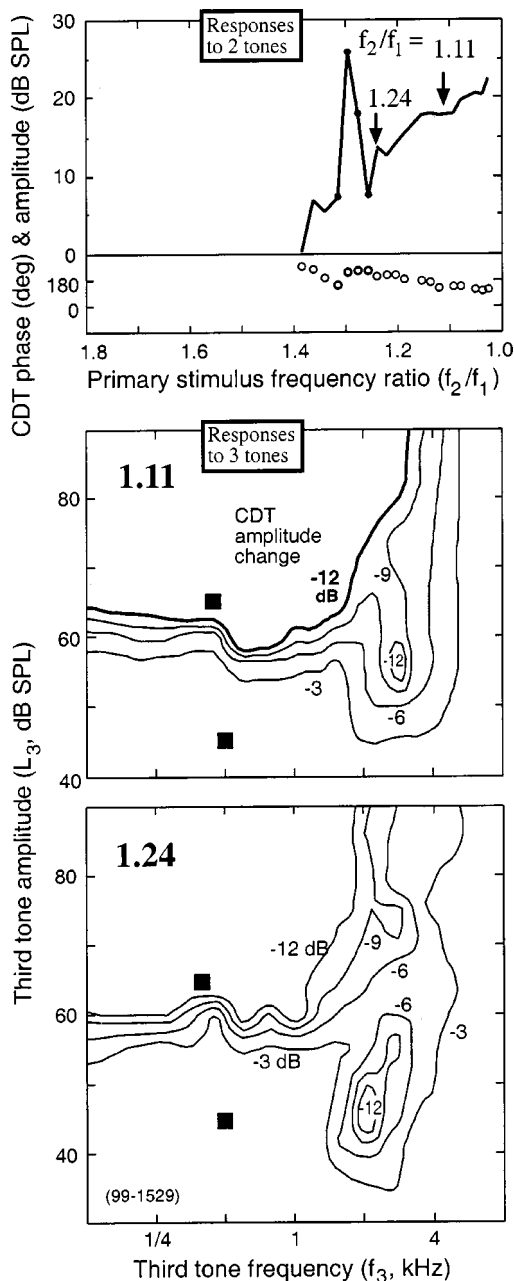


FIG. 8. Typical two- and three-tone frequency responses for the primary frequency $f_2 = 0.5$ kHz and for levels $L_1 \times L_2 = 65 \times 45$ dB SPL. The top pair of panels presents the two-tone frequency-response curve, with the horizontal axis the primary stimulus frequency ratio, f_2/f_1 . The direction of this axis is reversed so that low f_1 frequencies, and low-emission frequencies, lie to the left, as they do in similar plots in Fig. 1. The emission ($2f_1 - f_2$) amplitude is given by the upper curve, in dB SPL, with the associated phase angles shown immediately below by open circles. The phase angle reference is arbitrary. The contour maps in the lower panels present the three-tone responses for the frequency ratios (f_2/f_1) noted in the top panel by vertical arrows. Same conventions as previous maps (Figs. 4, 6, and 7). See the text for meaning of highlighted or heavier lines and data points.

complex suppression may have been present for 8-kHz primaries, just not as obvious as at lower f_2 frequencies.

In the 2-kHz map (center panel of Fig. 7), an increase in emission amplitude above the “unsuppressed” level was also found for certain values of the third-tone stimulus. This increase is indicated by the lighter contour line, also identified as the +3-dB contour. The same convention is followed

in subsequent maps whenever such an “enhancement” is found.

The routine strategy in this study was to conduct an intensive examination of the two- and three-tone responses associated with just one f_2 frequency in a given animal. For each animal, this typically resulted in a total of four two-tone frequency-response curves and 16 three-tone sets, each three-tone set constituting an emission suppression tuning curve (STC). A selection of typical responses is presented in Figs. 8–13, with each figure representing part of the data obtained from one animal.

With the exception of Fig. 10, each of Figs. 8–13 follows the same format. The top panel presents the two-tone frequency response curve (f_2/f_1 ratio function) for the emission at $2f_1 - f_2$. Note that the convention in the literature has been to report such data using the frequency ratio, f_2/f_1 . When holding f_2 constant and varying f_1 , it seems preferable to use the inverse ratio, f_1/f_2 (Mills and Rubel, 1997). This way, lower f_1 frequencies, and lower emission frequencies, would lie to the left in such a display, similar to the way these would normally be presented, e.g., in Fig. 1. As a compromise, so that the frequency ratio data can be immediately understandable to most readers, in the figures in this report the frequency-response curves have been displayed listing the conventional ratio f_2/f_1 , but with the axis direction reversed, so that low f_1 frequencies do lie to the left. In the upper panels in each figure, the emission amplitude response is shown by the upper continuous curve, with associated phase angles indicated by the open circles immediately below.

In each of Figs. 8–13, contour maps representing the three-tone responses are presented below the frequency response curve. The f_2/f_1 ratios chosen for display are noted by the vertical arrows in the frequency-response curves and by the number in the upper left of each contour map. The same conventions are used in each contour map as in Figs. 4, 6, and 7.

Responses for a primary stimulus frequency (f_2) of 500 Hz are illustrated in Fig. 8 for the primary stimulus level pair $L_1 \times L_2 = 65 \times 45$ dB SPL. As can be seen, the frequency-response curve was a relative maximum for nearly equal primary stimulus frequencies ($f_2/f_1 \approx 1$), and tended to decrease to the noise floor for frequency ratios of $f_2/f_1 > 1.3$. Three examples of such frequency-response curves were shown in Fig. 1, in the left column. There was a sharp peak in emission amplitude observed in two of these three animals at $f_2/f_1 \approx 1.3$. These sharp peaks appeared to be a real effect. The peaks were not due to a single “spectral line,” but typically had a finite width and relatively continuous phase angles across them. Note data points emphasized in Fig. 8 in the region of this peak. The sharp peaks sometimes found for $f_2 = 0.5$ kHz are reminiscent of peaks of similar sharpness shown in Fig. 1 for $f_2 = 16$ kHz and, especially, for $f_2 = 32$ kHz. However, for the present study, contour maps were not obtained for f_2/f_1 ratios at or near such peaks because of the difficulty of holding constant an emission amplitude which depended so sharply on the primary parameters.

The upper contour map in Fig. 8 illustrates the map

found for a primary stimulus frequency ratio of 1.11. The map shown was typical for this primary frequency; a similar map was shown in the top panel of Fig. 7, for the same frequency ratio and slightly lower primary stimulus levels. At 0.5 kHz, it was not possible to obtain three-tone responses for the CDT emission at primary frequency ratios above 1.3, because of the rapid decline in the two-tone emission amplitude (as illustrated in Figs. 1 and 8).

The contour maps (Fig. 8) demonstrate the general finding for such low-frequency primaries: The emission STCs were completely dominated by the complex response, and there was very little similarity to a neural FTC. There was not even a relative minimum in threshold in the region of the primary stimulus frequencies, for example, as one might hope to find from simple theoretical arguments (Mills, 1998). Compared to the region of the primaries, the threshold was nearly 20 dB lower for third-tone frequencies two and three octaves higher than the primary frequencies. In common with neural tuning curves, there was a low-frequency tail and thresholds did eventually rise rapidly for third-tone frequencies well above the primary frequencies. This was particularly true if one restricted consideration to the continuous lines representing strong suppression. For illustration, this line is indicated by a heavier weight in the middle panel of Fig. 8. This suggests that one may get “improved” results, i.e., maps which appear more like neural FTCs, by ignoring the closed contours which represent relative minima, or “holes” in the more complex maps. However, even the open contours in Fig. 8 appear to be distorted by the presence of the effect causing the complex suppression.

Typical two- and three-tone responses for the case $f_2 = 1$ kHz are shown in Fig. 9 for the primary stimulus levels $L_1 \times L_2 = 60 \times 40$ dB SPL, and in Fig. 10 for the equilevel case, $L_1 \times L_2 = 55 \times 55$ dB SPL. The frequency-response curve in the top panel of Fig. 9 was found to be divided into two approximately equal, broad peaks. Note that, even though there was a “notch” in the curve, i.e., an amplitude minimum at about $f_2/f_1 = 1.4$, there was no sharp phase change observed as the f_1 frequency was traversed across the notch region. This should be contrasted to the phase jump of about 180 deg usually seen at notches in input-output functions (Fig. 2). Two contour maps are shown here for each of the two broad peaks. Representing the peak for primary stimulus frequencies close together, the contour map for $f_2/f_1 = 1.11$ is shown in the second panel of Fig. 9; the map for $f_2/f_1 = 1.25$ is shown in Fig. 6. These maps have similar characteristics to those noted for the 0.5-kHz maps at similar small frequency ratios.

Results for the peak where the primary stimulus frequencies were farther apart is illustrated by the lower two maps in Fig. 9, with $f_2/f_1 = 1.44$ and 1.58. The complex suppression effects appeared weaker in this case. The maps had characteristics more like normal FTCs, and more like that for 8-kHz emission STCs (Fig. 4). That is, the most sensitive thresholds occurred in the region of the two primaries, at least for criterion decreases exceeding 3 dB. Further, there was a low-frequency tail. At high frequencies, the contours, at least the open contours representing strong suppression, moved rapidly upward, reaching the upper stimulus

limit only an octave above the primary frequency f_2 . The mildly complex contours observed for third-tone frequencies (f_3) between 1 and 3 kHz in Fig. 9 apparently were due to remaining effects of the complex suppression. However, it appears that, for these higher f_2/f_1 ratios at $f_2 = 1$ kHz, it took relatively higher third-tone levels (L_3) to effectively evoke the complex behavior, and it typically caused less change in emission amplitude when the complex behavior did occur.

At these low primary stimulus frequencies, it is particularly appealing to include in the investigation the behavior of the other cubic distortion product, $2f_2 - f_1$. Because its emission frequency is above the primary frequencies, its noise floor tends to be lower than that for the cubic distortion tone (CDT, $2f_1 - f_2$) emission, at least for low f_2 frequencies. This is particularly true as the f_2/f_1 ratio increases, causing a decrease in the CDT frequency. However, for the case of Fig. 9, where the primary stimulus amplitudes are unequal, i.e., for $L_1 \gg L_2$, the other cubic distortion product ($2f_2 - f_1$) is very weak and not useful. For approximately equal-level primaries, $L_1 = L_2$, however, both upper and lower cubic emissions are typically of moderate amplitude, and so the maps can be usefully compared. Typical results are presented in Fig. 10, for the case $f_2 = 1$ kHz and with the primary amplitudes both equal to 55 dB SPL.

The results displayed were typical for these two emission components for equal-level primaries across f_2 frequencies and f_2/f_1 ratios. The frequency-response curve (f_2/f_1 ratio function) for the $2f_2 - f_1$ emission was generally much flatter than that for the $2f_1 - f_2$ emission. For all frequencies, these two emission components were approximately equal in amplitude at small primary frequency ratios, that is, when $f_2 f_1 \rightarrow 1$. Note that this is expected mathematically, as long as the emission amplitude is a continuous function of (f_2/f_1) as (f_2/f_1) gets close to unity. The result is that, for $f_2 = 0.5$ kHz, the higher-frequency cubic emission ($2f_2 - f_1$) remained measurable at higher f_2/f_1 ratios, where the CDT emission was much weaker. In general, however, as the f_2 frequency increased, the CDT emission amplitude became relatively weak in the limit as $f_2/f_1 \rightarrow 1$ (e.g., Fig. 1). The consequence is that the higher-frequency distortion component ($2f_2 - f_1$) became relatively weaker than the CDT ($2f_1 - f_2$) at the mid f_2/f_1 ratios at higher f_2 frequencies.

The two contour maps shown in Fig. 10 illustrate maps for the two cubic emission frequencies measured in the same signal, at the frequency ratio $f_2/f_1 = 1.44$. These maps illustrate the general results for similar maps measured at all f_2 frequencies and a wide range of f_2/f_1 ratios. Compare the equal-level CDT contour map for $L_1 \times L_2 = 55 \times 55$ dB SPL in Fig. 10 to the unequal case in Fig. 9 with $L_1 \times L_2 = 60 \times 40$ dB SPL and $f_2/f_1 = 1.44$. Note that these two cases had approximately the same total energy in the primaries. As these two examples illustrate, when there were similar conditions the complex suppression effects were usually *more extreme* for equal-level primaries ($L_1 = L_2$) than for unequal levels with $L_1/L_2 = 20$ dB.

Now, compare the two contour maps in Fig. 10. The complex effects in the map of the emission at $2f_2 - f_1$ were generally more extreme than those taken at the same time at

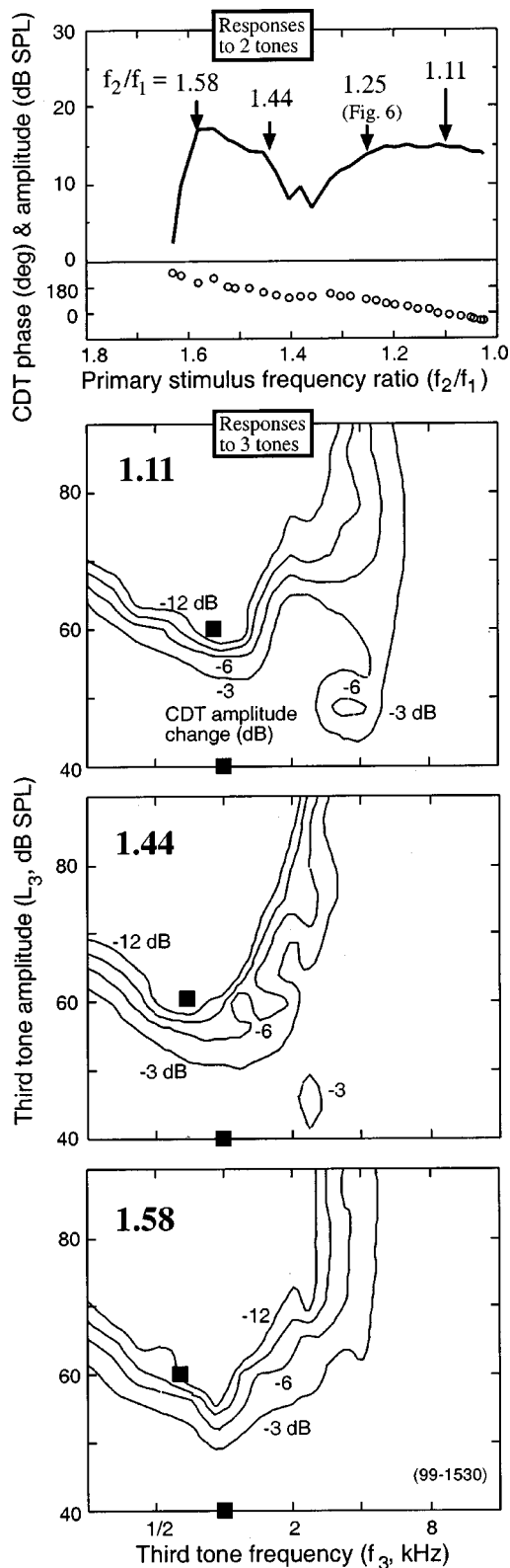


FIG. 9. Typical two- and three-tone frequency responses for primary frequency $f_2 = 1$ kHz and for levels $L_1 \times L_2 = 60 \times 40$ dB SPL. Same conventions as Fig. 8.

the frequency $2f_1 - f_2$. Maps for equal-level primaries were constructed for all f_2 frequencies and a range in frequency ratios, and this result was found to be general. That is, whenever the higher cubic emission ($2f_2 - f_1$) was measurable, its contour map showed complex behavior that was at least as

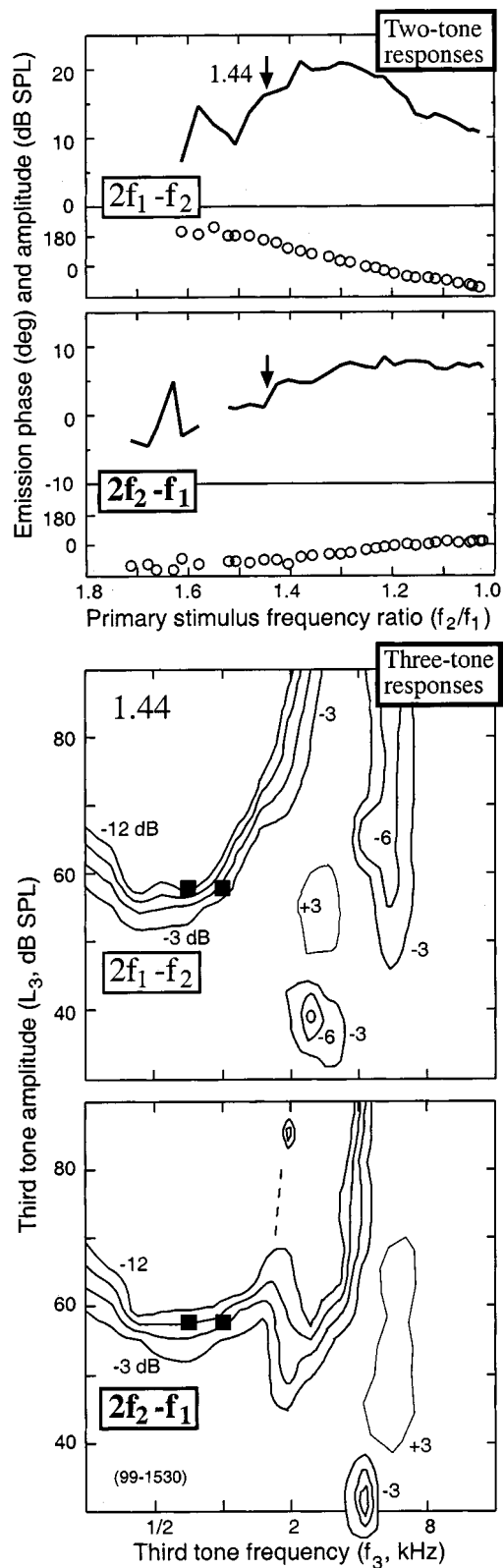


FIG. 10. Two- and three-tone frequency responses recorded at the same time as the two cubic distortion frequencies: the CDT emission at frequency $2f_1 - f_2$ (upper plot in each set) and the higher frequency cubic emission at $2f_2 - f_1$ (lower plot in each set). Same animal and f_2 frequency as in Fig. 9, but for equal level primaries, $L_1 = L_2 = 55$ dB SPL. Missing data points in the top panels indicate the emission signal was contaminated by noise. The same conventions are used as in previous figures. Note regions of emission enhancements that are often seen with equal-level primaries (lines denoted by lighter contours and labeled $+3$ dB).

extreme as that for the CDT emission ($2f_1 - f_2$) measured in the same signal.

Figure 11 presents typical results for the case $f_2 = 2$ kHz. The upper panel shows that there were two broad, prominent peaks in the two-tone frequency response curve, and possibly some smaller peaks present as well. Contour maps for three-tone responses are shown below, for increasing primary stimulus frequency ratios at $f_2/f_1 = 1.11$, 1.25, and 1.60. These showed a similar trend to the 1-kHz response, in that the effects of complex suppression appeared to decrease as the primary frequency ratio, f_2/f_1 , increased. Note that the major effects of the complex suppression occurred near $f_3 = 4$ kHz (middle two panels of Fig. 11).

The effects seen for primary frequencies of $f_2 = 4$ kHz are illustrated in Fig. 12. The trend seen in the two-tone frequency-response curves earlier continued, to the extent that the emission was nearly undetectable for small primary stimulus frequency ratios, leaving only a very broad major peak for higher-frequency ratios, i.e., from $f_2/f_1 = 1.4$ to 1.8. In comparison to the 2-kHz results (Fig. 11), the emission at this broad peak for $f_2 = 4$ kHz was quite strong for the same stimulus levels (i.e., $L_1 \times L_2 = 60 \times 40$ dB SPL), reaching an amplitude of about 25 dB SPL (also see Fig. 1, lower curves).

In contrast, the three-tone responses found in the $f_2 = 4$ kHz maps did not appear to be a simple continuation of earlier trends. The middle map in Fig. 12 illustrates most clearly the major complex effect found, being a very sharp increase in threshold for third-tone frequencies near 2 kHz. The high-frequency side of this increase was very sharp, but the low-frequency side was more gradual. Additional, close-spaced f_3 frequencies confirmed that the slope was nearly constant, as depicted.

It has been suggested that the gerbil middle ear is resonant for stimulus frequencies near 4 kHz, which acts to improve the transmission of low-frequency sound into the inner ear (Plassmann and Kadel, 1991). This would suggest that opening the bulla widely, as was routinely done in these studies, might alter the middle-ear transmission sharply near the resonant frequency. For this reason, studies in some animals were conducted with the bulla first vented (see Methods), then measurements repeated in the same animal after the bulla was opened widely. This was done for the animal in Fig. 12. As the bottom two maps indicate, venting the bulla so that it was in its more normal state reduced, but did not eliminate, the sharp increase in threshold near $f_3 = 2$ kHz. It also moved this peak to slightly lower f_3 frequencies.

Comparing the maps for the *vented* bulla at $f_2/f_1 = 1.44$ to that at 1.70 in Fig. 12 (second and fourth panels) illustrates again that higher primary stimulus frequency ratios led to three-tone responses which showed less extreme complex effects. It is also important to note that only for $f_2 = 4$ kHz did the fact that the bulla was vented or was open result in complex suppression effects which were markedly different between these two conditions.

One map for $f_2 = 8$ kHz has already been shown (Fig. 4); detailed two- and three-tone data for another animal are illustrated in Fig. 13. Maps at $f_2/f_1 = 1.25$ and 1.44 showed no evidence of any complex suppression effects. However,

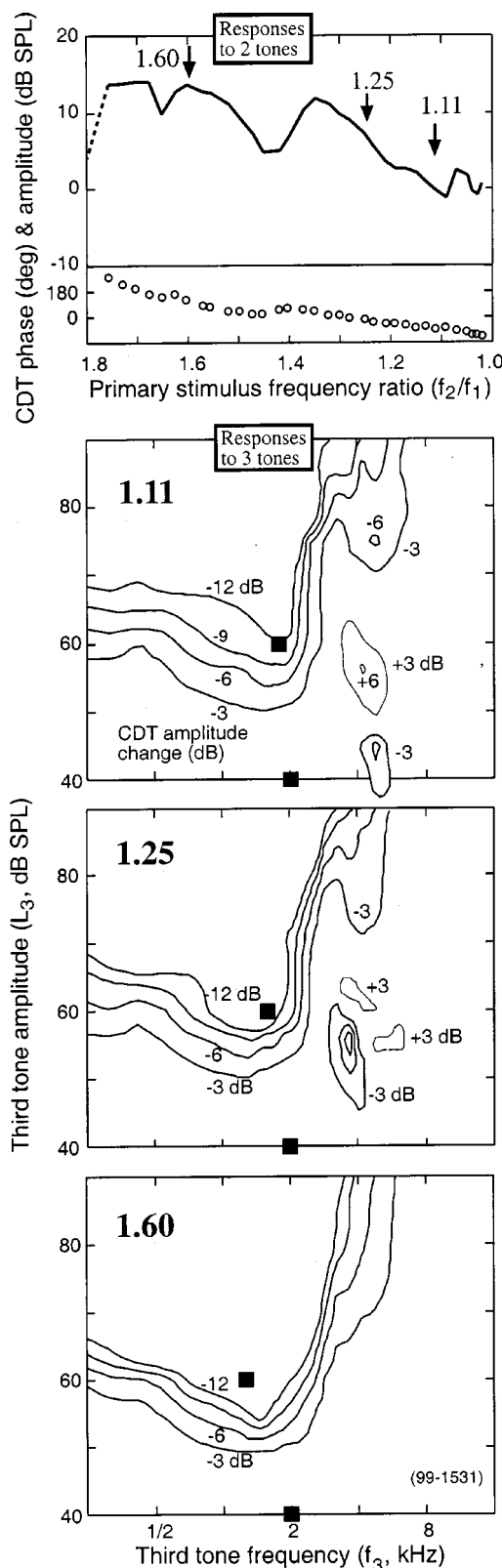


FIG. 11. Typical two- and three-tone frequency response curves for the primary frequency, $f_2 = 2$ kHz, and for stimulus levels $L_1 \times L_2 = 60 \times 40$ dB SPL. Same conventions as previous figures.

the modest distortion of contours in the $f_2/f_1 = 1.11$ map may indicate a remnant of such complex behavior. As at lower f_2 frequencies, the trend was for the maps to show fewer effects of complex suppression for higher f_2/f_1 ratios.

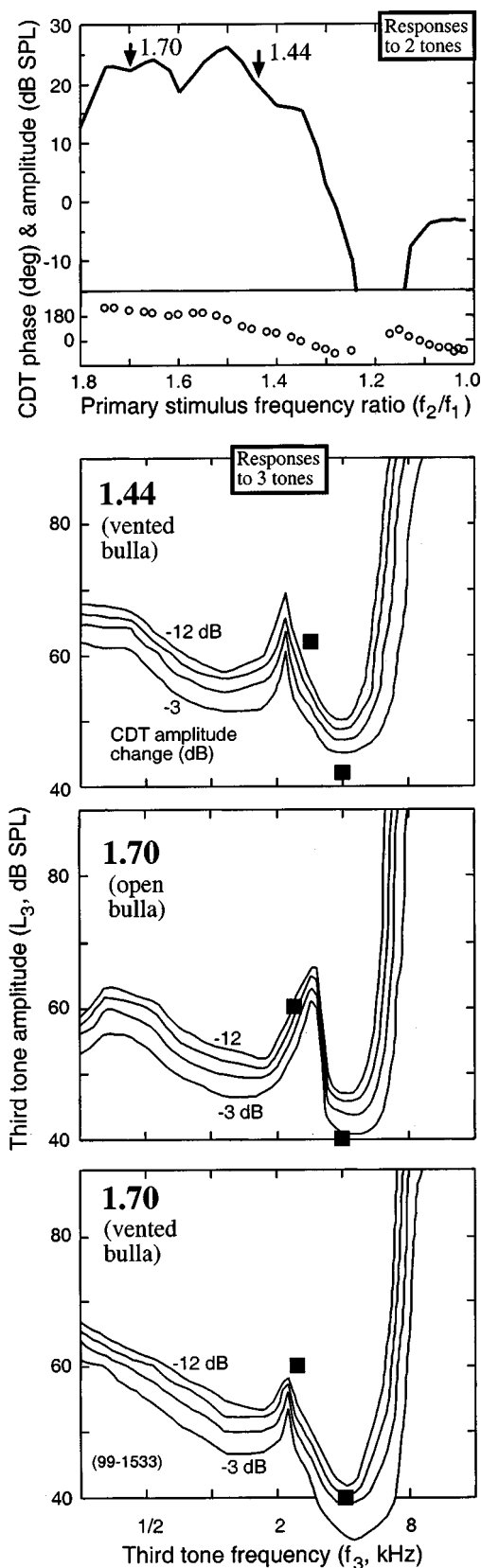


FIG. 12. Typical two- and three-tone frequency-response curves for the primary frequency, $f_2 = 4$ kHz, and for stimulus levels $L_1 \times L_2 = 60 \times 40$ dB SPL. Same conventions as previous figures. The middle map shows results at the primary stimulus frequency ratio $f_2/f_1 = 1.70$ for the bulla opened widely, as it was in other figures presented in this report. For comparison, the maps above and below were obtained with the bulla *vented* instead to obtain a more normal middle-ear transfer function. See Methods for details.

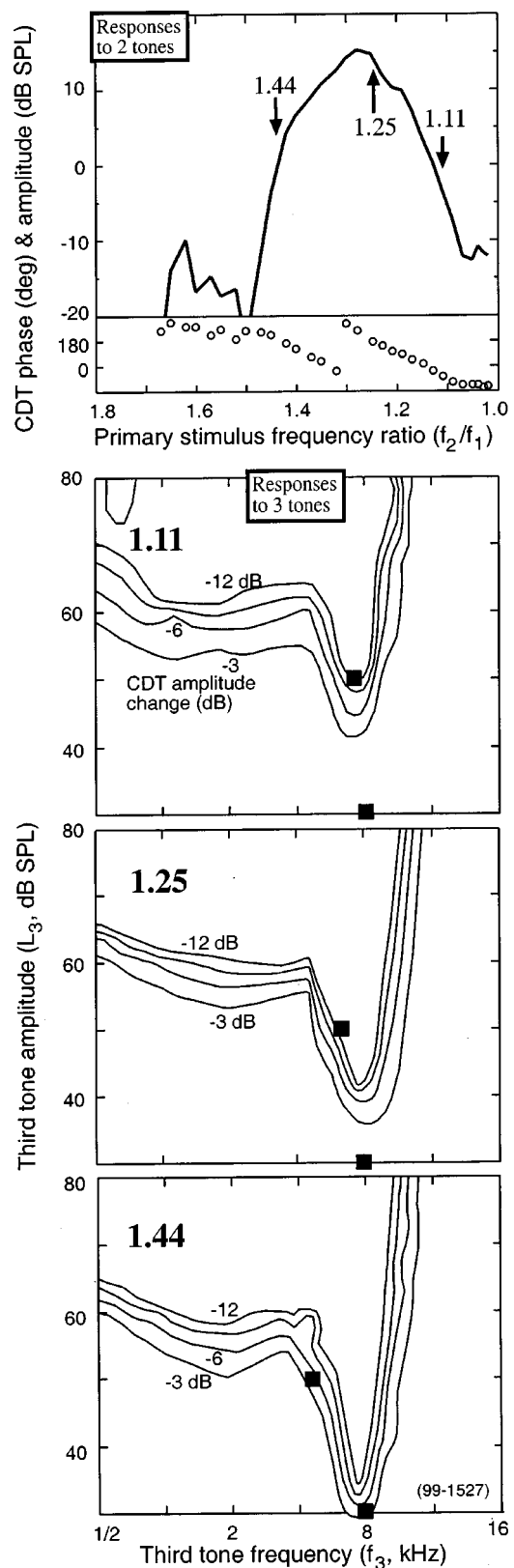


FIG. 13. Typical two- and three-tone frequency response curves for the primary frequency, $f_2 = 8$ kHz, and for stimulus levels $L_1 \times L_2 = 50 \times 30$ dB SPL. Same conventions as previous figures.

The results for $f_2 = 8$ kHz also support the other main trend seen in these observations, that the complex effects generally decreased as the f_2 frequency increased. Though not a focus of this study, two- and three-tone responses were also ob-

tained for f_2 frequencies of 16 and 22 kHz. Other than the very sharp peaks found in the two-tone curves, noted in Fig. 1, there were no obvious complex effects. That is, three-tone maps for higher f_2 frequencies (not shown) were found to be similar in appearance to those at 8 kHz.

III. DISCUSSION

A. Emission frequency responses in other species

The complex behavior detailed here in the gerbil is found to be relatively extreme, and as noted previously, this fact makes it particularly useful to study in this species. Results obtained in the gerbil can be compared to two- and three-tone frequency responses in selected mammalian species, where such responses have been obtained in detail.

1. Human

Two-tone frequency responses in human adults have been extensively investigated by Brown and colleagues (Gaskill and Brown, 1990; Brown *et al.*, 1993a, 1993b). Multiple peaks (as a function of f_2/f_1 ratio) have been noted in the emission amplitudes for frequencies of $2f_1-f_2$ and $3f_1-2f_2$. These appear similar to the multiple peaks seen in the gerbil, but generally were less extreme in humans—i.e., there was typically a smaller peak-to-trough ratio. These responses also appeared to be quite variable. Aspirin consumption, and even nasal infection, were reported to cause shifts in the observed peak frequency responses. Abdala (1996) showed multiple peaks in the individual f_2/f_1 responses of human adults and neonates, particularly for $f_2=1.5$ kHz compared to 6 kHz. The multiple peaks disappeared when the responses were averaged over the group. Such structure in individual responses also appeared to decrease as the stimulus ratio, L_1/L_2 , was increased.

There have been a number of studies of three-tone responses in human adults (Brown and Kemp, 1984; Cianfrone *et al.*, 1994; Kummer *et al.*, 1995; Harris *et al.*, 1992); and comparing human adults and neonates (Abdala and Sininger, 1996; Abdala *et al.*, 1996). Overall, human STCs appear generally to have shapes more similar to neural FTCs than to the complex STC behavior described here in the gerbil. However, complex behavior has been observed in specific cases. Multiple tips have been observed in neonates (Abdala *et al.*, 1996) and in human adults (Kummer *et al.*, 1995), but only for moderate levels of suppression (1–6 dB).

However, no specific search appears to have been made for complex behavior in human STCs similar to that in the gerbil and rabbit. Many of the published studies have included too small a range in parameters to have found such complex behavior if it existed. For example, for humans no published studies have systematically investigated f_3 frequencies more than an octave above f_2 , nor f_2 frequencies below 1 kHz. Some kinds of complex behavior could well have been missed completely if STCs were typically measured by “tracking” the emission decrease, i.e., following a contour experimentally, as is often done. Assuming that tracking begins for suppressor frequencies near the primary frequencies, such procedures are equivalent to following an open contour like the heavy line in Fig. 8. Using such a procedure, the sensitive, complex holes could be missed

completely. This may be the explanation for complex behavior not being reported previously in the gerbil itself (Brown and Kemp, 1984).

Another complexity of the response in human STCs was a modest suppression that occurred with third-tone frequencies near the emission frequency, $2f_1-f_2$ (Kummer *et al.*, 1995; Gaskill and Brown, 1996; Stover *et al.*, 1996; Heitmann *et al.*, 1998). This effect has been interpreted as due to “re-emission” at the $2f_1-f_2$ place. Such behavior in the human was relatively weak, causing at most a 3–4-dB decrease in emission amplitude. Similar behavior was not obvious in gerbil STCs at the –3-dB level. Weaker effects were not investigated, however, and could have been present. In any case, the behavior of the STC for f_3 frequencies near the emission frequency, $2f_1-f_2$, is a completely different matter from the complex behavior detailed here, which occurs for $f_3 > f_2$.

2. Rabbit

Two-tone frequency response curves in this animal typically have been found to have multiple peaks (Whitehead *et al.*, 1992). In the procedure used in that report, the frequency ratio, f_2/f_1 , was varied but the distortion frequency was fixed at either 2, 4, or 5.6 kHz. Multiple peaks were evident in both individual responses and in averaged (group) data. In spite of differences in procedure, the frequency response seems similar to that of the gerbil, except that the complex two-tone behavior appears to extend to higher stimulus frequencies in the rabbit.

Complex results have also been detailed in the rabbit for three-tone responses (Martin *et al.*, 1998, 1999). With primary stimulus frequencies about 3 kHz, anomalous behavior was found near 6 kHz for third-tone levels typically 10–15 dB lower than the thresholds found for third-tone frequencies near the primaries. The complex behavior included increases as well as decreases in primary emission amplitudes, and was reported for emission frequencies $2f_1-f_2$, $2f_2-f_1$, and $3f_1-2f_2$. As in the gerbil, the complex behavior appeared to be more extreme for small primary stimulus frequency ratios ($f_2/f_1 \rightarrow 1$). However, ratios above $f_2/f_1 = 1.25$ were not reported. Neither were primary stimulus frequencies below 3 kHz, so that it is not currently possible to compare gerbil and rabbit responses in detail.

3. Mole rat

Two-tone responses generally had two moderate, broad peaks which changed their relative strengths with stimulus levels (Kössl *et al.*, 1996). For f_2 frequencies from 0.6 to 1.5 kHz, the STCs had broad, or “U”-shaped minima. The lowest threshold typically occurred for f_3 frequencies one-to-two octaves above the primaries. For f_2 frequencies between 2 and 6 kHz, on the other hand, the STCs were complex and broadly similar to those here; that is, there were usually multiple, sharp tips. However, the most sensitive tip was usually the one nearest the f_2 stimulus frequency, unlike the results in gerbil and rabbit.

4. Guinea pig

Two-tone responses appeared to be relatively smooth, generally resulting in a single broad peak (Brown and Gaskill, 1990b). Complex behavior in three-tone responses was found that was “similar” to those in the rabbit, but results were reported only for one set of primary parameters (Martin *et al.*, 1999).

5. Opossum (*Monodelphis domestica*)

Strong multiple “periodic” peaks were found in the two-tone frequency response for this animal for $f_2 = 27.5$ kHz (Faulstich *et al.*, 1996). STCs were reported for f_2 frequencies from 8 to 60 kHz. There was modest evidence for multiple tips, but the thresholds of the tips nearest f_2 remained lowest.

6. Sum

In general, the complex behaviors in two- and three-tone emission responses studied here in the gerbil appear to be relatively common in mammals. However, there is considerable variation among species in the manifestation of such behavior. Rabbits and gerbils appear to show the behavior in an extreme form in both two- and three-tone frequency responses. However, some species show no or only moderately complex behavior in one or the other of these responses. Unfortunately, a detailed comparison is not feasible at the present time because published studies have been done with a wide variation in goals and procedures.

B. Implications for cochlear mechanics

There appear to be two main possibilities for the origin of the complex behavior as detailed in this report. The first is that the complex behavior is actually present in the frequency response of the cochlea itself, and the emissions are faithfully reporting this. This possibility seems unlikely on the evidence. In the gerbil, for example, neural FTCs do not appear to show such effects (Ohlemiller and Echteler, 1990). That is, if the complex behavior were occurring in the cochlear traveling wave, some population of neurons with best frequencies below 8 kHz should have two distinct tips, similar to those seen here for emission STCs. Further, if this possibility were true, suppression of evoked cochlear potential responses in the gerbil should also show two distinct tips, and these do not (e.g., Dolan *et al.*, 1985; Hellstrom and Schmiedt, 1996). More important, the behavior detailed here itself provides an argument that the complex responses primarily occur in the emissions themselves. That is, for the same f_2 frequency, the behavior of the three-tone emission response changes considerably as the primary frequency ratio (f_2/f_1) is changed, or if the primary stimulus levels ($L_1 \times L_2$) are changed. This occurs to the extent that, at a given f_2 frequency, primary stimulus parameters can usually be found so that the complex behavior in the three-tone responses is minimized, or apparently absent. Such findings were, of course, one of the main goals for the present investigation, and are summarized in the next section.

The tentative conclusion is that the complex behavior is present in the process of generating distortion product emis-

sions and/or in the propagation of the emissions out of the cochlea, and *not* in the tuning of the cochlea itself. On the one hand, this finding is welcome, because it suggests that the current basic understanding of cochlear function, the fundamental frequency-place relationships defined by the traveling wave, does not have to be modified. On the other hand, however, it actually makes it more difficult to understand how the complex behavior of the emissions is produced. That is, if the current understanding of the traveling wave behavior is correct, how can two-tone emissions from primary frequencies both near 1 kHz be so strongly affected by very weak stimuli near 3 and 4 kHz? The answer to this question is largely unknown. As noted above, however, it is an important question and it must be answered to establish the correct interpretation of otoacoustic emission measurements in the general situation.

An important clue to the underlying mechanism must be that substantial phase changes often accompany the anomalous suppression or enhancement observed. That is, in many situations, DPOAEs act as if the emission is composed of two or more components at different phase angles (Mills, 1997). If the relative phase between the two components is near 180 degrees, addition of a third tone could cause small changes in this phase which would result in large changes in the summed emission amplitude, similar in form to the changes often observed (e.g., Fig. 5).

C. Measurement of cochlear frequency responses using DPOAEs

The results reported here suggest the following guidelines for attempts to measure cochlear frequency responses using DPOAEs. Overall, the two-tone frequency responses, at least in the gerbil, appear not to be useful for this purpose (Fig. 1). For three-tone responses, the following guidelines appear to be relatively well-established for the Mongolian gerbil, and are offered as preliminary hypotheses to be investigated in other species. That is, in the gerbil, the complex behavior is observed to be less extreme, and the resulting STCs look more like FTCs, when the following conditions are sought:

- (1) The primary stimulus amplitude ratios are unequal, so that $L_1 \gg L_2$. In this report, the two conditions studied included equal level, $L_1 = L_2$, and $L_1/L_2 = 20$ dB. It is possible that even greater inequality in levels could be more useful, i.e., lead to “simpler” STC behavior.
- (2) The primary stimulus frequency ratio (f_2/f_1) is chosen to be as large as possible and still obtain measurable emissions. The result of the first two conditions is that one seeks a primary stimulus where the stimulus at f_1 is much lower in frequency but much higher in intensity than the f_2 primary. That is, one seeks primary stimulus parameters which have maximum *separation* between the two primary tones, in both frequency and intensity, while still resulting in measurable emission amplitudes.
- (3) Attention is directed preferentially to the contour lines representing strong suppression, i.e., by 12 dB or more. Even in this case, attention is best restricted to “open contours,” i.e., the contours representing the upper

boundary of the tuning curve (shown as a heavy line in Fig. 8, middle panel). This suggestion comes from the fact that the complex effects are usually relatively weak, and even when there is strong complex suppression, it is found in a "hole". That is, the contours representing the most extreme complex suppression effects are usually closed. Examples are seen in both maps in Fig. 8.

- (4) Even at low primary frequencies, the cubic distortion tone (CDT, $2f_1-f_2$) emission remains a better choice for obtaining emission STCs than does the other cubic emission at frequency $2f_2-f_1$.

Unfortunately, these requirements tend to result in relatively lower primary emission amplitudes, pushing the stimulus parameters to the boundaries of the parameter space which results in measurable emissions, and making measurements correspondingly difficult. The conclusion of this study is that following the above parameter choices generally results in emission suppression tuning curves (STCs) which appear more similar in form to neural frequency-tuning curves (FTCs), at least in the Mongolian gerbil. This does not, of course, mean that such STCs actually provide useful information on cochlear tuning, or that the correct interpretation of such emission STCs has the same meaning as a neural FTC. To establish such relationships requires, among other things, that direct comparisons be made between the results of emission STC measurements and other, more established measures of cochlear frequency tuning.

ACKNOWLEDGMENTS

Thanks to B. Warren for assistance with software and equipment, and to two anonymous reviewers for constructive comments on an earlier version of this manuscript. This research was supported by Grant No. RO1 DC 04077-01 from the National Institute for Deafness and Other Communication Disorders, National Institutes of Health.

Abdala, C. (1996). "Distortion product otoacoustic emission ($2f_1-f_2$) as a function of f_2/f_1 frequency ratio and primary tone level separation in human adults and neonates," J. Acoust. Soc. Am. **100**, 3726–3740.

Abdala, C. (1998). "A developmental study of distortion product otoacoustic emission ($2f_1-f_2$) suppression in humans," Hear. Res. **121**, 125–138.

Abdala, C., and Sininger, Y. S. (1996). "The development of cochlear frequency resolution in the human auditory system," Ear Hear. **17**, 374–385.

Abdala, C., Sininger, Y. S., Ekelid, M., and Zeng, F.-G. (1996). "Distortion product otoacoustic emission suppression tuning curves in human adults and neonates," Hear. Res. **98**, 38–53.

Brown, A. M., and Gaskill, S. A. (1990a). "Can basilar membrane tuning be inferred from distortion measurement?," in *The Mechanics and Biophysics of Hearing: Proceedings of a Conference held at the University of Wisconsin, Madison, WI, 25–29 June 1990*, edited by P. Dallos, C. D. Geisler, J. W. Matthews, M. A. Ruggero, and C. R. Steele (Springer, New York), pp. 164–169.

Brown, A. M., and Gaskill, S. A. (1990b). "Measurement of acoustic distortion reveals underlying similarities between human and rodent mechanical responses," J. Acoust. Soc. Am. **88**, 840–849.

Brown, A. M., Gaskill, S. A., Carlyon, R. P., and Williams, D. M. (1993a). "Acoustic distortion as a measure of frequency selectivity: Relation to psychophysical equivalent rectangular bandwidth," J. Acoust. Soc. Am. **93**, 3291–3297.

Brown, A. M., Gaskill, S. A., and Williams, D. M. (1992). "Mechanical filtering of sound in the inner ear," Proc. R. Soc. London, Ser. B **250**, 29–34.

Brown, A. M., and Kemp, D. T. (1984). "Suppressibility of the $2f_1-f_2$ stimulated acoustic emissions in gerbil and man," Hear. Res. **13**, 29–37.

Brown, A. M., Williams, D. M., and Gaskill, S. A. (1993b). "The effect of aspirin on cochlear mechanical tuning," J. Acoust. Soc. Am. **93**, 3298–3307.

Cianfrone, G., Altissimi, G., Cervellini, M., Musacchio, A., and Turchetta, R. (1994). "Suppression tuning characteristics of $2f_1-f_2$ distortion product otoacoustic emissions," Br. J. Audiol. **28**, 205–212.

Dallos, P., and Cheatham, M. A. (1976). "Compound action potential (AP) tuning curves," J. Acoust. Soc. Am. **59**, 591–597.

Dallos, P., Harris, D., Ozdamar, O., and Ryan, A. (1978). "Behavioral, compound action potential, and single unit thresholds: Relationship in normal and abnormal ears," J. Acoust. Soc. Am. **64**, 151–157.

Dolan, T. G., Mills, J. H., and Schmiedt, R. A. (1985). "A comparison of brainstem, whole nerve AP and single-fiber tuning curves in the gerbil: Normative data," Hear. Res. **17**, 259–266.

Eggermont, J. J. (1976). "Analysis of compound action potential responses to tone bursts in the human and guinea pig cochlea," J. Acoust. Soc. Am. **60**, 1132–1139.

Faulstich, M., Kössl, M., and Reimer, K. (1996). "Analysis of non-linear cochlear mechanics in the marsupial *Monodelphis domestica*: Ancestral and modern mammalian features," Hear. Res. **94**, 47–53.

Frank, G., and Kössl, M. (1995). "The shape of $2f_1-f_2$ suppression tuning curves reflects basilar membrane specializations in the mustached bat, *Pteronotus parnellii*," Hear. Res. **83**, 151–160.

Gaskill, S. A., and Brown, A. M. (1990). "The behavior of the acoustic distortion product, $2f_1-f_2$, from the human ear and its relation to auditory sensitivity," J. Acoust. Soc. Am. **88**, 821–839.

Gaskill, S. A., and Brown, A. M. (1996). "Suppression of human acoustic distortion product: Dual origin of $2f_1-f_2$," J. Acoust. Soc. Am. **100**, 3268–3274.

Harris, F. P., Probst, R., and Xu, L. (1992). "Suppression of the $2f_1-f_2$ otoacoustic emission in humans," Hear. Res. **64**, 133–141.

Heitmann, J., Waldmann, B., Schnitzler, H.-U., Plinkert, P. K., and Zenner, H.-P. (1998). "Suppression of distortion product otoacoustic emissions (DPOAE) near $2f_1-f_2$ removes DP-gram fine structure—Evidence for a secondary generator," J. Acoust. Soc. Am. **103**, 1527–1531.

Hellstrom, L. I., and Schmiedt, R. A. (1996). "Measures of tuning and suppression in single-fiber and whole-nerve responses in young and quiet-aged gerbils," J. Acoust. Soc. Am. **100**, 3275–3285.

Kemp, D. T. (1978). "Stimulated acoustic emissions from within the human auditory system," J. Acoust. Soc. Am. **64**, 1386–1391.

Köppl, C., and Manley, G. A. (1993). "Distortion-product otoacoustic emissions in the bobtail lizard. Suppression tuning characteristics," J. Acoust. Soc. Am. **93**, 2834–2844.

Kössl, M. (1993). "Evidence for a mechanical filter in the cochlea of the 'constant frequency' bats, *Rhinolophus rouxi* and *Pteronotus parnellii*," Hear. Res. **72**, 73–80.

Kössl, M., Frank, G., Burda, H., and Müller, M. (1996). "Acoustic distortion products from the cochlea of the blind African model rat, *Cryptomys spece*," J. Comp. Physiol. A **178**, 427–434.

Kössl, M., and Vater, M. (1996). "Further studies on the mechanics of the cochlear partition in the mustached bat. II. A second cochlear frequency map derived from acoustic distortion products," Hear. Res. **94**, 78–86.

Kummer, P., Janssen, T., and Arnold, W. (1995). "Suppression tuning characteristics of the $2f_1-f_2$ distortion-product otoacoustic emission in humans," J. Acoust. Soc. Am. **98**, 197–210.

Liberman, M. C., and Kiang, N. Y. S. (1978). "Acoustic trauma in cats: Cochlear pathology and auditory-nerve activity," Acta Oto-Laryngol. Suppl. **338**, 1–63.

Martin, G. K., Jassir, D., Stagner, B. B., and Lonsbury-Martin, B. L. (1998). "Effects of loop diuretics on the suppression tuning of distortion-product otoacoustic emissions in rabbits," J. Acoust. Soc. Am. **104**, 972–983.

Martin, G. K., Lonsbury-Martin, B. L., Probst, R., Scheinin, S. A., and Coats, A. C. (1987). "Acoustic distortion products in rabbit ear canal. II. Sites of origin revealed by suppression contours and pure-tone exposures," Hear. Res. **28**, 191–208.

Martin, G. K., Stagner, B. B., Jassir, D., Telischi, F. F., and Lonsbury-Martin, B. L. (1999). "Suppression and enhancement of distortion-product otoacoustic emissions by interference tones above f_2 . I. Basic findings in rabbits," Hear. Res. **136**, 105–123.

- Mills, D. M. (1997). "Interpretation of distortion product otoacoustic emission measurements. I. Two stimulus tones," J. Acoust. Soc. Am. **102**, 413–429.
- Mills, D. M. (1998). "Interpretation of distortion product otoacoustic emission measurements. II. Estimating tuning characteristics using three stimulus tones," J. Acoust. Soc. Am. **103**, 507–523.
- Mills, D. M., and Rubel, E. W. (1996). "Development of the cochlear amplifier," J. Acoust. Soc. Am. **100**, 428–441.
- Mills, D. M., and Rubel, E. W. (1997). "Development of distortion product emissions in the gerbil: Filter response and signal delay," J. Acoust. Soc. Am. **101**, 395–411.
- Norton, S. J., Bargones, J. Y., and Rubel, E. W. (1991). "Development of otoacoustic emissions in gerbil: Evidence for micromechanical changes underlying development of the place code," Hear. Res. **51**, 73–92.
- Ohlemiller, K. K., and Echteler, S. M. (1990). "Functional correlates of characteristic frequency in single cochlear nerve fibers of the Mongolian gerbil," J. Comp. Physiol. A **167**, 329–338.
- Plassmann, W., and Kadel, M. (1991). "Low-frequency sensitivity in a gerbilline rodent, *Pachyuromys duprasi*," Brain Behav. Evol. **38**, 115–126.
- Ryan, A. (1976). "Hearing sensitivity of the Mongolian gerbil, *Meriones unguiculatus*," J. Acoust. Soc. Am. **59**, 1222–1226.
- Schmiedt, R. A. (1989). "Spontaneous rates, thresholds and tuning of auditory-nerve fibers in the gerbil: Comparisons to cat data," Hear. Res. **42**, 23–36.
- Schmiedt, R. A., and Zwislocki, J. J. (1978). "Low-frequency neural and cochlear-microphonic tuning curves in the gerbil," J. Acoust. Soc. Am. **64**, 502–507.
- Stover, L. J., Neely, S. T., and Gorga, M. P. (1996). "Latency and multiple sources of distortion product otoacoustic emissions," J. Acoust. Soc. Am. **99**, 1016–1024.
- Whitehead, M. L., Lonsbury-Martin, B. L., and Martin, G. K. (1992). "Evidence for two discrete sources of $2f_1-f_2$ distortion-product otoacoustic emission in rabbit. Differential dependence on stimulus parameters," J. Acoust. Soc. Am. **91**, 1587–1607.

Effects of acoustic trauma on acoustic enhancement of electrically evoked otoacoustic emissions

Hideko Heidi Nakajima^{a)}

Hearing Research Center and Department of Biomedical Engineering, Boston University,
44 Cummington Street, Boston, Massachusetts 02215

Allyn E. Hubbard

Hearing Research Center and Department of Biomedical Engineering, Department of Electrical
and Computer Engineering, and Department of Otolaryngology, Boston University, 44 Cummington Street,
Boston, Massachusetts 02215

David C. Mountain

Hearing Research Center and Department of Biomedical Engineering and Department of Otolaryngology,
Boston University, 44 Cummington Street, Boston, Massachusetts 02215

(Received 14 July 1999; accepted for publication 26 January 2000)

Moderate acoustic trauma results in decreased cochlear sensitivity and frequency selectivity. This decrease is believed to be caused by damage to the cochlear amplifier that is associated with outer hair cells (OHCs) and their nonlinear electromechanical characteristics. A consequence of OHC nonlinearity is the acoustic enhancement effect, in which low-frequency electrically evoked otoacoustic emissions are enhanced by a simultaneous tone. The present study found that acoustic trauma reduced the acoustic enhancement effect and this reduction is correlated with the N_1 threshold at the electrode site. This result is consistent with the theory that trauma affects the mechanoelectric transduction process, thus affecting cochlear mechanical nonlinearity. Acoustic trauma also reduced the cochlear microphonic in a way that suggests that the number of functioning tension-gated channels and the stiffness of the gating springs were decreased. In some cases, the electromechanical transduction process was also found to be affected by acoustic trauma. © 2000 Acoustical Society of America. [S0001-4966(00)04904-3]

PACS numbers: 43.64.-q, 43.64.Jb, 43.64.Bt, 43.64.Wn [BLM]

INTRODUCTION

Moderate acoustic trauma has been shown to decrease cochlear sensitivity and frequency selectivity. This decrease is believed to be caused by damage to the cochlear amplifier (Davis, 1983; Kemp, 1978; Neely and Kim, 1983). Because the cochlear amplifier is associated with outer hair cells (OHCs), it appears that the OHCs are adversely affected by overstimulation (Cody and Russell, 1985, 1988). Furthermore, direct evidence of OHC structural damage due to acoustic trauma has been observed (Engström *et al.*, 1970; Hunter-Duvar, 1977; Liberman and Kiang, 1978).

Possible OHC mechanisms that can be affected by acoustic trauma are the forward (mechanoelectric) transduction process and the reverse (electromechanical) transduction process. OHC forward transduction can be studied using cochlear microphonic (CM) measurements, and low-frequency CM is believed to be proportional to the OHC forward transduction current (Dallos *et al.*, 1972; Patuzzi *et al.*, 1989a). Patuzzi *et al.* (1989b) used the CM to demonstrate that acoustic trauma altered the forward transduction process. OHC electromotility can be studied using electrically evoked otoacoustic emissions (EEOEs), which are believed to be proportional to reticular lamina motion (Xue *et al.*, 1996). Nakajima *et al.* (1996) studied the effect of acoustic trauma on EEOEs and found that acoustic trauma compromises the

forward transduction process, while reverse transduction seems to be relatively unaffected.

In the past, acoustic trauma has been shown to affect the compressive nonlinearity of the basilar-membrane motion, resulting in a more linear response (Sellick *et al.*, 1982; Patuzzi *et al.*, 1984). If the mechanical nonlinearity is largely due to the nonlinear forward transduction process (Mountain *et al.*, 1983), and if acoustic trauma damages forward transduction, then we would expect a decrease in cochlear mechanical nonlinearity.

The nonlinearity in the forward transduction process may also underlie the nonlinear property of EEOEs. The EEOE magnitude at the frequency of the electrical stimulus increases when a relatively loud tone is presented simultaneously (Mountain and Hubbard, 1989; Murata *et al.*, 1991; Xue *et al.*, 1993; Kirk and Yates, 1996). We refer to this phenomenon as the acoustic enhancement effect. If acoustic trauma damages forward transduction, then we would expect a decrease in this nonlinear interaction between sound and electrical stimulation. This prediction has been confirmed by the present study, in which we found that acoustic trauma reduced the acoustic enhancement effect.

I. METHODS

A. Surgical preparation and experimental setup

Adult Mongolian gerbils, *Meriones unguiculatus*, were anesthetized with an initial intraperitoneal injection of so-

^{a)}Electronic mail: hhn@bu.edu

dium pentobarbital (60 mg/kg) and maintained with 12 mg/kg as needed according to an institutionally approved protocol. The surgical procedure and the experimental setup, including acoustic and electrical stimulation and the recording of EEOEs, CM, summating potential (SP), and endocochlear potential (EP), were similar to those described in Nakajima *et al.* (1994, 1996). An ear bar with probe-tube microphone B&K type 4166 and earphone (Beyer DT-48) was attached to the bony auditory canal. A glass microelectrode filled with 1.5 M KCl with a 2- to 4- μm tip and 1- to 2-M Ω impedance was inserted into the scala media of the second cochlear turn to stimulate the cochlea with electrical current and to measure CM, SP, and EP. The return electrode for current stimulation and the reference electrode for electrical measurements were placed near the neck muscles. A silver electrode was placed near the round window and a return electrode was inserted near the neck muscles to measure the threshold curves for the tone-pip evoked N_1 waveform of the compound action potential.

The steady-state scala media CM responses were measured with stimulus tone intensities which varied from 20 to 90 dB SPL at 400 Hz. The waveforms were stored digitally, and the magnitude of the fundamental, second, and third harmonics were extracted by performing a fast Fourier transform.

SP was measured using frequencies of 800 Hz to 9 kHz and stimulus levels of 40 to 70 dB SPL. The stimulus frequency that evoked the most negative SP was used as an estimate of the characteristic frequency (CF) of the electrode location (Nakajima *et al.*, 1994).

N_1 responses to tone pips were measured for stimuli with 10- to 80-dB SPL intensities and 200-Hz to 20-kHz frequencies. The N_1 threshold curves for a threshold of 10 μV were then calculated for frequencies between 200 Hz and 20 kHz. Recordings were taken for various stimulus intensities and frequencies, and the thresholds were computed by interpolation.

Otoacoustic emissions were evoked by injecting an alternating current (approximately 7 $\mu\text{A}_{\text{peak}}$) into the scala media and the resulting signals from the calibrated probe-tube microphone were digitized. The emission, at the frequency of the stimulus current, was extracted by applying a Hanning window to the averaged microphone signal and performing a fast Fourier transform. Care was taken to ensure isolation between the electrical stimulation and microphone recording signals. A control measurement for EEOE was made by passing current through a microelectrode with its tip placed on the surface of the bulla to ensure that the microphone recording was at the acoustic noise level (acoustic noise level measured without electrical stimulation).

In this study, we focused our efforts on responses below CF. Because EEOE and CM are measurements of distributed responses, interpretation is simplified if the wavelength of the associated traveling wave is large compared to the electrical space constant. For this condition, the responses from the OHCs contributing to the measurements would be in phase. In contrast, interpretation of responses near CF (~ 2.5 kHz) would be more complex due to phase cancellation.

B. Transient effects of traumatizing tones

Overstimulating tones (100–110 dB SPL for 3–10 min) were presented sequentially from high to low frequencies (i.e., 1–2 presentations each of 4.5 kHz, 1.8 kHz, then 800 Hz). For each traumatizing tone exposure sequence, we measured N_1 , EEOE (without acoustic stimulation), and the acoustically enhanced electrically evoked otoacoustic emission (AEEOE). All three measurements were taken just before the traumatizing exposure (3 points) and just after (7 points). The EEOE and AEEOE were measured at 800 Hz. The simultaneous acoustic stimulus used to enhance the emission for the AEEOE measurement was 1.2 kHz and 70 dB SPL. Acoustic enhancement of emissions was defined as the ratio of the AEEOE to the EEOE. The N_1 measurements were made to assess the damage caused by the traumatizing tone. The frequency used (half an octave above the traumatizing tone) was chosen to correspond to the CF of the expected location for maximum damage. These measurements enabled us to ensure that the fast transient effects of the traumatizing tones were over by the time we made steady-state measurements.

Between the measurements of the transient effects of traumatizing tones, we made steady-state measurements. These steady-state measurements included EEOEs with simultaneous acoustic stimuli and CM. We also obtained N_1 thresholds between the presentations of traumatizing tones.

C. Maximum acoustic enhancement of EEOEs

To quantify the acoustic enhancement of the EEOEs, we measured the EEOEs (at frequency $f_E = 800$ Hz) while simultaneously applying a tone at a frequency ($f_A = 1200$ Hz). These frequencies for the electrical and acoustic stimuli were used because they are known to yield good enhancement in turn 2 of the gerbil. Acoustic enhancement is graded with the level of acoustic stimulus and typically saturates for high sound-pressure levels. For our purposes, we defined the “maximum acoustic enhancement” for the steady-state recordings as the difference between the maximum (usually saturated) emission and the emission at low sound stimulus level. Specifically, as shown in Fig. 1, we calculated the maximum acoustic enhancement as the difference between the saturated enhanced emission and the unenhanced emission. The saturated enhanced emission is defined as the average response from 70- to 80-dB SPL stimuli, and the unenhanced emission is defined as the average response from 10- to 30-dB SPL stimuli.

In some animals, slight variations (maximum of 1.7%) in the current stimulus level occurred. To correct for the variation, we assumed a linear relationship between EEOE magnitude and current stimulus for the frequency and small range of current used. This linear relationship was confirmed by an earlier study (Nakajima *et al.*, 1998) and in some of the cochleas used in this study. Other investigators, such as Murata *et al.* (1991), Ren and Nuttall (1995), and Ren *et al.* (1996), also observed a linear relationship.

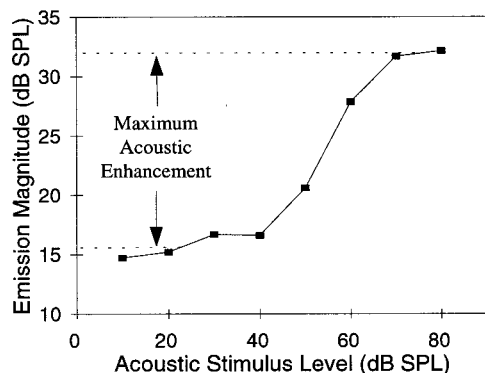


FIG. 1. The emissions were evoked by electrical current at $f_E = 800$ Hz while simultaneous acoustic stimuli at various levels were presented at $f_A = 1.2$ kHz. The abscissa is the acoustic stimulus level and the ordinate is the emission magnitude at the electrical frequency. "Maximum acoustic enhancement" is defined as the average saturated enhanced emission divided by the average nonenhanced emission, in dB.

D. Assessment of cochlear condition

The N_1 threshold curve and the EP were used to assess the cochlear condition throughout the experiment. Although the EP was monitored throughout the experiment, the most reliable measurements are at the very beginning when the electrode goes into the scala media, and at the end when the electrode is withdrawn, to measure the reference potential at the surface of the cochlea. This is because the potential can drift due to electrode polarization. To be considered a healthy cochlea, the N_1 threshold shift due to electrode insertion into the scala media had to be less than 10 dB. The EP had to remain high, except where noted, throughout the experiment (above 70 mV initially and above 50 mV at the end).

II. RESULTS

A. Transient changes in emissions

Figure 2 shows the transient changes in emission measurements for five traumatizing-tone sequences. The timing of the traumatizing tones is indicated by horizontal bars near the bottom of the figure and the frequencies of traumatizing tones are noted. The first row of data shows the AEEOE

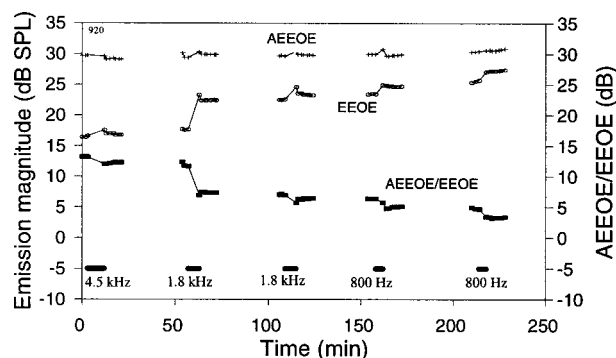


FIG. 2. Time-series emission measurements for AEEOE, EEOE, and acoustic enhancement (AEEOE/EEOE). Horizontal bars near the bottom indicate the timing and frequencies of traumatizing tones. Emissions were evoked by $7 \mu A_{\text{peak}}$ electrical current at $f_E = 800$ Hz, and the acoustic stimulus for AEEOE was 70 dB SPL at $f_A = 1.2$ kHz.

(with a simultaneous tone of 70 dB SPL), and the second row of data shows the EEOEs (with no acoustic stimuli). The third row of data is the acoustic enhancement effect (AEEOE to EEOE ratio in dB). Three points of control data are collected prior to overstimulating tone delivery and then seven points after. Just after the traumatizing-tone presentation, the emissions show a transient change and then level off at a value that changes slowly. Generally, the EEOE magnitude increases after acoustic trauma, as was shown in an earlier study (Nakajima *et al.*, 1996). The enhancement effect (AEEOE to EEOE ratio) was found to decrease after traumatizing-tone exposure.

B. Effects on steady-state EEOE with simultaneous tone stimuli

As shown in the transient traumatizing-tone sequence data, the enhancement effect decreases after traumatizing-tone exposure for steady-state measurements. However, the manner in which EEOE magnitudes (with simultaneous acoustic stimuli) changed after trauma varied with the intensity of simultaneous tone and varied across animals. We present three examples, each representing a type of traumatizing effect on EEOEs with simultaneous acoustic stimuli.

The upper left panel of Fig. 3 shows an example of the steady-state EEOE response magnitude as a function of simultaneous tone intensity, before and after several traumatizing-tone presentations. The initial input/output curve is depicted with solid square markers. The ordering of the recordings can be seen in the legend. When low stimulus tone intensities (10 to 30 dB SPL) were used, the emissions were not enhanced, and after each subsequent traumatizing-tone presentation, emissions increased monotonically. The sound level required to enhance the emission became higher as more acoustic trauma was produced. The maximally enhanced emission remained basically the same after acoustic trauma. This type of acoustic trauma effect on EEOEs with simultaneous acoustic stimuli will be referred to as *type I*.

A different EEOE response following acoustic trauma, referred to as *type II*, is shown in the upper panel in the middle column of Fig. 3. In general, for low intensities of the simultaneous tone, the emissions increased, as for the type I case. However, the last traumatizing tone actually decreased the emission slightly. At high enhancing tone intensities, the emission decreased after traumatizing tones. This decrease almost eliminated the enhancement effect for the stimulus tone intensities investigated. As seen in the curves before and after the first traumatizing tone, acoustic trauma shifted the onset of enhancement to higher stimulus tone levels.

The third example of acoustic trauma's effect on the acoustically enhanced EEOE (*type III* case) is shown in the upper right panel in Fig. 3. In this case, the curves demonstrated pronounced downward shifts with successive traumatizing tone deliveries.

A schematic diagram is presented to clarify the categorization of type I, II, III cases and the different stimuli involved (Fig. 4). In these experiments, both electrical (at f_E) and acoustic (at f_A) stimuli are used simultaneously to obtain EEOE (at f_E) that can be acoustically enhanced. The resulting EEOE is shown on the bottom left labeled "no trauma."

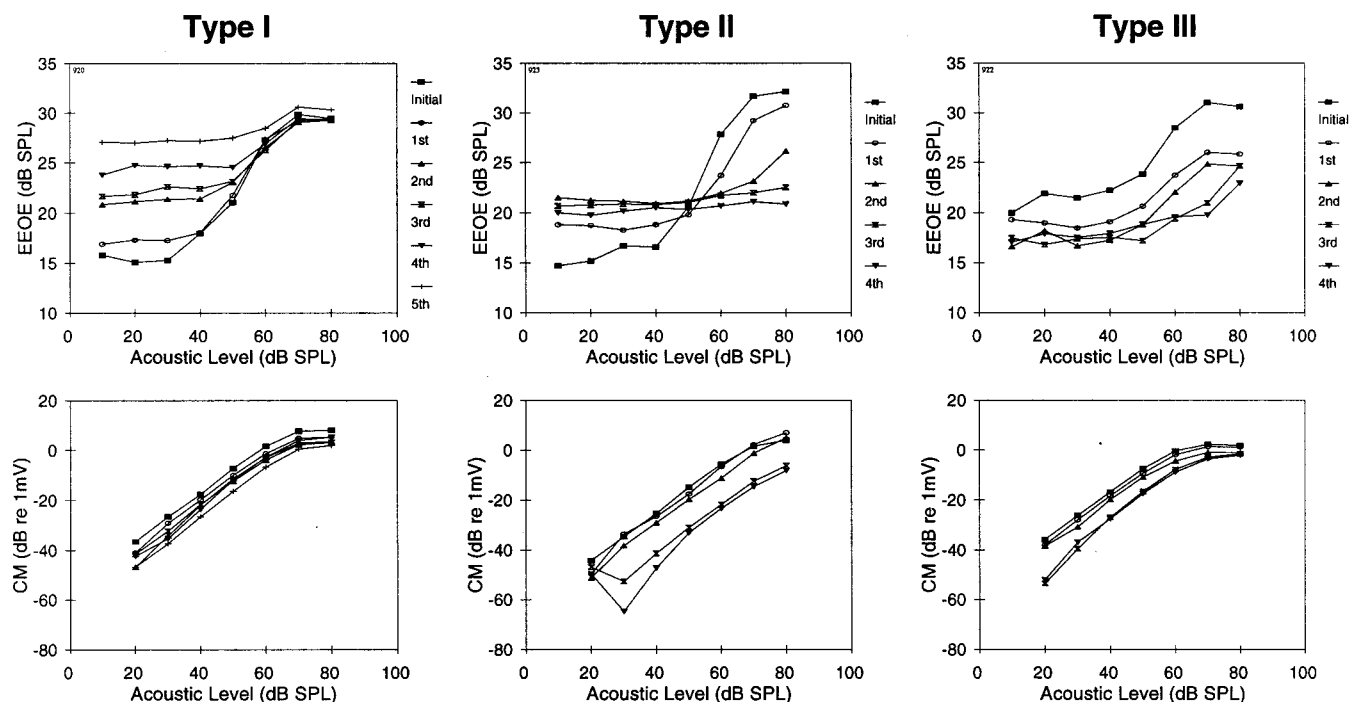


FIG. 3. Top row: steady-state EEOE response as a function of simultaneous acoustic stimulus intensity. The initial recording is plotted with solid squares, and the legend indicates the sequence of recordings taken after each acoustic overstimulation presentation. The electrical current was at $f_E = 800$ Hz, and the acoustic stimuli were at $f_A = 1.2$ kHz. Bottom row: corresponding steady-state CM response as a function of acoustic stimulus intensity ($f_A = 400$ Hz). For type I (leftmost column), the electrical current was $7 \mu A_{\text{peak}}$ during EEOE measurements and the EP was 84 mV at the beginning of the experiment and 66 mV at the end. For type II (middle column), the electrical current was $8 \mu A_{\text{peak}}$ during EEOE measurements and the EP was 80 mV at the beginning and 87 mV at the end of the experiment. For type III (rightmost column), the electrical current was $6 \mu A_{\text{peak}}$ during EEOE measurements and the EP was 49 mV initially and 37 mV at the end.

Acoustic trauma affects the acoustic enhancement of EEOE in three possible ways, resulting in type I, II, and III cases. The initial emission is plotted with a thin line, and the emission after acoustic trauma is plotted with a thick line.

The effect of acoustic trauma on the acoustically enhanced EEOE appears to depend on the condition of the cochlea. Although initial N_1 thresholds were usually good, the EP could vary. Generally, cochleas with type I and II

responses had EPs above 50 mV throughout the experiment, while cochleas with type II responses had low EPs. Out of 18 experiments, 5 exhibited responses of the type I case, 4 experiments were of the type II case, and all 9 had high EPs. Six experiments were the type III case, and all but one animal out of the six had low EP. Other experiments (not shown) exhibited a combination of responses due to acoustic trauma. Three experiments initially produced type II results,

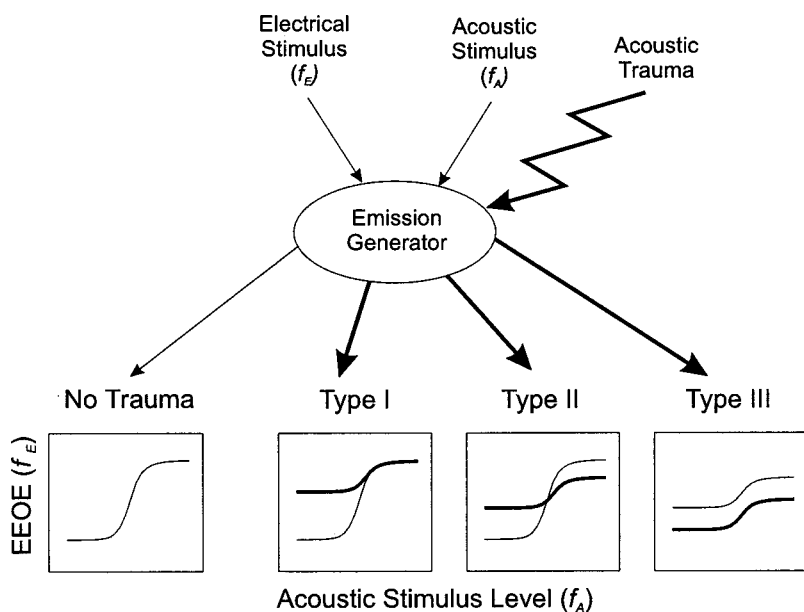


FIG. 4. Effect of acoustic trauma on acoustic enhancement of EEOEs. Enhancement of EEOE is produced by the simultaneous stimulation of acoustic stimulus at f_A with the electrical stimulus at f_E . The EEOE before acoustic trauma is shown in the bottom-left curve. With acoustic trauma, the EEOE changes in three possible ways: type I, II, and III. The after-trauma curves are plotted with thick lines and the initial EEOEs are plotted with thin lines.

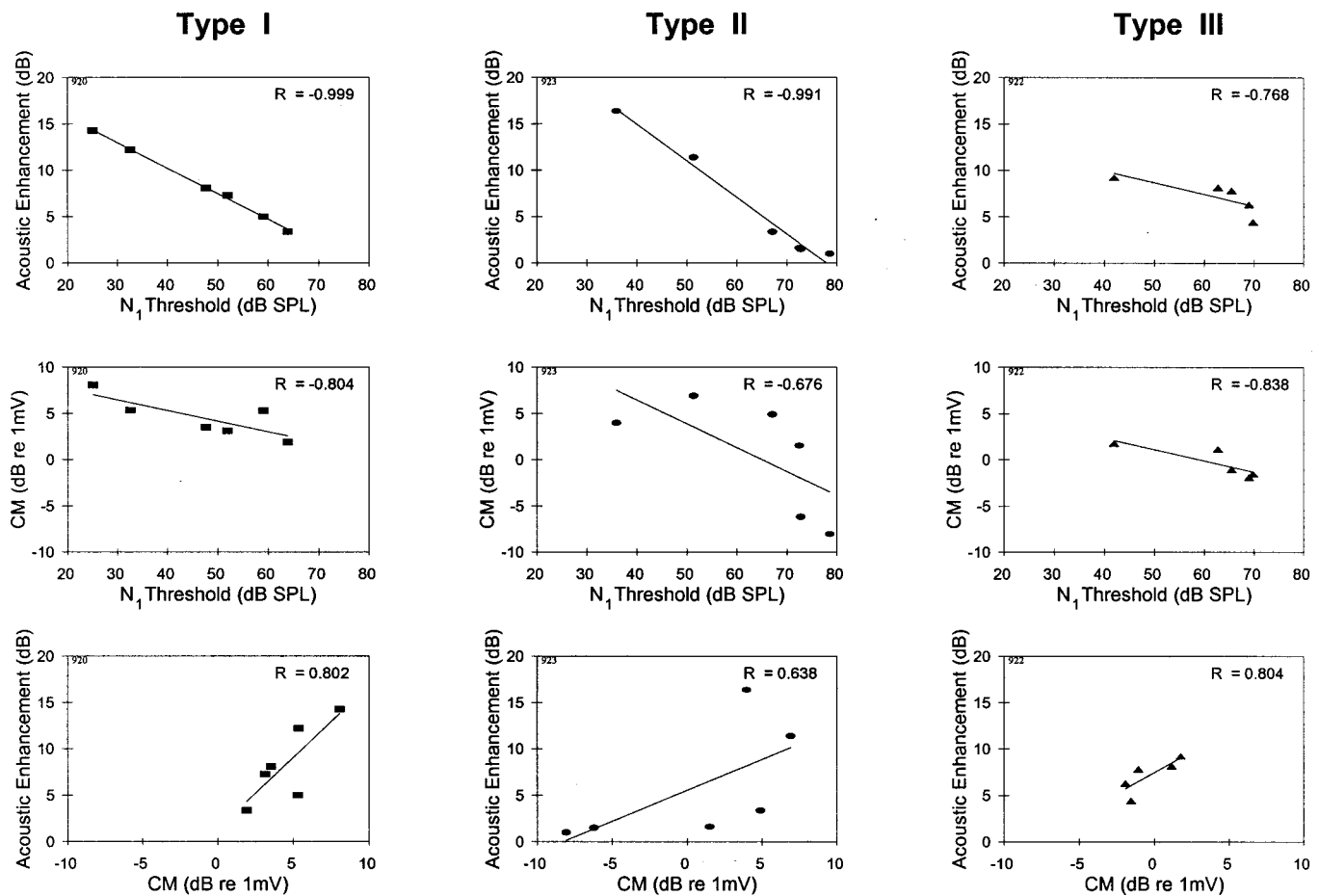


FIG. 5. Data in the type I, II, and III columns are from the same animals as in Fig. 3. Top row plots maximum acoustic enhancement of emission versus N_1 threshold. Middle row plots CM versus N_1 threshold. Bottom row plots maximum acoustic enhancement versus CM. Type I and II were cochleas with high EP, while III was a cochlea with low EP. The correlation coefficient, R , is noted on the right-top corner of each plot.

and then later type III. Two out of these three cochleas had low EPs by the end of the experiment.

C. Effects on CM

The effect of acoustic trauma on CM was similar to that reported in our previous study (Nakajima *et al.*, 1996). The CM input/output curves are plotted under the emission data in Fig. 3. The CM magnitude at 400 Hz initially saturated around 70 dB SPL (solid square markers). Generally, acoustic trauma decreased CM magnitude and the largest decreases occurred at lower acoustic stimulus levels.

D. Correlation between enhancement, CM, and N_1 threshold

The changes in maximum enhancement and N_1 thresholds due to acoustic trauma exhibited a strong correlation. The top panels in Fig. 5 are plots of the maximum acoustic enhancement ($f_E = 800$ Hz) as a function of the N_1 threshold at 2.5 kHz (the CF of the electrode location) for the three animals presented earlier in Fig. 3. A regression line is plotted and the correlation coefficient (R) noted on the top-right corner. Type I and type II (first two columns), which were cochleas with high EP, show strong correlation between en-

hancement and N_1 threshold. Type III (third column), cochlea with low EP, shows a less strong correlation between enhancement and N_1 .

The second row of panels in Fig. 5 are plots of the correlations between the 400-Hz CM (with 80-dB SPL stimuli) and the N_1 thresholds at the CF of the electrode location. The effect of acoustic trauma on N_1 thresholds is generally much larger than on CM. This is similar to results reported by Patuzzi *et al.* (1989b). Comparing the effects of acoustic trauma on acoustic enhancement (top row) and CM (middle row) for cochleas with high EP (type I and II), correlation between enhancement and N_1 threshold is better than CM and N_1 threshold, especially for the type II case. For the cochlea with low EP (type III), the initial enhancement and CM magnitude was low, and the first traumatizing-tone presentation increased the N_1 threshold to a large value. Subsequent traumatizing-tone presentations did not change the amount of enhancement, CM, or N_1 threshold much further. Thus, the relatively high R value may be misleading in the type III case.

The bottom panels of Fig. 5 are plots of the relationship between acoustic enhancement and CM. For the cochleas with high EP type I has stronger correlation than type II. For the low-EP cochlea (type III), the correlation is relatively good, but both the enhancement and CM were initially low and acoustic trauma did not have a large effect.

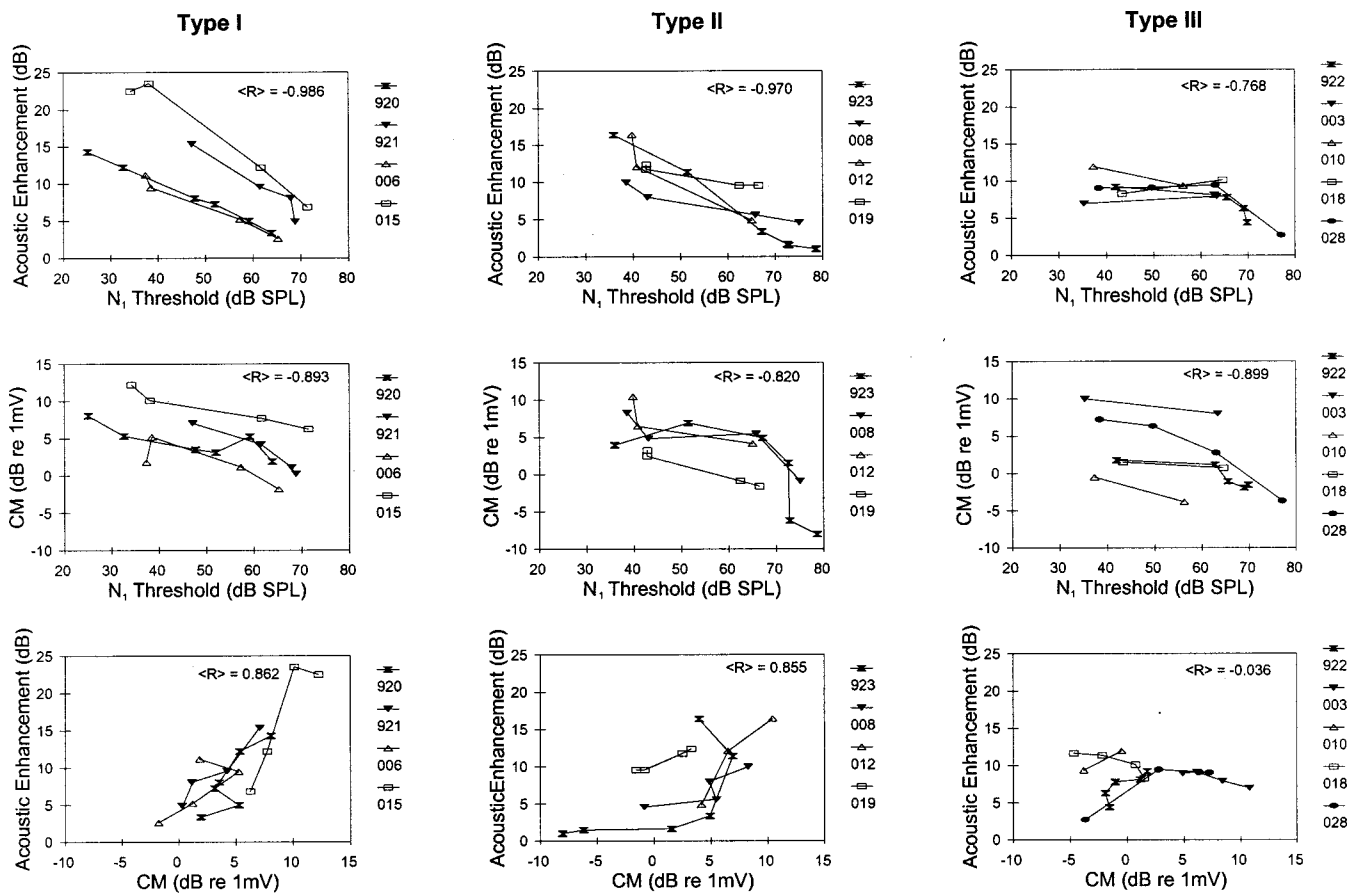


FIG. 6. These plots are similar to Fig. 5, but include data from several animals for each of the traumatized-emission types, organized by column. Top-row panels are plots of maximum acoustic enhancement of emission versus N_1 threshold. Middle-row panels are plots of CM versus N_1 threshold. Bottom-row panels are plots of maximum acoustic enhancement versus CM. The average correlation coefficient, $\langle R \rangle$, is indicated on the right-top corner of each plot, and the animal numbers are in the legends.

Figure 6 plots acoustic enhancement versus N_1 threshold, CM versus N_1 threshold, and acoustic enhancement versus CM for several animals that were acoustically traumatized. The average correlation coefficient ($\langle R \rangle$) is on the upper-right corner of each plot. Looking at acoustic enhancement versus N_1 threshold plots (first row), the type I and II's acoustic enhancement generally decrease with each acoustic trauma presentation and correlate well with the incremental increase in N_1 threshold. The CM versus N_1 plots (second row) show less correlation than the enhancement versus N_1 plots for cochleas with type I enhancement. The correlation of the CM to N_1 is even weaker for cochleas with type II enhancement. In the acoustic enhancement versus CM plots (third row), type I and II show a general trend where the decrease in acoustic enhancement correlates with the CM decrease after each acoustic trauma presentation. For both the CM versus N_1 threshold and enhancement versus CM, the type I has higher correlation than type II. The type III, however, begins with a low acoustic enhancement; thus, acoustic trauma does not always decrease the enhancement further. Also in type III, the N_1 threshold sometimes increases after acoustic trauma above the maximum SPL used; thus, the threshold is unknown in these cases (only two animals' correlation coefficients are averaged). There does not seem to be a general relationship between enhancement and CM for the type III case.

III. DISCUSSION

A. Effects of acoustic trauma

We observe considerable variability in the effect of acoustic trauma on EEOE enhancement. We believe that the variability is not the result of cochlear degradation due to the invasiveness of the experimental procedure because the type of EEOE change due to acoustic trauma does not correlate with initial N_1 threshold. This variability is not surprising, since many investigators have shown that cochleas vary in their response to acoustic trauma (see Borg *et al.*, 1995 for a review).

We chose to group the effect of acoustic trauma on acoustic enhancement into three categories. The type I case is characterized by the low-level, unenhanced emissions increasing in magnitude, while the maximally acoustically enhanced emissions remained basically unchanged (cf. the upper-left panel of Fig. 3). The type II case is characterized by increased low-level, unenhanced emissions, and also decreased maximally enhanced emissions (cf. the upper-middle panel of Fig. 3). Both type I and type II cases occurred in cochleas with high EP. The type III case is characterized by a decrease in all emission magnitudes, the low-level, unenhanced, as well as the maximally enhanced emissions (cf. the upper-right panel of Fig. 3). This type of effect generally occurred in cochleas with low EP.

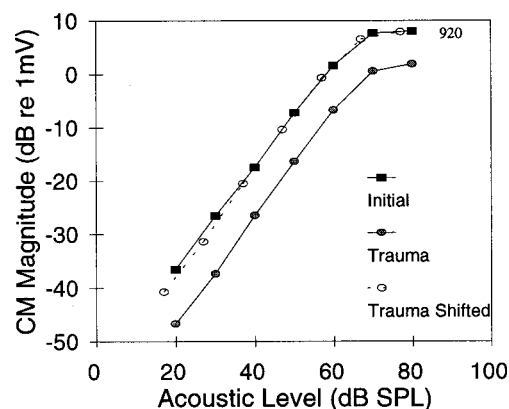


FIG. 7. The CM magnitude for $f_A = 400$ Hz. The initial recordings are depicted with solid square markers and recordings after all traumatizing tones were presented are depicted with filled circular markers. The open circular markers connected with dotted lines represent the CM after trauma that has been shifted up 6 dB and left 3 dB to superimpose with the initial CM.

All three types of acoustic-trauma effects had some common features. Most notably, acoustic trauma generally resulted in decreased acoustic enhancement of EEOEs. Also, the low-frequency CM magnitude generally decreased after acoustic trauma, with more decrease at low sound levels than at high sound levels. In other words, the CM input-output curve shifted down as well as to the right after acoustic trauma. Figure 7 shows that if the CM input-output curve after acoustic trauma is shifted up 6 dB and left 3 dB, it can be superimposed onto the initial CM.

B. Acoustic enhancement hypotheses

1. Feedback hypothesis

To interpret the results of EEOE experiments, we have used a feedback hypothesis in which the forward and reverse transduction processes of the OHCs form the feedback loop (Mountain *et al.*, 1983; Nakajima *et al.*, 1996). This hypothesis has been used to explain the measured responses of EEOEs, electrically evoked basilar-membrane motion, CM, and the thresholds of the N_1 waveform of the whole auditory-nerve response. Figure 8 shows a schematic of a negative feedback system representing the forward and reverse transduction processes. At low frequencies (lower than the given location's CF), we expect feedback to be negative. To understand this, consider that when the basilar membrane moves towards the scala media, the hair bundles are displaced toward the spiral ligament (towards the taller stereocilia), due to the shear between the reticular lamina and the tectorial

membrane. This displacement depolarizes the OHCs. *In vitro* studies have demonstrated that depolarization of OHCs results in their contraction (Kachar *et al.*, 1986), which will pull the reticular lamina back down and towards the ligament and decrease the deflection of the hair bundles. Therefore, for frequencies below the CF of a given location, the feedback is negative and decreases reticular lamina motion. However, for frequencies near the CF, the feedback is hypothesized to be positive due to phase shifts in the feedback loop. The positive feedback would increase the hair-bundle motion at frequencies near CF. This phase shift could occur in any or all of the blocks represented in Fig. 8. A phase lag can be from either mechanical or electrical process(es). We believe that this feedback accounts for the high selectivity and sensitivity of the cochlea.

Data from several EEOE experiments support the hypothesis that the feedback is negative for low frequencies and positive for frequencies near CF. Manipulations, such as the early stages of death where the EP drops (Nakajima *et al.*, 1994), and acoustic trauma (Nakajima *et al.*, 1996) support the feedback hypothesis. Both these manipulations (death and acoustic trauma) are expected to affect the forward transduction process which would open the feedback loop, resulting in an increase in EEOE below CF and decrease in EEOE near CF.

When a tone is presented simultaneously with a low-frequency electrical current stimulus, the EEOE magnitude is enhanced (Mountain and Hubbard, 1989; Murata *et al.*, 1991; Xue *et al.*, 1993; Kirk and Yates, 1996). In the context of the negative feedback hypothesis, a moderately loud simultaneous acoustic stimulus would saturate the nonlinear forward transduction process, decreasing the total gain in the feedback loop. In this manner, the normal effect of the negative feedback would be removed, and the reticular lamina motion would increase. Near CF, where the feedback is positive, a simultaneous acoustic stimulus would decrease the EEOE magnitude. Xue (1993) tested this prediction, and the feedback model was consistent with his results. He found that acoustic enhancement of EEOEs occurred for electrical frequencies below the CF of the electrode location. Near CF, the acoustic enhancement of EEOEs decreased and often became acoustic suppression. Similar experiments were conducted by Kirk and Yates (1996) in the guinea pig cochlea. Their results differed from Xue's, in that they did not observe a transition from enhancement to suppression as the emission frequency approached the CF of the stimulation site. These results led Kirk and Yates to question the negative feedback hypothesis.

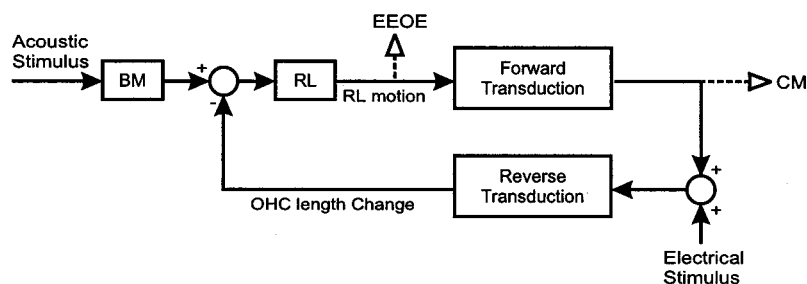


FIG. 8. A schematic of the negative feedback hypothesis for low-frequency electrical stimulation. Sound produces motion at the basilar membrane (BM) and reticular lamina (RL). The OHC hair-bundle motion is converted to receptor current by the forward transduction process. The receptor current is converted to OHC length change by the reverse transduction process. At low frequency, the feedback works to decrease the RL motion. The measured signals, EEOE and CM, as well as the electrical current and sound stimuli to evoke them, are shown within the context of the feedback system.

2. Conductance change hypothesis

Yates and Kirk (1997) subsequently proposed an alternative hypothesis for the acoustic enhancement effect. Their hypothesis states that if the majority of the stimulus current passes through the apical transduction channels and if simultaneous acoustic stimulation increases the average apical conductance, then the flow of stimulus current through the cell would increase. This increase in current flow would result in an increase in the component of the basolateral potential at the electrical frequency, resulting in an enhanced EEOE. However, our modeling studies indicate that enhancement from an increase in average apical conductance is small and does not account for the large enhancement observed, while a model utilizing the negative feedback hypothesis can closely simulate acoustic enhancement (Nakajima *et al.*, 2000).

3. Reflection hypothesis

Nuttall and Ren (1995) and Ren and Nuttall (1995) pioneered the study of the EEOE evoked by extracochlear electrical stimulation near the round window. The EEOE evoked by extracochlear current differs from the EEOEs evoked by scala media current injection in that the extracochlearly evoked EEOE frequency response has more peaks and notches. It may be that their type of emission is the result of contributions from a large fraction of the frequency place map. Alternatively, Ren and Nuttall (1998b, 1999) have hypothesized that EEOEs consist of two components: a short-latency component originating from near the electrode, and a long-latency component originating from the best place for the stimulus frequency. The long-latency component travels from the electrode location to the best place and then is reflected back towards the stapes. The peaks and notches observed in the frequency response are hypothesized to be the result of phase cancellation between the two components.

Ren and Nuttall (1998a) studied the effect of simultaneous acoustic stimulation on the extracochlearly evoked EEOE and found that sound altered the fine structure of the frequency response. Their interpretation of this result is that sound changed “the intracochlear reflection pattern of an electrically evoked traveling wave.” We believe that such a mechanism does not play a major role in our experiments because we find little evidence for reflections.

Furthermore, Ren and Nuttall (1998a) showed that Furosemide did not enhance extracochlearly evoked EEOEs. However, Hubbard and Mountain (1990) and Murata *et al.* (1991) have shown enhancement of below-CF EEOE evoked by current injected into the scala media, as would be expected by the negative feedback hypothesis.

C. Interpretation of data using the negative feedback hypothesis

1. Interpretation of EEOE data

We can interpret the observed effects of acoustic trauma on acoustic enhancement of EEOE using the negative feedback model. To do this, we make predictions about the effects of acoustic trauma on the forward and reverse transduc-

tion processes. These predictions are then tested against measured variations in CM and N_1 . It is important to note that the basis for the classification scheme of the acoustic trauma types I, II, III relies only on how acoustic trauma affects acoustic enhancement.

For the type I case, the unenhanced EEOE (low-level simultaneous acoustic stimulation) increases, while the enhanced EEOE (high-level simultaneous acoustic stimulation) remains the same (upper-left panel of Fig. 3). This result would represent a situation where the traumatizing tone primarily affects the forward transduction gain. Thus, the effectiveness of the negative feedback is reduced, allowing the originally unenhanced emission to increase. However, reverse transduction seems to remain intact because the maximally acoustically enhanced emission magnitude remains the same.

In the type II case, acoustic trauma caused the low-level EEOEs to increase and the high-level EEOEs to decrease (upper-middle panel of Fig. 3). Similar to the type I case, the low-level EEOEs are believed to have increased due to the decrease in forward transduction gain. However, in the type II case the maximally enhanced EEOEs decrease, suggesting that the reverse transduction gain is also affected.

The type III case exhibited decreases for both low-level EEOEs and high-level EEOEs (upper-right panel of Fig. 3). The decreases suggest that trauma significantly reduced the reverse transduction gain in type III cochleas. This type of acoustic trauma effect was typical in cochleas with EP that dropped during the experiment and that had low enhancement before trauma. Therefore, the forward transduction process may have already been compromised before the acoustic trauma presentation.

The change in the low-level EEOE observed in this study is consistent with the changes reported by Nakajima *et al.* (1996). The previous study found increases in unenhanced EEOEs below CF due to acoustic trauma in high-EP cochleas, similar to the type I and type II responses reported here. However, in this study we also explored the effect of acoustic trauma on EEOEs that were acoustically enhanced, allowing us to draw conclusions about the OHC mechanisms being affected by acoustic trauma. Low-EP characteristics were also consistent between the Nakajima *et al.* (1996) study and the present one. Kirk and Yates (1998) also observed similar acoustic trauma results to Nakajima *et al.* (1996). In one animal, Kirk and Yates observed increases in EEOEs, similar to our type I and type II cases. In several animals, Kirk and Yates observed decreases in EEOEs, similar to our type III case.

2. Interpretation of N_1 and CM data

We find in this study that the decrease in acoustic enhancement of the low-frequency EEOE is closely correlated with the increase in N_1 thresholds at CF for the high-EP animals (type I and II, Fig. 6). The negative feedback hypothesis predicts that acoustic enhancement of emissions is related to the low-frequency loop gain. N_1 thresholds at the CF of the electrode site give us information about the gain of the cochlear amplifier at the CF. Thus, the high correlation between enhancement and N_1 thresholds is in agreement

with the feedback hypothesis, where acoustic trauma would be expected to increase reticular lamina motion for low frequencies and decrease motion at CF.

The saturated high-level CM is proportional to the maximum receptor current, reflecting the number of functioning tension-gated channels in OHCs. Acoustic trauma generally decreases CM, and this decrease in CM does not always correlate strongly with N_1 threshold changes at CF. In the type I case, we hypothesize that trauma affects mostly the forward transduction process and there is good correlation between CM decrease and N_1 threshold increase (Fig. 5). However, in the type II case where acoustic trauma affects both the forward and reverse transduction processes, there is a lower correlation between CM and N_1 threshold changes.

If acoustic trauma affected only the forward transduction process, the CM is expected to be strongly correlated with N_1 threshold. That is, damage only to the forward transduction process would result in a CM decrease and a N_1 threshold increase. On the other hand, if only the reverse transduction process was damaged, then we would expect a CM increase and an N_1 threshold increase. Therefore, if acoustic trauma affects both the forward and reverse transduction process, then the CM would be expected to show less correlation with N_1 , as was seen in the type II data. This is because CM would decrease due to compromised forward transduction and increase due to compromised reverse transduction, and these effects on CM would cancel each other out except at high acoustic levels. In contrast, acoustic enhancement would be expected to show strong correlation with N_1 for type I and type II cases because both acoustic enhancement and N_1 threshold would decrease due to compromise of either the forward or reverse transduction process within the negative feedback loop.

D. Possible mechanisms of damage to forward transduction

Because our data lead us to believe that moderate acoustic trauma in healthy cochleas mainly affects forward transduction, we consider two possible mechanisms by which the forward transduction process may be compromised.

One possibility is that acoustic trauma decreases the number of active forward transduction channels (Patuzzi *et al.*, 1989b). This hypothesis is illustrated in the upper-left panel of Fig. 9, where the apical conductance of the OHC is plotted as a function of cilia displacement. Decreasing the number of active channels decreases the maximum conductance, as shown by the arrow.

The other possibility is that acoustic trauma alters the sensitivity of individual channels instead of altering the number of channels. Howard and Hudspeth (1988) proposed that mechanical stimuli stretch elastic components of the hair cell, the gating springs, which pull directly on the channels and open them. In their model, the single-channel sensitivity depends on the stiffness of the gating springs. The arrows in the lower-left panel of Fig. 9 illustrate how forward transduction can change due to the change in gating-spring stiffness. The slope of the transducer function decreases, while

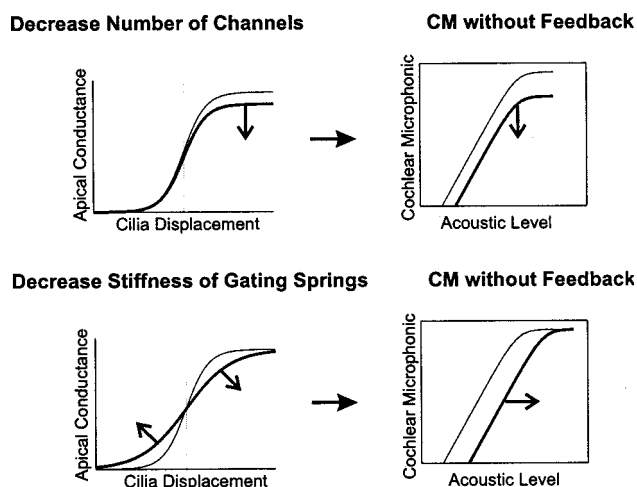


FIG. 9. Illustration of possible ways the forward transduction process can change due to acoustic trauma (left column), and its effect on CM (right column) without feedback. The upper-left panel shows how the forward transduction process would change if the number of active apical channels decreased. The lower-left panel illustrates the effect of decreasing the stiffness of the gating springs in the gating-spring model for forward transduction. With no feedback, CM would shift down due to decreased apical conductance (upper-right panel). With no feedback, CM would shift to the right due to decreased gating-spring stiffness (lower-right panel).

the maximum and minimum conductance values remain the same.

How these changes in forward transduction will affect low-frequency CM is illustrated in the right column of Fig. 9. First, consider what would happen to CM if the system did not have negative feedback. The decrease in the number of active apical channels due to acoustic trauma would lead to a downward shift in the CM input–output curve, as illustrated in the upper-right panel of Fig. 9. An example of this type of behavior was reported by Patuzzi *et al.* (1989b), who observed that the CM input–output curve shifted down vertically after acoustic trauma.

The second possible consequence of acoustic trauma is the decrease in gating-spring stiffness. Because the gating springs are hypothesized to be associated with the stereocilia tip links, mild acoustic trauma may be altering the tip-link stiffness. In a system without negative feedback, the CM input–output curve would shift to the right, as illustrated in the lower-right panel of Fig. 9.

If we now consider a system with negative feedback, the CM changes due to alterations in forward transduction changes are more complicated. For the case where there is a decrease in the number of active forward transduction channels, there will be a larger CM decrease at high stimulus levels than at low stimulus levels (cf. the upper-right panel of Fig. 10). This is because at lower acoustic levels, the negative feedback is active and would reduce the sensitivity of the closed-loop system to a reduction in forward transduction gain, while at high acoustic levels, the feedback is compromised and the reduction in forward transduction gain would be directly apparent in the CM.

In the case where the gating-spring stiffness decreases due to acoustic trauma, the CM would be affected by the negative feedback system, as illustrated in the lower right panel of Fig. 10. The resulting CM input–output curve

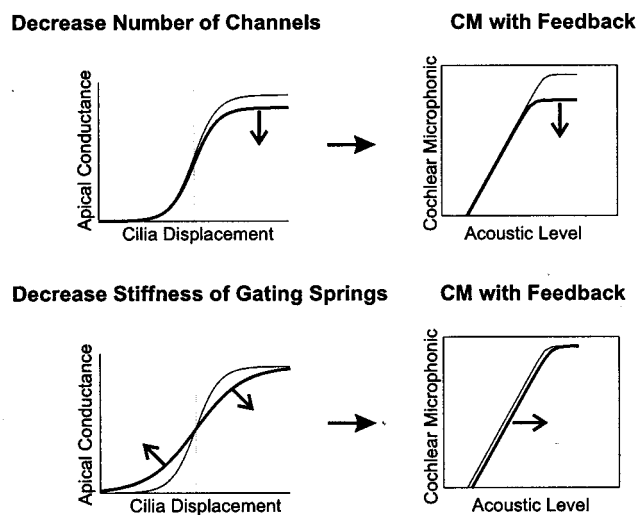


FIG. 10. Illustration of possible ways the forward transduction process can change due to acoustic trauma (left column), and its effect on CM (right column) with feedback. The upper-left panel shows how the forward transduction process would change if the number of active apical channels decreased. The lower-left panel illustrates the effect of decreasing the stiffness of the gating springs in the gating-spring model for forward transduction. With feedback, the saturated CM would decrease due to a decrease in apical conductance (upper-right panel). With feedback, the CM would shift slightly to the right due to decreased gating-spring stiffness (lower-right panel).

would shift to the right, but not as much as in a system without negative feedback.

In our experimental results, the CM decreased more at low stimulus levels than at high due to acoustic trauma. Figure 7 shows that if the CM input-output curve after acoustic trauma is shifted up by 6 dB and left by 3 dB, it can be superimposed on the initial CM curve. Therefore, it is likely that a combination of a decrease in the number of apical channels (causing the high-level CM to shift down) and a decrease in the stiffness of the gating springs (causing the CM curve to shift to the right) are responsible for the recorded changes in the CM input-output curve after acoustic trauma.

E. Possible mechanisms of damage to reverse transduction

Acoustic trauma can sometimes decrease emissions, as shown in the type II and III cases (Fig. 3). A decrease in the EEOEs could be due to trauma's effect on the reverse transduction process, such as inactivation of the motor molecules. Another possibility is OHC hyperpolarization, resulting in a decrease in the reverse transduction gain. Acoustic trauma could also decrease the ac basolateral membrane potential, which drives the reverse transduction process. It is also possible that the OHCs are mechanically destroyed. Each of these possibilities is considered below.

Acoustic trauma might compromise reverse transduction by the inactivation of motor molecules of the OHC. Inactivation could be the result of conformational changes in the molecules to some voltage-insensitive state, blocking reverse transduction. Motor molecule inactivation seem unlikely because OHCs presumably undergo substantial stress during

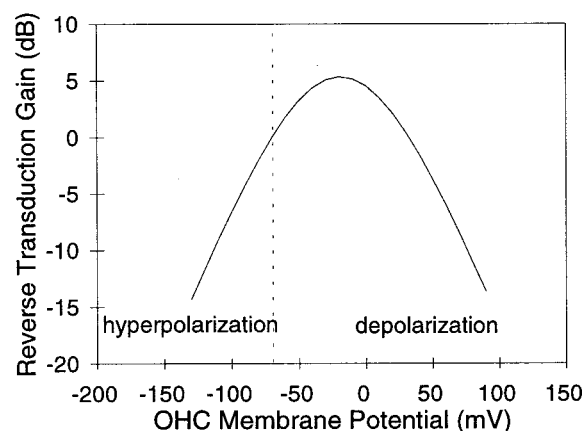


FIG. 11. A plot of the OHC reverse transduction gain versus OHC membrane potential. The resting membrane potential of -70 mV is shown by the vertical dotted line. If the membrane potential is hyperpolarized, the reverse transduction gain would decrease (adapted from Fig. 1B, Roddy *et al.*, 1994).

isolation procedures, yet isolated OHCs show robust reverse transduction (Brownell *et al.*, 1985; Kachar *et al.*, 1986; Ashmore, 1987).

Acoustic trauma might result in a change in resting membrane potential, which would change the reverse transduction gain (Fig. 11). If trauma hyperpolarizes the OHC, then the operating region of the OHC length-voltage curve would be shifted to a less sensitive region. However, this seems unlikely because Cody and Russell (1985) found that OHCs exhibited a sustained depolarization of membrane potential after acoustic trauma, which would increase reverse transduction gain.

The depolarization observed by Cody and Russell (1985) may be due to an increase in the OHC basolateral conductance. Such a conductance increase would result in a decrease in the ac basolateral membrane potential produced by the injected current. Emissions would then decrease because the reverse transduction process is driven by the ac basolateral membrane potential. If there is an increase in the basolateral membrane conductance and forward transduction remains undamaged, then we might expect a small increase in CM magnitude and a decrease in the emission magnitude. An example of this phenomenon is shown in Fig. 3 for the type II case. The type II emission decreases at high levels for the first trauma, while the corresponding CM increases at the high levels for the first trauma. Increase in CM after acoustic trauma has also been reported previously (Nakajima *et al.*, 1996).

Another possible effect of acoustic trauma, resulting in EEOE decrease, is the mechanical destruction of OHCs. Anatomical studies on acoustic trauma have revealed that changes observed in the physiology can be correlated with changes in cochlear structure. Liberman and Kiang (1978) observed that the first row of OHCs typically showed the most extensive damage after 1 to 4 h of 100 to 120 dB SPL of noise (25- to 400-Hz bandwidth centered between 1.5 and 12 kHz, white noise, or low pass). Stereocilia orderliness has been correlated with increased thresholds in nerve fibers after 2 h of 108- to 116-dB SPL noise (width of 50 Hz to 4.5 kHz)

centered between 0.75 to 6 kHz (Liberman and Beil, 1979). Acute swelling and vacuolization and “blebbing” or bulging of the endolymphatic surface of hair cells have been observed after an exposure of narrow-band noise centered between 0.5 and 15 kHz for 1 to 2 h at intensities ranging from 100 to 125 dB SPL (Liberman and Mulroy, 1982). These anatomical studies typically used traumatizing noise for periods that were long compared to our short traumatizing-tone presentations. However, Liberman and Mulroy (1982) found that acute acoustic trauma produced by pure-tone or narrow-band noise 10 to 30 min at frequencies between 1 and 6 kHz and intensities of 100 to 110 dB SPL resulted in minimal hair cell loss, but with several cases of scattered OHCs showing nuclear swelling and/or cytoplasmic rupture. Therefore, in some of our experiments, the traumatizing tones may have resulted in mechanical damage or loss of OHCs.

IV. SUMMARY

Moderate acoustic trauma causes a decrease in the acoustic enhancement of EEOEs and the decrease in enhancement correlates strongly with an increase in N_1 threshold. Our interpretation of the data is that acoustic trauma affects the forward transduction process by decreasing the number of functioning apical channels and by decreasing the stereocilia gating-spring stiffness. Acoustic trauma can also affect the reverse transduction process by increasing the OHC basolateral membrane conductance.

ACKNOWLEDGMENTS

Lisa Olson and Ram Naidu provided many helpful comments. This work was supported by NSF and NIDCD.

- Ashmore, J. F. (1987). “A fast motile response in guinea pig outer hair cell: The cellular basis of the cochlear amplifier,” *J. Physiol. (London)* **388**, 323–347.
- Borg, E., Canlon, B., and Engström, B. (1995). “Noise-induced hearing loss, literature review and experiments in rabbits,” *Scand. Audiol. Suppl.* **24**, 1–147.
- Brownell, W. E., Bader, C. R., Bertrand, D., and de Ribaupierre, Y. (1985). “Evoked mechanical responses of isolated cochlear outer hair cells,” *Science* **277**, 194–196.
- Cody, A. R., and Russell, I. J. (1985). “Outer hair cells in the mammalian cochlear and noise-induced hearing loss,” *Nature (London)* **315**, 662–665.
- Cody, A. R., and Russell, I. J. (1988). “Acoustically induced hearing loss: Intracellular studies in the guinea pig cochlea,” *Hear. Res.* **35**, 59–70.
- Dallos, P., Billone, M. C., Durrant, J. D., Wang, C. Y., and Raynor, S. (1972). “Cochlear inner and outer hair cells: Functional differences,” *Science* **177**, 356–358.
- Davis, H. (1983). “An active processing of cochlear mechanics,” *Hear. Res.* **9**, 79–90.
- Engström, H., Andes, H. W., and Bredberg, G. (1970). “Normal structure of the organ of Corti and the effect of noise-induced cochlear damage,” in *Sensorineural Hearing Loss. A Ciba Foundation Symposium*, edited by G. W. E. Wolstenholme and J. Knight (J & A Churchill, London), Vol. 197, pp. 127–156.
- Howard, J., and Hudspeth, A. J. (1988). “Compliance of the hair bundle associated with gating of the mechanoelectrical transduction channels in the bullfrog’s saccular hair cell,” *Neuron* **1**, 189–199.
- Hubbard, A. E., and Mountain, D. C. (1990). “Hair cell forward and reverse transduction: differential suppression and enhancement,” *Hear. Res.* **43**, 269–272.
- Hunter-Duvar, I. M. (1977). “Morphology of the normal and acoustically damaged cochlea,” in *Scanning Electron Microscopy*, Vol. 2 (ITT Research Institute, Chicago), pp. 421–428.
- Kachar, B., Brownell, W. E., Altschuler, R. A., and Fex, J. (1986). “Electrokinetic shape changes in cochlear outer hair cells,” *Nature (London)* **322**, 365–367.
- Kemp, D. T. (1978). “Stimulated acoustic emissions from within the human auditory system,” *J. Acoust. Soc. Am.* **64**, 1386–1391.
- Kirk, D. L., and Yates, G. K. (1996). “Frequency tuning and acoustic enhancement of electrically evoked otoacoustic emissions in the guinea pig cochlea,” *J. Acoust. Soc. Am.* **100**, 3714–3725.
- Kirk, D. L., and Yates, G. K. (1998). “4-Aminopyridine in scala media reversibly alters the cochlear potentials and suppresses electrically evoked oto-acoustic emissions,” *Audiol. Neuro-Otol.* **3**, 21–39.
- Liberman, M. C., and Beil, D. G. (1979). “Hair cell condition and auditory nerve response in normal and noise-damaged cochleas,” *Acta Oto-Laryngol.* **88**, 161–176.
- Liberman, M. C., and Kiang, N. Y. S. (1978). “Acoustic trauma in cats: Cochlear pathology and auditory-nerve activity,” *Acta Oto-Laryngol.* **358**, 1–63.
- Liberman, M. C., and Mulroy, M. J. (1982). “Acute and chronic effects of acoustic trauma, cochlear pathology and auditory nerve pathophysiology,” in *New Perspectives on Noise-Induced Hearing Loss*, edited by R. P. Hammernik, D. Henderson, and R. Salvi (Raven, New York), pp. 105–136.
- Mountain, D. C., and Hubbard, A. E. (1989). “Rapid force production in the cochlea,” *Hear. Res.* **42**, 195–202.
- Mountain, D. C., Hubbard, A. E., and McMullen, T. A. (1983). “Electromechanical processes in the cochlea,” in *Mechanics of Hearing*, edited by E. de Boer, and M. A. Viergever (Delft University Press, Delft), pp. 119–126.
- Murata, K., Moriyama, T., Hosokawa, Y., and Minami, S. (1991). “Alternating current induced otoacoustic emissions in the guinea pig,” *Hear. Res.* **55**, 201–214.
- Nakajima, H. H., Hubbard, A. E., and Mountain, D. C. (2000). “A physiologically-based nonlinear active feedback model of cochlea mechanics,” in *Recent Developments in Auditory Mechanics*, edited by H. Wada *et al.* (World Scientific, Singapore), pp. 202–208.
- Nakajima, H. H., Mountain, D. C., and Hubbard, A. E. (1998). “Nonlinear characteristics of electrically evoked otoacoustic emissions,” *Hear. Res.* **122**, 109–118.
- Nakajima, H. H., Olson, E. S., Mountain, D. C., and Hubbard, A. E. (1994). “Electrically-evoked otoacoustic emissions from the apical turns of the gerbil cochlea,” *J. Acoust. Soc. Am.* **96**, 786–794.
- Nakajima, H. H., Olson, E. S., Mountain, D. C., and Hubbard, A. E. (1996). “Acoustic overstimulation enhances low-frequency electrically-evoked otoacoustic emissions and reduces high-frequency emissions,” *Aud. Neurosci.* **3**, 79–99.
- Neely, S. T., and Kim, D. O. (1983). “An active cochlear model showing sharp tuning and high sensitivity,” *Hear. Res.* **9**, 123–130.
- Nuttall, A. L., and Ren, T. (1995). “Electromotile hearing: Evidence from basilar membrane motion and otoacoustic emissions,” *Hear. Res.* **92**, 170–177.
- Patuzzi, R., Johnstone, B. M., and Sellick, P. M. (1984). “The alteration of the vibration of the basilar membrane produced by loud sound,” *Hear. Res.* **13**, 99–100.
- Patuzzi, R. B., Yates, G. K., and Johnstone, B. M. (1989a). “The origin of the low-frequency microphonic in the first cochlear turn of guinea pig,” *Hear. Res.* **39**, 177–188.
- Patuzzi, R. B., Yates, G. K., and Johnstone, B. M. (1989b). “Changes in cochlear microphonic and neural sensitivity produced by acoustic trauma,” *Hear. Res.* **39**, 189–202.
- Ren, T., and Nuttall, A. L. (1995). “Extracochlear electrically evoked otoacoustic emissions: a model for *in vivo* assessment of outer hair cell electromotility,” *Hear. Res.* **92**, 178–183.
- Ren, T., and Nuttall, A. L. (1998a). “Acoustical modulation of electrically evoked otoacoustic emission in intact gerbil cochlea,” *Hear. Res.* **120**, 7–16.
- Ren, T., and Nuttall, A. L. (1998b). “Multiple origins of the electrically evoked otoacoustic emission,” in *Association for Research in Otolaryngology Abstracts*, 21, p. 179, <http://www.aro.org>.
- Ren, T., and Nuttall, A. L. (1999). “Acoustical self-interference and magnification in the cochlea,” in *Association for Research in Otolaryngology Abstracts*, 22, p. 120, <http://www.aro.org>.
- Ren, T., Nuttall, A. L., and Miller, J. M. (1996). “Electrically evoked cubic distortion product otoacoustic emissions from gerbil cochlea,” *Hear. Res.* **102**, 43–50.

- Roddy, J., Hubbard, A. E., Mountain, D. C., and Xue, S. (1994). "Effects of electrical biasing on electrically-evoked otoacoustic emissions," *Hear. Res.* **73**, 148–154.
- Sellick, P. M., Patuzzi, R., and Johnstone, B. M. (1982). "Measurement of the basilar membrane motion in the guinea pig using the Mossbauer technique," *J. Acoust. Soc. Am.* **72**, 131–141.
- Xue, S. (1993). "Measurement of basilar membrane motion and otoacoustic emissions in response to electrical and acoustic stimulation," Ph.D. thesis, Boston University, Boston, MA.
- Xue, S., Mountain, D. C., and Hubbard, A. E. (1993). "Acoustic enhancement of electrically-evoked otoacoustic emissions reflects basilar membrane tuning: Experiment results," *Hear. Res.* **70**, 121–126.
- Xue, S., Mountain, D. C., and Hubbard, A. E. (1996). "Electrically-evoked otoacoustic emissions: direct comparisons with basilar membrane motion," *Aud. Neurosci.* **2**, 301–308.
- Yates, G. K., and Kirk, D. L. (1997). "Modulation of electrically-evoked emissions by forward transduction gates in the outer hair cell," in *Association for Research in Otolaryngology Abstracts*, **20**, p. 123, <http://www.aro.org>.

Auditory-nerve-fiber responses to high-level clicks: Interference patterns indicate that excitation is due to the combination of multiple drives

Tai Lin

Eaton-Peabody Laboratory of Auditory Physiology, Department of Otolaryngology, Massachusetts Eye and Ear Infirmary, 243 Charles Street, Boston, Massachusetts 02114 and Harvard-MIT Division of Health Sciences and Technology, and Research Laboratory of Electronics, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139

John J. Guinan, Jr.^{a)}

Eaton-Peabody Laboratory of Auditory Physiology, Department of Otolaryngology, Massachusetts Eye and Ear Infirmary, 243 Charles Street, Boston, Massachusetts 02114; Harvard-MIT Division of Health Sciences and Technology, and Research Laboratory of Electronics, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139; and Department of Otology and Laryngology, Harvard Medical School, Boston, Massachusetts 02115

(Received 4 October 1999; accepted for publication 24 January 2000)

There has been no systematic study of auditory-nerve-fiber (ANF) responses to high-level clicks despite the advantages of clicks in revealing the natural resonances of a system. Cat single ANFs were studied using clicks up to 120 dB pSPL. Peri-stimulus-time (PST) histograms of responses were corrected for refractory effects, and compound PST (cPST) histograms were formed from rarefaction- and condensation-click PSTs. At low levels the responses followed the classic picture with each cPST appearing to be from a single resonant system followed by low-pass filtering that reduces high-frequency synchrony. In fibers across all characteristic frequencies, there were significantly different patterns at high click levels including several nonclassic features and “phase reversals,” i.e., a peak in the rarefaction-click PST at low levels was replaced at high levels by a peak at the same latency in the condensation-click PST. There were two separate regions of nonclassic features and phase reversals, which indicates that auditory-nerve fibers are excited by the combination at some stage in the cochlea of at least three excitation drives derived from the acoustic stimulus. These data support the interpretation that the cochlear partition vibrates in multiple resonant modes with each mode producing one excitation drive and that the mix of modes varies with sound level. © 2000 Acoustical Society of America. [S0001-4966(00)00205-8]

PACS numbers: 43.64.Pg, 43.64.Kc, 43.64.Nf [LHC]

INTRODUCTION

In the classic view, auditory-nerve-fiber responses to clicks are due to a resonant response of the basilar-membrane at the fiber characteristic frequency (CF) that is half-wave rectified by the inner-hair-cell (IHC) synapse. For low-CF fibers (CFs <4 kHz), peristimulus-time (PST) histograms of click responses show multiple peaks separated by 1/CF, and the peaks from rarefaction and condensation clicks neatly interleave (e.g., the histograms from the lowest click levels in Fig. 3) (Kiang *et al.*, 1965). High-CF fibers (CFs ≥4 kHz) responding to low-level clicks typically show only a single PST peak that is almost identical for rarefaction and condensation clicks (Kiang *et al.*, 1965), with the loss of high-frequency synchrony being due to low-pass filtering in IHCs and at the IHC-ANF synapse that reduces the ac component of the response.

Despite the classic view, various reports show nonclassic click responses, particularly for high-level clicks (Kiang *et al.*, 1965; Pfeiffer and Kim, 1972; Antoli-Candela and

Kiang, 1978; Versnel *et al.*, 1992, 1997; Schoonhoven *et al.*, 1994). High-CF fibers in response to high-level clicks show short-latency PST peaks that have different latencies for rarefaction and condensation clicks (Kiang *et al.*, 1965; Schoonhoven *et al.*, 1994). For some low-CF fibers, waxing and waning has been observed in click-response PSTs, and a model has been presented to explain this (Pfeiffer and Kim, 1972). However, the model assumed that these auditory-nerve fibers innervated multiple outer hair cells (OHCs), an assumption that does not fit with present knowledge of cochlear anatomy and of the fibers contacted with pipette electrodes (Liberman, 1982; Ryugo, 1992). No study has attempted to systematically document nonclassic responses across CFs and at both low and high click levels, and with the possible exception of the short-latency peaks in high-CF fibers, there is no viable explanation for the presence of nonclassic click responses.

An important motivation for the exploration of click responses comes from the advantage of clicks over tones in showing the natural resonances of a system. In steady-state responses to a tone, all structures respond at the frequency imposed by the tone (typically with different amplitudes and phases) even if their own natural resonant frequencies are

^{a)} Author to whom correspondence should be addressed; electronic mail: jjg@epl.meei.harvard.edu

different. In addition, distortion can produce other frequencies, i.e., harmonics, but these are still not the natural frequencies of the system. In contrast, clicks are an excellent stimulus for studying the resonances of a system because they deliver a burst of energy and then allow the system to vibrate on its own. Thus, evidence of the cochlea's natural resonant frequencies should be present in auditory-nerve-fiber responses to clicks.

We hypothesized that there might be multiple resonances to be revealed by clicks because the vibration of the organ-of-Corti complex (including the basilar and tectorial membranes) may consist of a mixture of many different vibration patterns, or "resonant modes," each with its own spatial distribution and natural resonant frequency (Zhang *et al.*, 1997; Mountain, 1998; Mountain and Cody, 1999; see also Richter and Dallos, 1998; Karavitsaki and Mountain, 1998). A simple example of such a system is the vibration of a plucked violin string, which is a mixture of vibrations at fundamental and harmonic frequencies. However, the organ of Corti, with its intricate three-dimensional structure, is likely to have resonant modes that are related in more complicated ways than just being a harmonic series. In addition, since the cochlea is nonlinear, its resonant vibration patterns may also vary with sound level. Thus, the natural frequencies shown in the click responses may also vary with sound level.

I. METHODS

A. Animal preparation and procedures

The basic methods were as in Kiang *et al.* (1965). Cats were anesthetized with Dial in urethane following a protocol approved by the Animal Care Committee of the Massachusetts Eye and Ear Infirmary. The ear canals were surgically exposed, the auditory bullae opened and a small hole (a few mm diameter) made in the bony septum to reduce the acoustic resonance of the middle-ear cavities. The auditory nerve was exposed by aspirating part of the cerebellum, and single fibers were recorded with a micropipette electrode in the internal auditory meatus. Accurate spike timing was obtained using a custom-made device that detected spikes higher than a preset level, but with the output time set to the spike peak. Auditory-nerve compound action potentials (CAPs) were monitored with an electrode near the round window. An automated tone-pip audiogram determined the sound level at which 0.5, 1, 2, 4, 8, 16, 22.6 and 32 kHz tones evoked criterion (5 μ V) CAPs.

Broadband noise bursts were used as a search stimulus. After making contact with a fiber, the data-gathering sequence was: (1) obtain a tuning curve and determine CF, (2) record the spontaneous firing for 20 s and calculate the spontaneous rate (SR), and (3) run click-level series. Each click-level series had a constant polarity (condensation or rarefaction) and randomized click levels to even out the effects of adaptation. Typically, 11 click levels were used and 500 clicks (1 every 30 ms) were averaged at each level, so that a single randomized run took about 3 min. When the fiber was contacted long enough (approximately 20% of fibers with data), a second tuning curve was collected. These tuning curves indicate that, for the fibers used in this paper, there

was little or no sensitivity change due to our use of high-sound-level clicks. Fibers were grouped into low ($SR \leq 0.5$), medium ($0.5 < SR < 18$), and high ($SR \geq 18$ spikes/s) categories (Liberman, 1978), but no differences across SR categories were found in any of the properties reported here.

B. The click stimulus

The acoustic assembly consisted of a 1 in condenser earphone [Bruel & Kjaer (B&K) 4145] and a 1/4 in condenser microphone (B&K 4136 or Larson Davis 2520) with a probe tube whose tip was placed a few mm from the tympanic membrane (similar to the assembly in Fig. 2.2 of Kiang *et al.*, 1965). Click intensity is expressed in peak-equivalent sound pressure level (pSPL, the SPL of a tone with the same peak sound pressure as the click) at the tympanic membrane, calculated by deconvolving (using MATHCAD V7 and/or LABVIEW 3.1, with similar results) the sound pressure waveform measured at the end of the probe tube and the probe-tube impulse response (the impulse response was calculated by a reverse Fourier transform of the measured probe-tube frequency-transfer function). Clicks (normally 90 μ s square pulses every 30 ms) were adjusted in amplitude in 0.1 dB steps to compensate for the earphone square-law nonlinearity (this was important to insure that rarefaction and condensation clicks were at the same pSPL). The parameters of the reverse-square-law compensation were determined empirically from measurements of the earphone output at various attenuations measured in a cavity. In the initial series of experiments (animals # 24–29), the assembly was used without acoustic damping in order to obtain the largest stimulus intensity possible, but in later experiments (animals 30–37) cotton was put in the assembly to reduce the acoustic ringing. Examples of click waveforms from early and late cats are given in Fig. 1. Although these clicks are not perfect acoustic impulses, we do not think any of the response features that we will present are due to imperfections of the click (see Sec. III).

C. Room echo

After the initial series of experiments we discovered that a room echo with a latency over 10 ms added acoustic energy at 3–6 kHz. These echoes could be seen in averages from the microphone in the ear canal, and they could be blocked by damping material placed over the hole in the bulla, or sheets of acoustic foam on the walls of the room. The sound path appears to be: sound in the ear canal vibrates the tympanic membrane, creating a sound in the open bulla cavity which radiates into the room, reflects from the walls of the room and returns to the bulla cavity, where it vibrates the tympanic membrane and is transmitted into the cochlea by the middle ear. The peak at 3–6 kHz appears to be due to the radiation (and receiving) properties of the hole in the tissue and the bulla. A similar frequency dependence was found for acoustic crosstalk from one ear to the other in a somewhat similar preparation [see Fig. 4(a) and 4(b) of Guinan *et al.*, 1972]. This echo caused some fibers with CFs in the 3–6 kHz range to have responses at the highest click levels with time-locked, bursted firing at latencies of 10–70

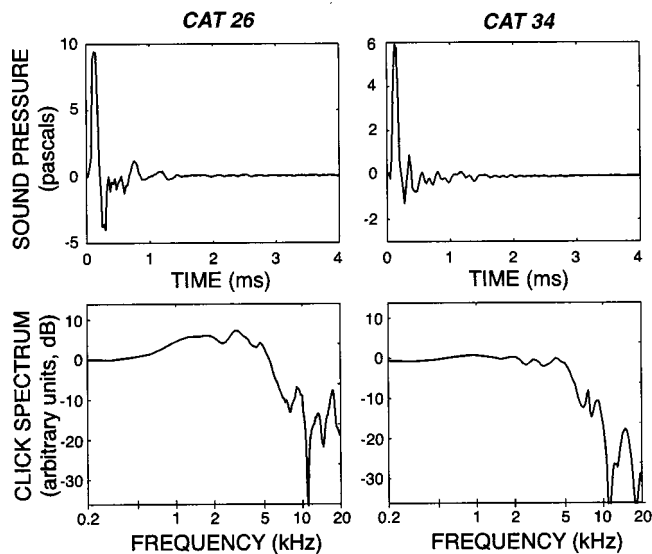


FIG. 1. Click sound pressure at the tympanic membrane with no damping material in the acoustic assembly (left) and with cotton damping in the acoustic assembly (right). Sound pressure at the tympanic-membrane was calculated from the sound pressure measured at the probe tube microphone by deconvolution using the measured transfer characteristic of the probe tube and microphone. The spectra at the bottom were calculated from the waveforms at the top (which extended for 6.4 ms); the two spectra are on the same scale.

ms. The echo affected only auditory-nerve fibers with CFs in the 3–6 kHz range on cats 24–29 and any fiber in this range showing evidence of responses with latencies over 10 ms has been removed from the data pool. The echo was effectively removed in later cats by lining the walls of the soundproof room with acoustic damping material.

D. Compound histograms

From the spike and click times stored at each click level, we calculated peri-stimulus-time (PST) histograms. Compound PST (cPST) histograms were formed by combining the rarefaction-click PST histogram with positive values and the corresponding condensation-click PST histogram with negative values (Fig. 2). Compound PST histograms provide a better estimate of the overall waveform of IHC receptor potentials and stereocilia motion than individual PSTs alone which presumably resemble rectified versions of these variables (Pfeiffer and Kim, 1972).

E. Recovered probability

Two aspects of auditory-nerve firing can distort PST histograms: refractory properties and a tendency for spikes to be paired at 3/4 ms intervals. So that the click responses would reveal the cochlear drive to the auditory-nerve fibers as accurately as possible, we computed recovered-probability PST histograms (Gray, 1967). A recovered-probability PST histogram shows the probability that a spike will occur in a time bin at a given latency, calculating this probability using only time bins (and spikes that occur in these bins) at least P seconds after a previous spike. As the exclusion period, P , is made larger, the effects of refractoriness (or any other spike-time related factor) are minimized at the expense of having

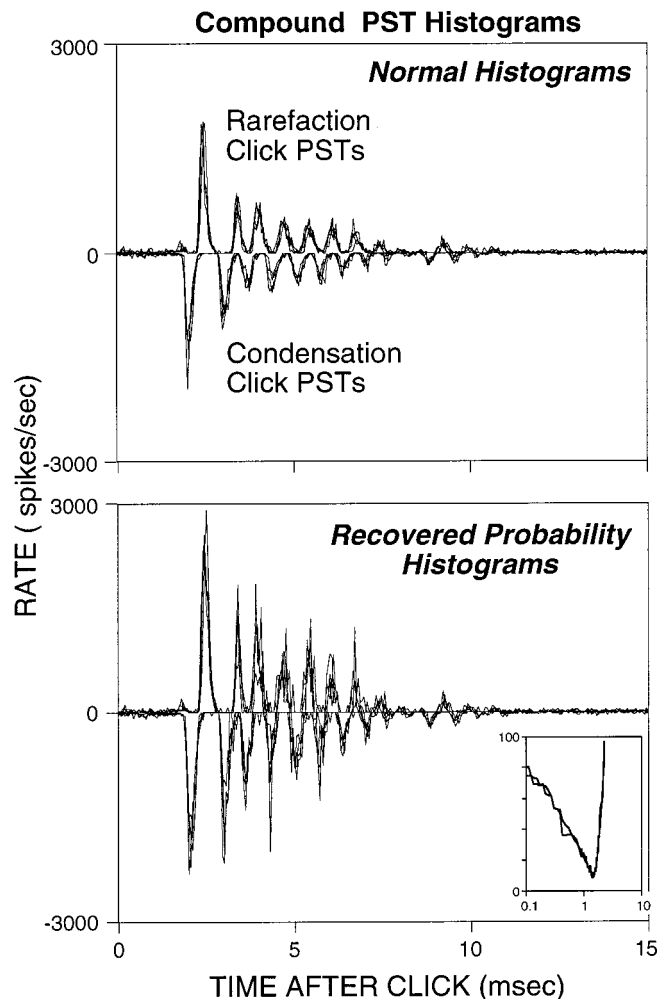


FIG. 2. Top: The formation of compound histograms. Bottom: The same compound histograms after the calculation of recovered probabilities. Each main panel shows four compound histograms obtained during a half-hour period from the same fiber. Compound PST histograms were formed by combining rarefaction-click peri-stimulus-time (PST) histograms with positive values (i.e., plotted upward) and the condensation-click PST histograms with negative values (i.e., plotted downward). Inset: Tuning curves obtained at the beginning and near the end of data taking on this fiber. Horizontal scale: Frequency (kHz). Vertical scale: Tone threshold (dB SPL). Fiber 26-7, CF: 1.40 kHz, SR: 64.6 spikes/s.

fewer total spikes and therefore more randomness in the resulting histogram. Gray (1967) found that a P of 20 ms was necessary to insure a fully recovered PST, but such a long P often removed most of the spikes and made the resulting PSTs very noisy. Fortunately, much shorter P 's yielded PSTs that appeared similar to those from fully recovered PSTs in the timing and presence of the peaks. For the figures and data analysis, we used recovered histograms with a P of 3 ms. This appeared to maximize the number of spikes while preserving the cPSTs shapes shown by recovered histograms with a P of 20 ms. Figure 2 shows an example of cPSTs before and after the recovered probability calculation. The biggest effect on the shapes of the cPSTs was to increase the amplitude of peaks after the first peak.

For most fibers and click levels, recovered-probability PSTs had the same features (i.e., peaks at the same times) as the regular PSTs, but some fibers had noticeable differences. For most fibers with very low CFs, each phase of the ‘‘raw’’

PST displayed a double peak, but the recovered-probability PST showed only a single peak. Presumably this “peak splitting” was due to the long excitation cycle in low-CF fibers (e.g., at 250 Hz, a half period is 2 ms) which allows time for a fiber to fire, to be refractory, to recover, and to fire again, all during a single excitation peak. For a few fibers, a waxing and waning of the response was more evident in the recovered probability cPSTs than in the raw cPSTs.

Finally, our use of a short repetition period (30 ms) produced strong adaptation which was probably an important aid in obtaining good recovered-probability PSTs with only 500 click repetitions. In a fully rested preparation, a high-level click evokes a spike in the first or second peak of the histogram with a very high probability, thereby making it difficult to obtain spikes at later times in the PST that were not preceded within 3 ms by a previous spike. This problem is greatly reduced in heavily adapted fibers. Note that the calculation of recovered probabilities removes short-term refractory effects but does not remove longer-term adaptation effects.

II. RESULTS

Auditory-nerve-fiber responses to clicks showed a variety of patterns, many of which were clearly not due to the response of a single oscillating system. An overview of typical click response patterns across CF regions is given in Figs. 3 and 4. Figure 3 shows compound PST histograms as a function of click level from six auditory-nerve fibers. In Fig. 3, response patterns are often difficult to grasp because seeing the patterns requires integrating information along three dimensions: latency, click level, and firing rate. To make the patterns easier to see, the data of Fig. 3 are redisplayed in Fig. 4 in a three-dimensional plot with cPST rate coded by color. The rate data in Fig. 4 have been linearly interpolated along the click-level axis to make it easier to trace the timing of cPST peaks across click levels.

At low click levels, auditory-nerve-fiber responses showed the classic pattern. For fibers with CFs < 4 kHz, the PSTs had multiple peaks with interpeak intervals of $1/\text{CF}$ and with rarefaction and condensation peaks interleaved so that the cPST histograms showed a simple oscillatory response pattern [e.g., the lowest click levels in Figs. 3 and 4, (a)–(e)]. For fibers with $\text{CF} \geq 4$ kHz, each low-level PST had a single peak, and rarefaction and condensation peaks were at the same time so that the cPST histograms were approximately symmetric about the zero line [Fig. 3(f), levels ≤ 70 dB pSPL]. [In Fig. 4, combining rarefaction- and condensation-click PSTs to produce a single color-coded “cPST rate” removes the part of the cPSTs that is symmetric about the zero line. Thus, Fig. 4(f) does not show the low-level response of this high-CF fiber.]

At high click levels, auditory-nerve-fiber responses were unremarkable in their firing rates; the interesting behavior was in the timing of the responses. As click level increased from threshold to over 100 dB pSPL, the response rate increased and saturated [Fig. 5(a)], presumably because adaptation limited the maximum average firing rate. However, despite this saturation in rate, the PSTs at high click levels changed with level because the timing of their peaks changed

(Figs. 3 and 4). Some insight into the degree of change across click levels can be obtained by comparing the PSTs at adjacent click levels on a bin-by-bin basis (details in Fig. 5 caption). As shown in Fig. 5(b), at low to medium click levels, PSTs showed only moderate changes that were due primarily to randomness in the histograms and to earlier peaks of the classic PST response being emphasized as click level increased. At high click levels, however, there were large changes from one click level to the next because there were major changes in the timing of individual peaks. Figure 5(b) shows that almost all fibers had much bigger changes from one level to the next at high levels compared to low levels, but does not indicate what these changes are.

To gain understanding of what changes produce the pattern of Fig. 5(b), we examined the data in plots similar to those in Figs. 3 and 4 and found many features of the responses that didn’t fit with a classic, smoothly oscillating pattern. Several of these nonclassic features were seen across many fibers and seemed worth noting. We have not found any automated way to detect these features so we visually scanned each cPST and noted by hand when the nonclassic features occurred. This was done with a high criterion, i.e., a nonclassic feature was noted only when we were confident that the feature observed could not be accounted for by a noisy response from a single resonant system followed by a low-pass filter. This could be done with a high degree of consistency, as judged by an initial test in which two observers applied the same criteria for nonclassic features to a subset of the data and agreed upon $> 70\%$ of the total number (which includes those found by either observer) of nonclassic features.

A. Anomalies and reversals

It is useful to separate the nonclassic features into two categories, “anomalies” and “reversals.” “Anomalies” are nonclassic response features (i.e., features that cannot be due to a single resonant system followed by a low-pass filter) that can be seen in the cPST from a *single* click level. “Reversals” are changes in the phase of the click (rarefaction or condensation) that produced the response at a given latency. Reversals can only be seen by comparing cPSTs across click levels. A third nonclassic feature was visible in the cPSTs: “glides,” a slow change of the interpeak interval (or oscillatory frequency) with time (de Boer and Nuttall, 1997; Carney *et al.*, 1999; see also Møller and Nilsson, 1979; Recio *et al.*, 1998). “Upward glides” were particularly visible for fibers with CFs of 2–3 kHz [Figs. 3 and 4, panels (d), (e)] and “downward glides” could be seen in very-low-CF fibers [Figs. 3 and 4, panel (a)]. Despite the presence of glides, we will focus on anomalies and reversals because we felt that they most directly show phenomena relevant to multiple resonances in the organ-of-Corti complex. We will reconsider this point in Sec. III.

We distinguished four types of anomalies and noted the click level and the latency of each occurrence of these anomalies in our data. The four types of anomalies are: “double peaks”: two successive peaks of the same polarity; the latency noted for this was the midpoint between the peaks [e.g., Fig. 3(b) at 106 dB, 3.4 ms]. “Overlapping

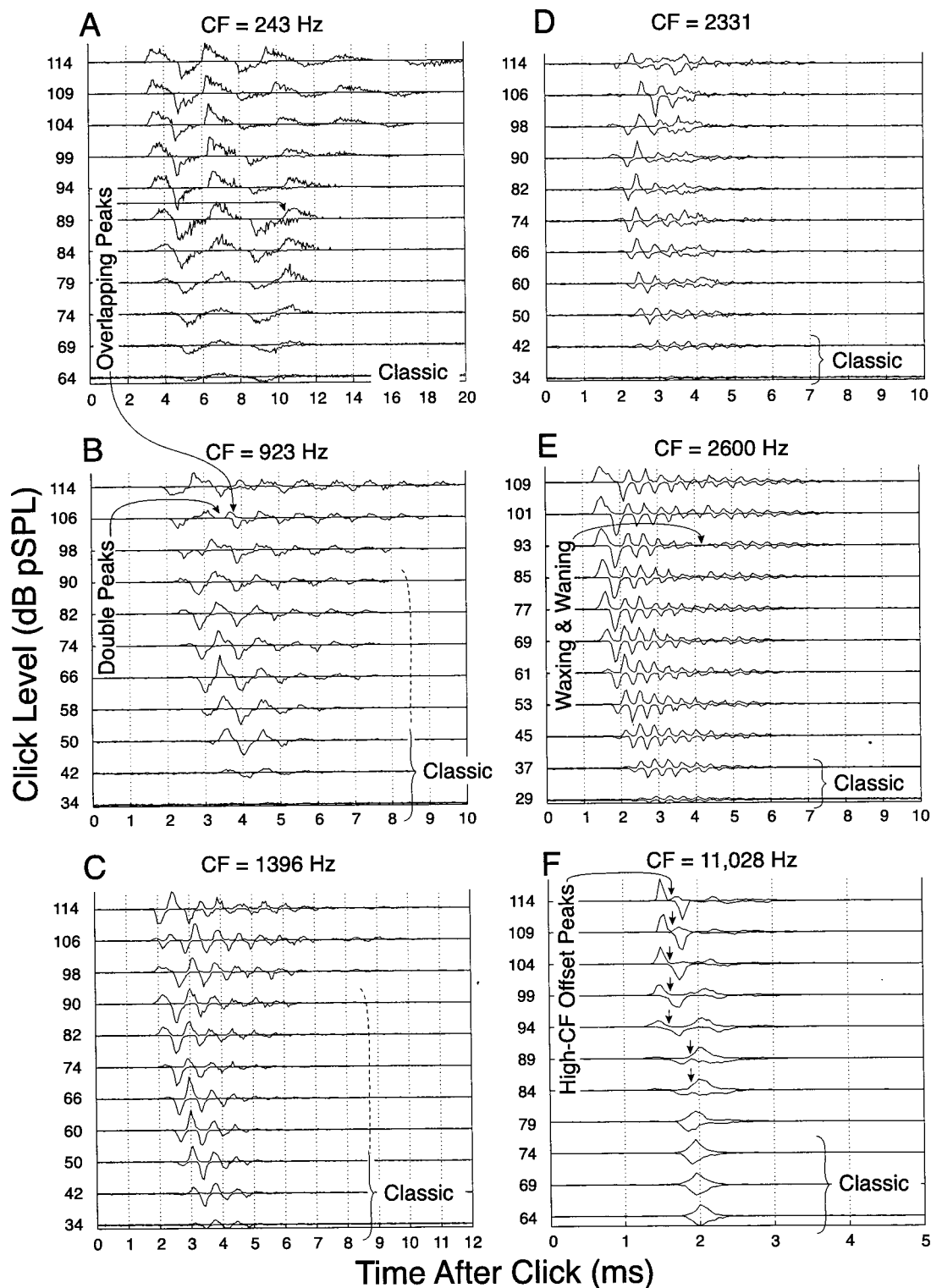


FIG. 3. Compound PST (cPST) histograms from click-level series from six auditory-nerve fibers. Each panel shows cPSTs arranged in ascending click level as indicated by the numbers at left. In each panel all cPSTs are on the same scale. Labels indicate levels with “classic” cPSTs and the anomalies pointed out in the text (an arrow indicates the latency noted for each anomaly). The cat-unit number, characteristic frequency (CF) in kHz, threshold to tones at CF in dB SPL, and spontaneous rates (SR) are:

Panel	Unit	CF	Threshold	SR
A	26–12	243	25.6	81.4
B	26–23	923	15	41.6
C	26–7	1396	9.4	64.6
D	25–45	2331	14	82.7
E	34–17	2600	0.5	101.1
F	25–43	11028	18.3	11.3

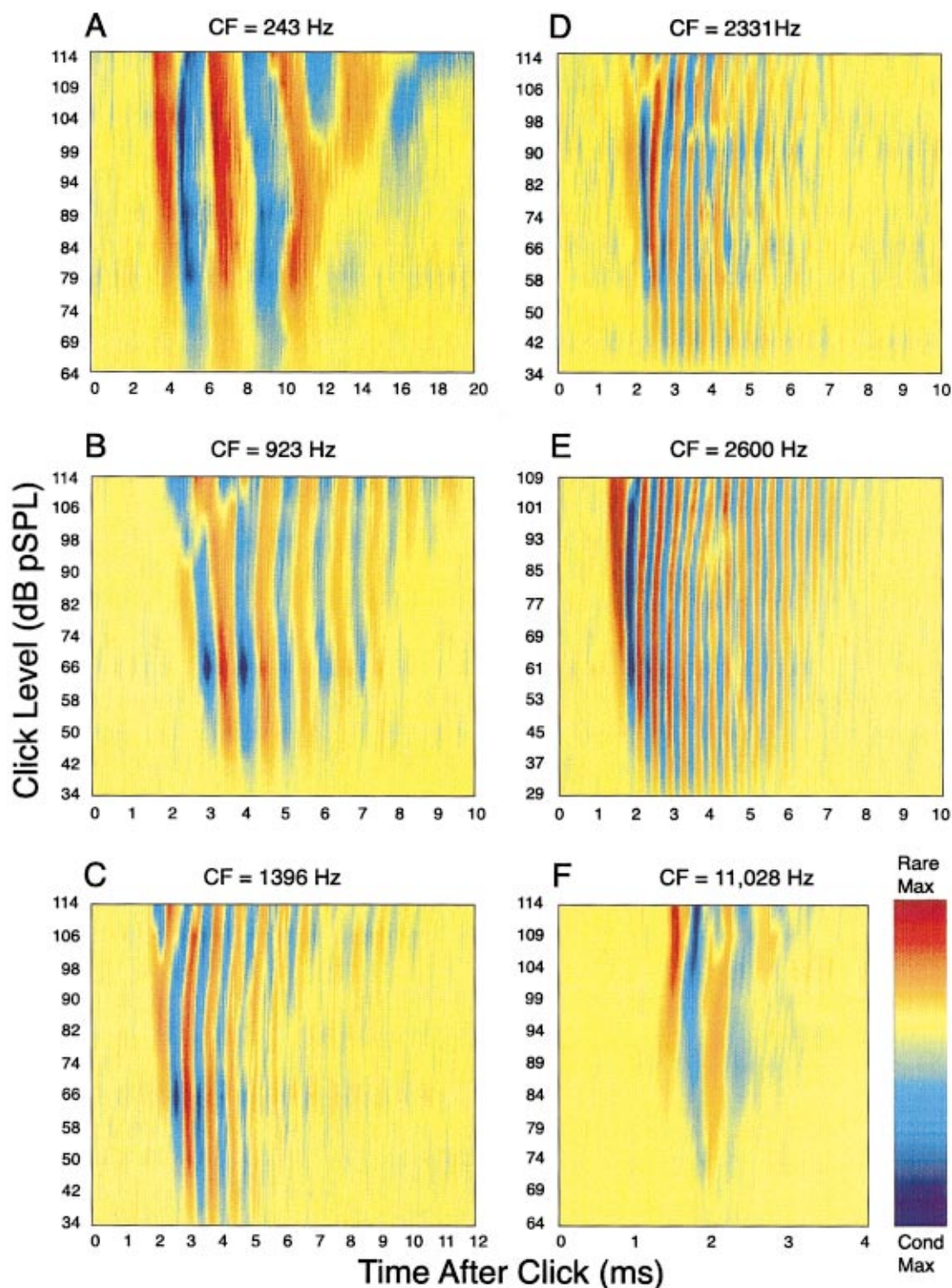


FIG. 4. Color-coded, interpolated plots of cPST histograms from click-level series from six auditory-nerve fibers. Same data as in Fig. 3. In forming color-coded plots, rates from rarefaction (positive values) and condensation (negative values) histograms were added to form a single "cPST rate." Each panel was separately normalized with yellow set to zero, the most positive value of the cPSTs coded as deep red, and the most negative value of the cPSTs coded as deep purple. The color-code bar in F applies to all panels. To efficiently use the color scale, the square root of the $|\text{cPST rate}|$ was used as the input to the color encoder.

peaks": rarefaction and condensation peaks overlap more than expected from normal responses at the fiber's CF (this applied only for $\text{CF} < 4$ kHz because above 4 kHz overlap is expected); the latency noted was the time of maximal overlap [e.g., Fig. 3(a) at 89 dB, 10.2 ms; Fig. 3(b) at 106 dB, 3.7 ms]. "High-CF offset peaks" ($\text{CF} > 4$ kHz): Rarefaction and

condensation peaks are not at the same time (simultaneous peaks are expected for a high-frequency (i.e., CF) oscillation followed by rectification and a low-pass filter so that responses are due mostly to the IHC dc receptor potential [e.g., Fig. 3(f) at ≥ 84 dB]; the latency noted was the midpoint of the first two opposite-sign peaks. "Waxing and waning":

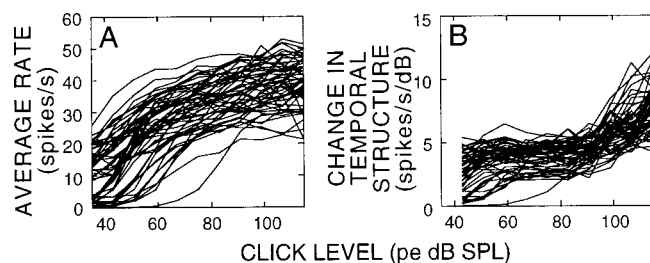


FIG. 5. Firing rate (A) and temporal change (B) versus click level in auditory-nerve-fiber responses. Temporal change was quantified by (1) at each bin of the cPST histogram, subtracting the response found at the next lower click level, (2) summing the absolute values of these differences for each click level, and (3) plotting the result versus click level (panel B). This temporal change measure is sensitive to changes in the timing of the peaks but ignores parts of the response that stay the same from one level to the next.

After the initial increase and decrease in the envelope of the response, the envelope increased and decreased at least once more; the latency noted was the middle of the first wane [e.g., Fig. 2(a) at 8.2 ms, and Fig. 3(e) at 93 dB, 4.2 ms].

A plot of the latencies of anomalies versus fiber CF is shown in Fig. 6. To remove latency variations across fibers due to cochlear delay, synaptic delay, and spike travel time to the recording electrode, for each fiber we subtracted the latency of the shortest-latency cPST peak seen at any click level on that fiber [these are shown in Fig. 6(a)]. As shown in Fig. 6(b), anomalies occurred over the whole range of CFs. All fibers with CFs > 4 kHz showed high-CF-offset-peaks anomalies. Most fibers with CFs < 4 kHz showed at least one of the other anomalies. The latencies of anomalies were distributed in two clusters, a short-latency group with latencies mostly within 2 ms of the shortest click latency, and a longer latency group [Fig. 6(b)]. Waxing and waning occurred mostly at longer latencies; the other anomalies were mostly at shorter latencies.

Peak “reversals” were determined from the cPST click-level series (such as those in Fig. 3) as latencies at which there was a peak from rarefaction clicks at one level and a peak from condensation clicks at another level. Sometimes the phase reversal was not complete. For the phase change to be counted as a “reversal,” the phase change had to be more than $3/8$ period at a given latency. The click level at which the peak reversal was said to occur was taken to be when the change was $1/4$ period from the low-level peak (i.e., when the cPST was zero). This is similar to the definition used by Wong *et al.* (1998) in distinguishing reversals for tones. The $1/4$ period definition of the reversal level was still used even if the phase change at a given latency was well over $1/2$ period [e.g., at 3.5 ms in Figs. 3(e) and 4(e)]. When necessary, we calculated the reversal level by interpolating between adjacent-level cPSTs. Reference to the “period” of the cPST makes sense for most latency ranges in which the cPST looks more or less like a sinusoidal oscillation. However, there were times when the high-level cPSTs were distinctly nonsinusoidal. In such cases, the criterion for a “reversal” was met if, at the latency of a cPST peak at low click levels, the response to the opposite polarity click at a high click level had a substantial amplitude (e.g., at least $1/2$ of nearby peak amplitudes). For many cPST level series, the

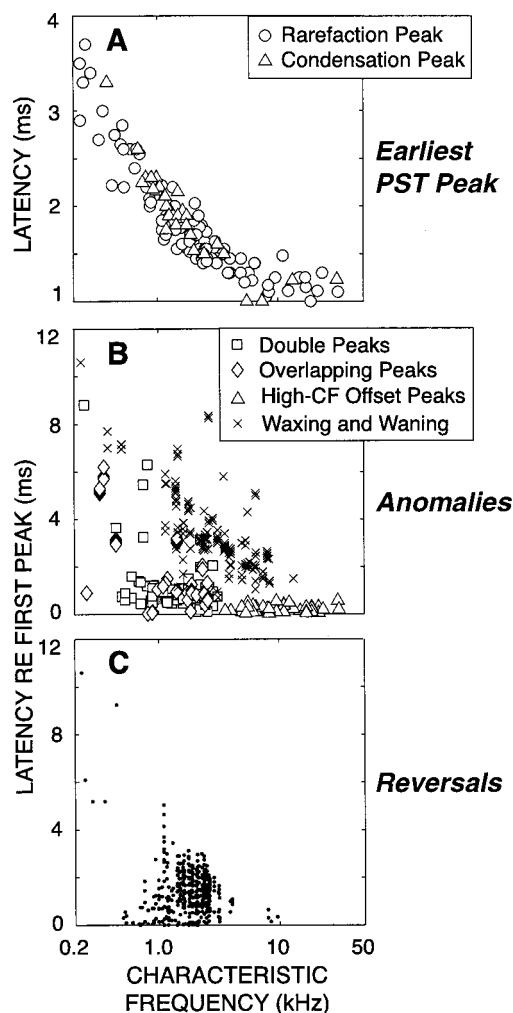


FIG. 6. Latencies of anomalies and reversals. (a) The latency of the earliest cPST peak across all click levels, coded by whether it occurred with rarefaction (○) or condensation (△) clicks. (b) The latency after the earliest PST peak of “anomalies” (nonclassic cPST regions, see text). Inset: key for the type of anomaly. (c) The latency after the earliest PST peak of peak reversals.

reversals appeared to extend to a certain latency and then stop [e.g., Figs. 3(e) and 4(e)]. For others it appeared that more reversals might have been distinguished if there had been more responses averaged so that the signal-to-noise ratio of the cPSTs would have been better.

A plot of the latencies of reversals versus fiber CF is shown in Fig. 6(c). Again, to make the latencies as comparable as possible across fibers, we have subtracted the latency of the shortest-latency cPST peak [Fig. 6(a)]. Figure 6(c) shows that most reversals were for fibers with CFs between 0.5 and 4 kHz with the greatest concentration near 2 kHz. Most reversals had latencies below 4 ms, but this must be interpreted with the caveat that decay of the response limited our ability to discern reversals at long latencies. Figure 6(c) shows that three fibers with reversals had CFs over 4 kHz [one fiber had reversals at two latencies and produced two points in Fig. 6(c)]. The reversals in these fibers were in the high-level, short-latency part of the response. Figure 7 shows an example of such a reversal.

For each fiber, the nonclassic “threshold” (the lowest

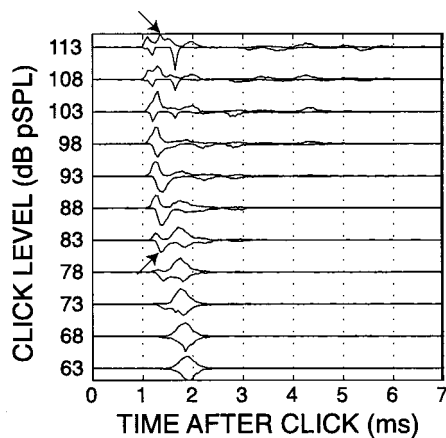


FIG. 7. A level series of compound-PST (cPST) histograms from a high-CF fiber that shows a reversal in an early, high-level cPST peak (arrows). Fiber 27–30: CF=8035 Hz, threshold=20.7 dB SPL, SR=40.5 spikes/s.

sound pressure level at which there was a nonclassic response, either an anomaly or a reversal) is plotted versus fiber CF in Fig. 8. The tone thresholds of these fibers are also plotted to provide perspective (click thresholds were not obtained but would be substantially above the tone thresholds). Figure 8 shows that nonclassic behavior extends down to click levels of 50 dB pSPL in some fibers (the lowest threshold points are waxing and waning), but for most fibers nonclassic behavior was first seen at 90 dB pSPL or above. Also noteworthy is that these nonclassic thresholds extend in a broad sweep across the whole range of CFs and that fibers with CFs above and below 4 kHz fit in approximately the

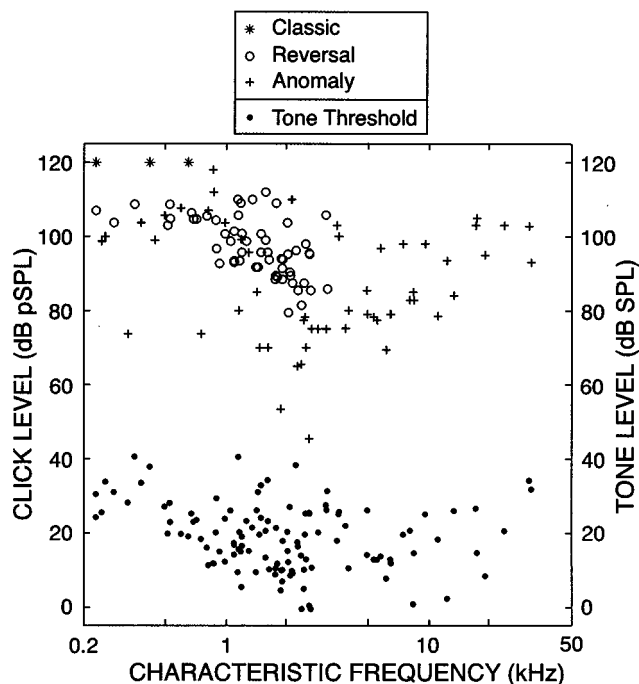


FIG. 8. The lowest sound pressure level at which there was a nonclassic response in a fiber, either an anomaly or a reversal, versus fiber CF. Points from three fibers that did not show any nonclassic responses ("classic" in the key) are shown at 120 dB and their CF. The solid points are tone thresholds for each fiber. The sound-level scales for clicks and tones are shown separately to emphasize that they are different measures.

Anomalies

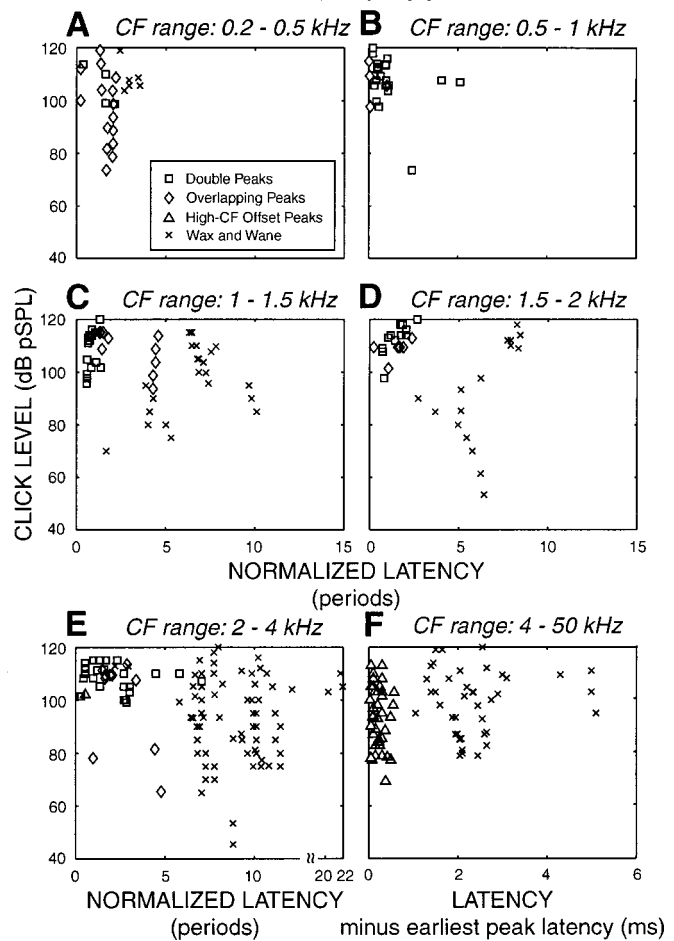


FIG. 9. Click level versus latency for "anomalies" in six CF bands. Latencies in panels (a)–(e) were normalized by subtracting the latency of the earliest PST peak and expressing the result in CF periods.

same distribution. Also included in Fig. 8 (plotted at their CF and at 120 dB pSPL) are the three fibers (out of 121) for which no nonclassic response was found at the click levels used.

B. Sound pressure level versus latency of nonclassic responses

A particularly revealing way to look at the data is to plot the sound pressure level of an anomaly or reversal versus its normalized latency (Figs. 9 and 10). For fibers with CFs < 4 kHz, the clearest patterns are produced if the latency of the first peak is subtracted and the remaining latency is expressed in CF periods. The result of this normalization is that the latency measure shows in which cycle of the response the anomaly or reversal occurs. However, for fibers with CFs > 4 kHz, expressing the latency in CF periods doesn't produce any discernible pattern so the latencies of these fibers are given in milliseconds.

Figures 9 and 10 show that for mid-CF fibers, nonclassic responses occurred primarily in two separate regions. The region that has anomalies and reversals over the largest range of CFs is a very short-latency, high-level region (Figs. 9 and 10, the region at latencies less than two periods and click levels over 95 dB pSPL). Level series with nonclassic behav-

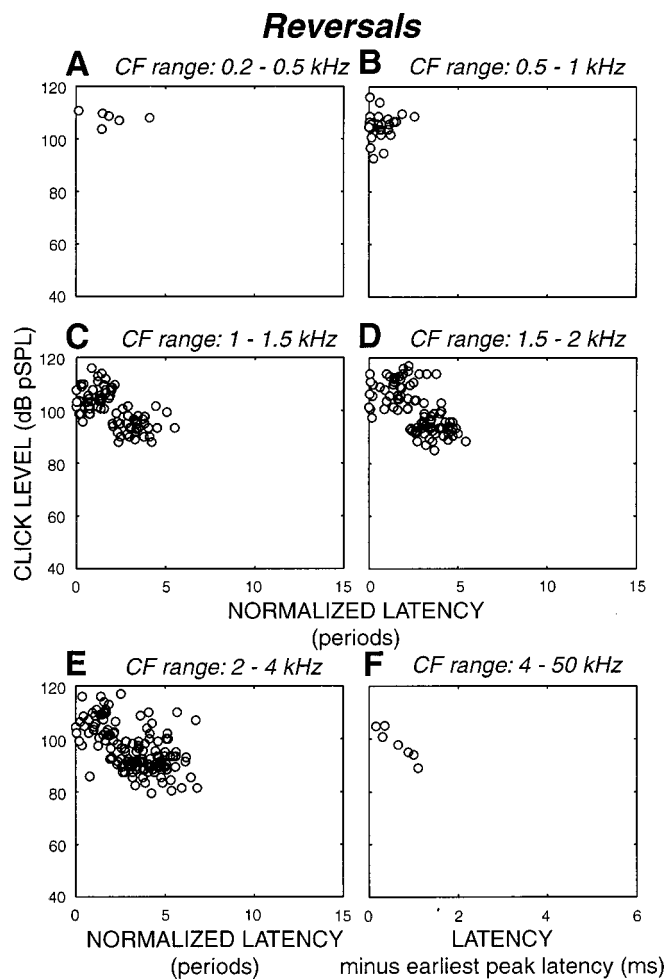


FIG. 10. Click level versus latency for "reversals" of cPST peaks, in six CF bands. Latencies in panels (a)–(e) were normalized by subtracting the latency of the earliest PST peak and expressing the result in CF periods.

ior in this region are shown in Figs. 3 and 4, panels (b)–(d). A particularly clear pattern is shown by the fiber in panel (b). In this fiber the earliest peak reverses at about 95 dB pSPL and subsequent peaks reverse at successively higher click levels. There were, however, other patterns of level versus latency for reversals in this region.

A second region of anomalies and reversals occurred at longer latencies and lower levels [Figs. 9 and 10, (c)–(e)]. Fibers with CFs of 1–2 kHz had anomalies and reversals both in the high-level, short-latency region, and in a long-latency, moderate-level region [Figs. 9 and 10, (c) and (d)]. Fibers with 2–4 kHz CFs had nonclassic behavior in the same two regions, but a greater proportion of fibers had nonclassic behavior in the long-latency, moderate-sound-level region [Figs. 9(e), 10(e)]. Many of these fibers had click-level series in which the period of the response changed with click level over a 10–20 dB range and one cycle disappeared as click level increased [panel (e) in Figs. 3 and 4]. At click levels in the middle of this transition, the response often waxed and waned [e.g., Fig. 3(e), 85 and 93 dB pSPL] as if the response were due to the mixing of two excitatory drives of slightly different frequencies. At latencies before the first wane, peaks in the low-level cPSTs usually reversed at high levels. Such reversals are the predominant source of rever-

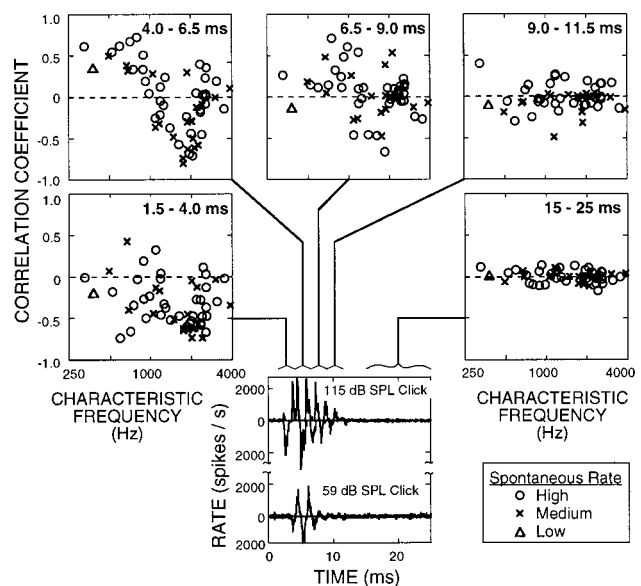


FIG. 11. The correlation of low- and high-level responses in five latency regions. Correlation coefficients were calculated for each fiber by comparing low-level (59 dB) and high-level (115 dB) cPSTs on a bin by bin basis across the latency range indicated in each panel.

sals in the long-latency, moderate-level region shown in Figs. 10(c)–(e). Note that waxing and waning is present even in fibers with CFs > 4 kHz [Fig. 9(f)]. In these fibers, waxing and waning is manifest by a multiburst response that is identical in rarefaction and condensation PSTs.

The primary nonclassic behavior of fibers with CFs > 4 kHz was the presence of high-CF offset peaks [panel (f) in Figs. 3 and 4]. As noted earlier, there were a few cases in which early nonclassic peaks reversed with click level. Figure 10(f) shows that these reversals all occurred within 2 ms of each fiber's shortest latency.

C. Low-level to high-level correlations

To provide an objective and sensitive measure of the extent to which fibers reversed the phase of their response, for each fiber we calculated the correlation of the response to low-level clicks with the response to high-level clicks. As noted earlier, it was difficult to determine with confidence whether individual peaks had reversed at long-latencies when the amplitudes of the peaks were small. Correlations overcome this by using data from many peaks. For the low click level, we chose 59 dB pSPL, which appeared to be low enough to be in the low-level region but high enough to have a strong response (see Fig. 5). For the high click level, we used the highest level. Noisiness of the data was still a problem if we did the correlation on a cycle by cycle basis, so we calculated the correlations over longer periods, 2.5 ms, skipping the first 1.5 ms where responses at the lowest click level were usually zero.

The correlation between cPSTs at low and high click levels is shown in Fig. 11. A fiber in which the responses at low and high levels were the same would have a correlation of +1, whereas one in which the responses were reversed would have a correlation of –1. Responses dominated by noise have zero correlation (e.g., at latencies of 15–25 ms

when the spikes were mostly spontaneous activity). In the shortest latency region (1.5–4 ms) almost all fibers showed negative correlations between their low and high level responses. This is primarily due to the presence of short-latency reversals in almost all fibers (Figs. 6 and 10). The 4–6.5 ms latency region is particularly interesting. Fibers with CFs less than about 1 kHz have positive correlations. They appear to have lost the short-latency reversal, and it has not been replaced by a longer-latency reversal. This is consistent with the picture of Fig. 10 which shows only high-level, short-latency reversals for fibers with CFs less than 1 kHz. In contrast, fibers with CFs in the 1–3 kHz range have negative correlations consistent with their responses being reversed from low to high click levels due to the long-latency, moderate-level reversals shown in Fig. 10(c)–(e). At longer latencies, this pattern subsides. Thus, Fig. 11 provides alternate and objective evidence for the patterns shown in Fig. 10.

III. DISCUSSION

Our data show that auditory-nerve-fiber responses to clicks display complex patterns that are clearly different from the classic view of click responses. With response patterns that seem so unusual, the question must be raised of whether they might have been produced by some sort of artifact. One possible source of artifact is the acoustic stimulus. Our acoustic source produces clicks that are relatively punctate in time (Fig. 1) but there is some ringing at 2–3 kHz (considerably more in the animal than in an artificial cavity). It seems unlikely, however, that this acoustic ringing is an important factor in producing the nonclassic features because (1) reversals are due to an inherently nonlinear mechanism and cannot be accounted for by any spectral or temporal properties of the clicks, (2) the latency after the first PST peak of anomalies and reversals depends on fiber CF (Figs. 6, 9, and 10) rather than on the temporal properties of the click, (3) the periods of the click responses at pSPLs above and below reversals are tied to fiber CF, and (4) our later cats (30–37) had increased damping in the acoustic assembly and reduced stimulus ringing, but both early and late cats had similar patterns of anomalies and reversals. Another potential source of artifact is acoustic trauma. In preliminary experiments we used a variety of click durations and found that clicks of 200–300 μ s duration produced acoustic trauma. However, the 90 μ s clicks used in the experiments reported here produced little or no acoustic trauma as judged by tone-pip audiograms and cochlear thresholds on single fibers. Furthermore, we noticed no differences in response patterns between fibers obtained early in an experiment and those obtained two days later in the same animal. In short, our experiments were done on intact ears, with good thresholds and adequately punctate clicks, so we conclude that the nonclassic responses we found represent the normal responses of the cat cochlea.

A. Comparisons with previous experimental data using clicks

There is clear evidence in earlier reports for each type of nonclassic response we have distinguished. First, consider

responses of low-CF fibers (CFs < 4 kHz). Although nonclassic response patterns were seldom mentioned in earlier publications, there is evidence in published figures for reversals and every type of anomaly. “Double peaks” are shown by Pfeiffer and Kim (1972, Fig. 6) and Versnel *et al.* [1992, Fig. 2(e)], “overlapping peaks” by Pfeiffer and Kim (1972, Figs. 4 and 5) and Versnel *et al.* [1992, Fig. 2(e)], “waxing and waning” by Pfeiffer and Kim (1972, Figs. 7 and 8), and “high-CF offset peaks” by Kiang *et al.* (1965), Versnel *et al.* (1992, 1997) and Schoonhoven *et al.* (1994). Evidence of reversals is shown by Kiang *et al.* [1965, Fig. 5.8 (at 3.7 ms) and Fig. 5.11 (at 3.5 and 9 ms)] and by Antoli-Candela and Kiang (1978, Fig. 10, the fibers with CFs of 1.7 and 2.3 kHz), and by Versnel *et al.* [1992, Fig. 2(e), at 5.0 ms].

Although most nonclassic patterns in low-CF fibers were not previously recognized, “high-CF-offset-peaks” anomalies in high-CF fibers have been both documented and explored. High-level clicks evoke short-latency responses in high-CF fibers that have different latencies for rarefaction and condensation clicks, while low-level clicks evoke longer-latency responses that have the same latency for rarefaction and condensation. An association between these click-response components and features of the tuning curve has been proposed (Schoonhoven *et al.*, 1994). In a two-drive model with filters that mimic tuning-curve tips and tails, the short-latency peaks are due to the excitation drive that produces tuning-curve tails and the long-latency peaks are due to the drive that produces tuning-curve tips (Schoonhoven *et al.* 1994; see also Versnel *et al.*, 1992, 1997). Finally, there has been no previous report of reversals of the short-latency click-response peaks (as in Fig. 7).

B. Comparisons with studies using tones

For high-CF fibers, if we consider the short-latency cPST peaks as being due to the tail excitation drive (Schoonhoven *et al.*, 1994), then short-latency click responses can be compared to tail-frequency tone responses. With low-frequency tones in cats, phase reversals at high sound levels have not been seen in high-CF fibers, although there is ample evidence for such reversals in low-CF fibers (Stankovic and Guinan, submitted; Gifford and Guinan, 1983; Liberman and Kiang, 1984; Kiang, 1990). In contrast, with low-frequency tones in the chinchilla, reversals are present in every high-CF fiber, sometimes more than one per fiber (Ruggero *et al.*, 1996b). Perhaps we have seen reversals in cat high-CF fibers with clicks but not tones because we are able to use higher sound levels with clicks than with tones and still avoid acoustic trauma. Perhaps the reversals we found in the tail-associated click-response peaks of high-CF fibers may be representative of the behavior of most cat fibers at high enough levels. If so, then cats and chinchillas would be similar.

For low-CF fibers, our nonclassic click responses are likely to be related to phase reversals produced by tip-frequency tones in the 80–100 dB SPL range (Kiang *et al.*, 1969; Kiang and Moxon, 1972; Gifford and Guinan, 1983; Liberman and Kiang, 1984; Kiang, 1990; Ruggero *et al.*, 1996b; also see Sokolich *et al.*, 1976; Sellick *et al.*, 1982). These phase reversals have been hypothesized to be pro-

duced by the addition at some stage of the cochlea of two excitation drives derived from the sound stimulus that are out of phase with each other and have different growth rates (Kiang, 1984, 1990). At low sound levels drive 1 is largest and determines the output phase, and at high sound levels drive 2 is largest and determines the output phase; the abrupt phase change occurs at the sound level where the drives are equal and cancel. Although it seems obvious that clicks should produce a corresponding response pattern, the exact pattern is not obvious because clicks do not have a phase in the same way as tones.

Insight into the corresponding click-response pattern can be gained from a two-drive model of tone-response phase reversals (Goldstein, 1990, 1991, 1995) that has been extended to clicks (Lin and Goldstein, 1995). The results with the two excitation drives 180° out of phase are most relevant.¹ At the lowest and highest click levels, the model click responses ring at one frequency, the resonant frequency of the dominant drive at that level. At low levels, actual click data ring at one frequency (the fiber CF), as in the model response. However, at high levels above a short-latency transition region, we have not seen actual click data that ring at only one frequency. This may be because we have not used high-enough click levels or because such one-frequency, high-level responses do not exist. At intermediate levels, the model click response is complex and shows cancellations of peaks that were present at lower click levels and peak reversals from low to high levels. According to the model, which peak reverses and is canceled, and to what extent, depends on the resonant frequencies, amplitudes, phases, delays and decay rates of the two drives, and can vary with click level (even though the model shows a cancellation and abrupt phase change at a single sound level for tones).

To see whether our actual click-response patterns might be accounted for by a two-drive model, we varied the parameters of the two-drive model of Lin and Goldstein (1995). For parameters in the range considered by Lin and Goldstein (1995), the two-drive model predicts click responses that show selective cancellation of peaks within the first two response periods, as in the short-latency reversals. For very different parameters, response patterns with reversals that resemble the long-latency reversals can be produced by the model. We have not, however, found any parameters that produce realistic looking patterns of reversals in both the short-latency and long-latency regions for a single set of parameters. Thus, the two-drive model in its present form does not seem to fit all of the reversal features in the data, and it is not clear whether any two-drive model can be made to fit them all.

C. Interpreting complex click responses: Low-CF fibers

If the two-drive model cannot produce the complex click-responses, what does? The only viable explanation appears to be that three (or more) excitation drives combine to produce the responses. Each region showing reversals appears to be due to the interaction of two out-of-phase excitation drives. For low-CF fibers, one region of reversals is the short-latency, high-level region present in a large fraction

of low-CF fibers. As suggested by the Lin and Goldstein (1995) model with parameters in the range they considered, reversals in the short-latency, high-level region may correspond to the phase reversals seen at high levels with tones (Kiang *et al.*, 1969; Kiang and Moxon, 1972; Gifford and Guinan, 1983; Liberman and Kiang, 1984; Kiang, 1990; Ruggero *et al.*, 1996b). Thus, the short-latency, high-level region appears to be due to two interacting excitation drives. The second region of reversals is the long-latency, moderate-level region. This region was most apparent in fibers with CFs 1–3 kHz (Figs. 3, 4, and 9–11) and appears to represent the transition between the response being driven at one frequency for low-level clicks and at another frequency for high-level clicks. The presence of waxing and waning of the response at click levels in the transition region argues for this transition being due to two interacting excitation drives with slightly different frequencies, as opposed to a single excitation drive that undergoes a slow change in its resonant frequency as click level increases. Thus, this long-latency, moderate-level region is also due to two interacting excitation drives. An important point is that some of the fibers that had long-latency reversals also had reversals in the short-latency region [panels (c) and (d) in Figs. 3 and 4]. Thus, for some fibers the observed responses must be produced by at least three excitation drives: (1) the CF drive, (2) a drive that interacts with the CF drive to produce the short-latency, high-level reversals, and (3) a drive that interacts with the CF drive at longer latencies to produce the waxing and waning and the longer-latency reversals.

There is previous evidence for three excitation drives interacting to produce the responses of auditory-nerve fibers. Low-CF auditory nerve fibers have tuning-curve tips that are lobed as if they resulted from two interacting excitation drives (Liberman and Kiang, 1978). In addition, phase-versus-frequency plots from low-CF, auditory-nerve fibers show two (or more) frequency regions with approximately straight-line slopes (Pfeiffer and Molnar, 1970; also see Allen, 1983) as if the excitation to the auditory-nerve fiber were due to the sum of two (or more) excitation drives that have different group delays. Furthermore, the frequencies at which the group delays change are the same frequencies as the seams between the lobes in the tuning curves (Kiang, 1984). Thus, the tuning-curve lobes appear to be due to the interaction of two excitation drives that have different group delays. Fibers in this same frequency region show phase reversals for high-level tones (Liberman and Kiang, 1984). It is likely that the excitation drive that produces the high-level response is different from both of the excitation drives that interact to produce the lobed structure of the tuning curves. One line of evidence for this is that the application of furosemide abolishes tone component 1 (i.e., it removes both of the excitation drives that interact to produce the low-level response) but does not remove tone component 2 (the high-level response) (Sewell, 1984a, 1984b). Taken together, the click and tone data complement each other and make a powerful case that three excitation drives interact to produce the drive to low-CF auditory-nerve fibers. Furthermore, the data suggest that two of these excitation drives are not just high-level phenomena but are present near threshold.

A tentative correspondence between the click and tone excitation drives can be suggested. Presumably, the excitation drive that produces the low-level click responses with periods near CF (drive CF), is the same as the excitation drive that produces responses to low-level tones near CF. Based on our model results, we hypothesize that the excitation drive that produces the response at click levels above the short-latency, high-level reversals (drive H), also produces the high-level tone component 2. This leaves the excitation drive that interacts with drive CF to produce long-latency, moderate-level transitions in the click responses and presumably the waxing and waning of the click responses (drive L), to be identified with the excitation drive that interacts with the drive CF to produce sidelobes in low-frequency tuning curves. One caveat is that the manifestations of these drives in tones and clicks occur over somewhat different CF ranges so it is not clear whether the same excitation drives are involved throughout the whole CF range.

D. Interpreting complex click responses: High-CF fibers

Our working hypothesis is that responses of high-CF fibers are due to separate tip and tail excitation drives, and that each of these is made up of two, or more, drives. The modeling work of Schoonhoven *et al.* (1994; see also Versnel *et al.*, 1992, 1997) provides an excellent justification that the “tip drive” produces the low-level, long-latency click responses, and the “tail drive” produces the high-level, short-latency (i.e., high-CF-offset-peaks) click responses. It might be said that high-CF-offset-peaks and tail responses are due to “glides” (de Boer and Nuttall, 1997; Recio *et al.*, 1998; Carney *et al.*, 1999). However, the term “glides” is just a name for the phenomena without a putative mechanism for the production of the “glide.” We hypothesize that glides are due to sequential activation of multiple excitation drives. For upward “glides,” low-frequency drive(s) are activated first and then near-CF drives; for downward “glides” higher-than-CF drives are activated before CF drives. Thus our working hypothesis is not an alternative hypothesis to a “glides” hypothesis, but is a hypothesis for the origin of “glides.”

Evidence that the tip-drive of high-CF fibers is composed of two interacting excitation drives comes from waxing and waning in basilar-membrane and auditory-nerve-fiber data. Near the 8–10 kHz place in the chinchilla, basilar-membrane responses to low-level clicks sometimes showed a waxing and waning that appears to be due to the beating of two excitation drives that have slightly different resonant frequencies (Ruggero *et al.*, 1996a; Recio *et al.*, 1998). These basilar-membrane-motion data look very similar to click-response cPSTs from auditory-nerve fibers with CFs 2–4 kHz that show waxing and waning. We also found evidence for waxing and waning in the envelopes of responses from fibers with CFs > 4 kHz [Fig. 9(f)]. Thus, both basilar-membrane and auditory-nerve data support the hypothesis that responses in the tip region of high-CF tuning curves are due to the interaction of two excitation drives with slightly different resonant frequencies.

Although not compelling, there is evidence that tail-associated responses are also due to the interaction of two or more excitation drives. First, in response to tail-frequency tones, some auditory-nerve-fibers show rate-vs-level functions with two regions separated by a plateau, with the low-level region inhibited more by medial-efferent stimulation than the high-level region (see Fig. 6 of Stankovic and Guinan, 1999). One explanation is that the response below the plateau is due to an excitation drive that is strongly affected by efferent stimulation whereas the response above the plateau is due to an excitation drive that is little affected by efferent stimulation. Second, chinchilla high-CF auditory-nerve-fibers stimulated with low-frequency tones show one or two abrupt phase changes which suggests that two, or perhaps three, excitation drives interact to produce the response (Ruggero *et al.*, 1996b). Finally, our data show that for some high-CF fibers, a part of the high-level, short-latency click response reverses at very high click levels. This suggests that the tail drive is composed of two excitation drives that can be out-of-phase and have different growth rates. Note that the existence of two tail-frequency drives would mean that at least four separate drives are required to account for the data from high-CF fibers.

In summary, existing data support the hypothesis that the drive that excites both low-CF and high-CF auditory-nerve fibers is obtained from the combination, at some stage in the cochlea, of at least three excitation drives derived from the acoustic stimulus.

E. The origin of multiple excitation drives

What do the multiple excitation drives correspond to physically in the cochlea? One view has been that the excitation drives represent mechanical (or possibly electrical) drives that are carried over separate physical paths in the cochlea and add at some point, leading to the excitation of auditory-nerve fibers (e.g., Kiang, 1990; Liberman and Kiang, 1984; Gifford and Guinan, 1983). Consider first the possibility that one of the excitation drives is electrical, e.g., CM acting extracellularly on IHCs to change the IHC transmembrane voltage and the release of transmitter (Sellick *et al.*, 1982; Russell and Sellick, 1983; Ruggero and Rich, 1983; Ruggero *et al.*, 1986; Guinan and Gifford, 1988; Cody and Mountain, 1989; Cheatham and Dallos, 1989, 1998). However, CM is much smaller than IHC receptor potentials except for very low-frequency sounds in the extreme tails of tuning curves (Russell and Sellick, 1983; Cody and Mountain, 1989; Mountain, 1989). While CM may act as a minor excitation drive, it is unlikely that it becomes the dominant drive at any latency or click level, so CM is unlikely to account for any of the excitation drives observed here.

If an electrical path is ruled out, the separate-paths hypothesis requires a physically separate way for each excitation drive to bend IHC stereocilia. Movements of the reticular lamina, the tectorial membrane, and fluid flow through the reticular-lamina/tectorial-membrane space (caused, for instance, by squeezing the inner sulcus) are potential drives for bending IHC stereocilia (see Freeman and Weiss, 1990a, 1990b; de Boer, 1993; Dallos *et al.*, 1996). Thus, three excitation drives might be accommodated by the separate-paths

hypothesis. However, the presence of four drives in high-CF fibers would be a problem for this hypothesis.

A more comprehensive way of interpreting the origin of multiple excitation drives is that the structures of the cochlear partition (basilar membrane, organ of Corti and tectorial membrane) move by a combination of vibrational modes (Zhang *et al.*, 1997; Mountain, 1998; Mountain and Cody, 1999). In the simplest case, one excitation drive corresponds to one mode of vibration. With this view, there is no need for physically separate paths to provide separate excitation drives with separate resonant properties; instead multiple excitation drives involve the same cochlear structures vibrating in multiple ways. The separate-paths hypothesis can be considered to be an extreme case of cochlear motion in modes, one in which the motion for each mode (or most of the motion) is carried by different structures. Another view, that the tectorial membrane provides a resonance and perhaps an excitation drive (Zwislocki, 1980; Allen and Neely, 1992), can also be considered as a subset of the hypothesis that the structures of the cochlear partition vibrate in modes.

Evidence that vibration of cochlear-partition structures is a mixture of multiple vibrational modes comes from a variety of mechanical measurements. In live guinea pigs, basilar-membrane motion varies as a function of radial location in a pattern that is not compatible with a simple bowing of the basilar membrane (Xue *et al.*, 1993; Nilsen and Russell, 1999; Nuttall *et al.*, 1999; but also see Cooper, 1999). Evidence for multiple resonances in cochlear structures comes from measurements of Reissner's membrane, tectorial-membrane and reticular-lamina motion which show multiple group delays and multiple peaks in plots of amplitude and phase versus frequency (Cooper and Rhode, 1995; Gummer *et al.*, 1996; Khanna and Hao, 1999). Evidence for multiple resonances has also been found in the time dimension; basilar-membrane motion in response to clicks at the 8–10 kHz place in the chinchilla reverses in phase over time as if it were due to two interacting resonances (Recio *et al.*, 1998). Finally, direct observations in excised cochleas have found vibration of the organ of Corti in spatially complex patterns that change with the stimulus frequency, i.e., exactly what is expected for cochlear vibration in resonant modes (Karavitaki and Mountain, 1998; Richter and Dallos, 1998).

Taken together, the above observations provide strong evidence that the motion of cochlear structures is by vibration in modes, but they do not show whether more than one mode is involved in determining the normal cochlear output, i.e., the firing of auditory-nerve fibers. Our data provide clear evidence that three or more drives with different resonant properties (i.e., three modes) have important roles in shaping the output of the intact, normal cochlea. We believe that our auditory-nerve observations, together with the previous measurements of cochlear motions, bring to a critical mass the evidence in favor of the hypothesis that multiple resonant modes of cochlear motion play important roles in the normal operation of the cochlea.

F. Waxing and waning: from reflections or from beating?

Waxing and waning are sometimes seen in otoacoustic emissions (OAEs) from clicks (Kemp, 1978), which raises the question of whether the waxing and waning in our click responses might be due to multiply reflected traveling waves. With this hypothesis, (1) energy from the click proceeds apically in a basilar-membrane traveling wave to produce the initial response, and (2) at each frequency some of this energy is reflected by coherent reflection at its characteristic place, travels basally, is reflected again at the stapes, and returns to its characteristic place where it influences the continuing click response. However, it seems unlikely that multiply reflected waves could produce the systematic transitions we see in our data. In particular, consider the fiber of panel (e) of Figs. 3 and 4, in which 99 dB clicks produced waxing and waning, lower-level clicks produced responses with ringing at CF, and higher-level clicks produced responses with ringing at a frequency slightly below CF. If multiply reflected waves produced the waxing and waning at 99 dB, it is difficult to see why multiply reflected waves would not also produce substantial disturbances in the responses at lower click levels, or how these multiple reflections might produce the change in response frequency from low to high levels. It seems much more economical to explain these data as we already have, as due to a transition at 99 dB between two excitation drives with different resonant frequencies, with the waxing and waning produced by beating of responses at these two frequencies. Nonetheless, we do not rule out multiply reflected waves as having some influence on our click responses. Basilar-membrane round-trip travel times in the cat (e.g., Shera and Guinan, 1999) are well within the latencies of the click-response cPSTs (Figs. 3, 4, and 6). Since the relative amplitudes of click-evoked OAEs is largest at low levels, we would expect that such influences would be greatest at low click levels.

If waxing and waning is due to the beating of two excitatory drives, then it might be possible to use the wax and wane period to determine the beat frequency and therefore the frequency difference between the two excitatory drives. However, in most cases waveforms with "waxing and waning" show only one clear wane (as in Fig. 2), not a series of peaks and dips, so the period of the beating is not obvious. If we assume that the click excites the two drives in phase and they first become 180° out of phase at the first wane, then we can calculate the frequency ratios of the two excitatory drives from the times of the first wane [\times 's in Figs. 6(b) and 9]. This gives us frequency ratios that range from approximately 0.02 to 0.2 (i.e., about 0.03 to 0.3 octave). However, these figures should be interpreted with caution because it is not at all clear that our assumptions are correct.

G. Implications of multiple excitation drives and resonant modes

With the working hypothesis that multiple resonances that correspond to vibrational modes play important roles in the operation of the cochlea, an obvious question is what vibration pattern corresponds to each excitation drive? At

present we cannot answer this question, but existing data suggest that there are substantial differences between the vibration patterns of different excitation drives. The two excitation drives that interact to produce waxing and waning and the long-latency reversals in auditory-nerve-fiber click responses (drives CF and L) appear to be present in basilar membrane motion, judging from the data of Recio *et al.* (1998). In contrast, measurements of basilar-membrane responses to tones in good preparations do not show abrupt phase changes and dips, even in animals such as the chinchilla, where auditory-nerve-fiber responses show these features [e.g., Ruggero *et al.*, 1996b; however, Patuzzi *et al.* (1984) reported a basilar-membrane-motion reversal in a damaged preparation]. Nonetheless, many recordings of inner-hair-cell (IHC) receptor potentials and some recordings of OHC receptor potentials show abrupt phase changes and dips in response amplitude (Dallos and Cheatham, 1989; Cody and Mountain, 1989; Cheatham and Dallos, 1998). It appears that the vibrational mode responsible for drive H, the excitation drive that causes the high-level phase reversals, is large at the top of the organ of Corti (at IHC and OHC stereocilia) but is normally small at the basilar membrane. Thus, the vibration pattern for drive H appears to be substantially different from the vibration patterns for drives CF and L.

The working hypothesis that the excitation drives correspond to cochlear resonant modes sheds new light on the relationship of tip and tail components to cochlear active processes. Tip and tail components have been thought to reflect basically the same coupling between basilar-membrane motion and the bending of IHC stereocilia, the difference being that basilar-membrane motion was amplified by active OHC processes to produce the tip, but not amplified in producing the tail. As we have already argued (Secs. III C and III D), tip and tail components are likely to be due to separate excitation drives with each drive corresponding to one or more vibrational modes. With this view, active processes (i.e., OHC motility) might affect tail as well as tip modes. The finding of efferent inhibition in part of the tail supports this hypothesis (Stankovic and Guinan, 1999). An economical hypothesis is that medial-efferent fast inhibition affects only one process, a process that reduces OHC fast motility, which then (1) reduces cochlear amplification in the tip thereby producing the well-known efferent inhibition in the tip (Guinan, 1996) and (2) reduces one of two tail-frequency modes, a mode that has a broad but higher-frequency distribution than the overall tail, thereby accounting for the frequency offset of efferent tail-frequency inhibition and the selective inhibition below a plateau in tail-frequency tone-level functions (Stankovic and Guinan, 1999).

An important implication of our working hypothesis is that standard cochlear models have serious shortcomings. Cochlear models with a single CF resonance do not adequately represent the two excitation drives that are present at low click levels in both low and high-CF auditory-nerve fibers, and they do not represent separate tail-frequency resonances. Models that incorporate a second resonance (usually attributed to the tectorial membrane) may do better, but they

still fall short. The standard assumption of cochlear macro-mechanical models, that motion of the fluids surrounding the cochlear partition provides energy flow along the length of the basilar membrane, is unchanged by the presence of vibrational modes within the cochlear partition. However, the coupling of this energy to micromechanical processes (i.e., to multiple micromechanical vibrational modes) is likely to be more complex than the previously accepted view that energy is coupled through a simple bowing of the basilar membrane. The presence of multiple vibrational modes that vary with level and perhaps CF suggests that there are distortions of cochlear cross sections with different squeezing of the fluid spaces within the cochlear partition. Squeezing of these fluid spaces may provide an additional mechanism for carrying energy along the length of the cochlea (de Boer, 1993; Hubbard *et al.*, 1999). Intra-partition and extra-partition longitudinal fluid couplings may have different propagation velocities and account for the different group delays seen in low-CF auditory-nerve fibers and in mechanical measurements (Pfeiffer and Molnar, 1970; Cooper and Rhode, 1995; Gummer *et al.*, 1996; Khanna and Hao, 1999).

Our data and working hypothesis bring many additional questions into focus. What are the resonant frequencies, phases, group delays, and damping of each excitation drive? Arguably, the resonant properties of the excitation drives are more important than their vibration patterns because excitation of these resonances produces the output of the cochlea no matter what their vibration patterns are (of course, knowing the vibration patterns should provide insight into the mechanisms that produce and influence the resonances). What controls the partitioning of energy between excitation drives? The change with sound level in which excitation drive is dominant shows that cochlear nonlinearities have a strong influence on partitioning energy among the resonant modes. Which nonlinearities are involved in partitioning energy among the modes? How does the presence of multiple vibrational modes affect calculations of cochlear-partition effective mass in cochlear models? How is the generation of otoacoustic emissions affected by the presence of multiple modes of cochlear partition motion? The new view that cochlear motion consists of a mixture of multiple vibrational modes, each producing a drive to auditory-nerve fibers, has tied together many previously disparate observations, but it has also led to a host of new questions.

IV. CONCLUSIONS

- (1) Our click data, and past data from tones, make a powerful case that at least three excitation drives determine the firing of auditory-nerve fibers.
- (2) Our results, along with recent measurements of cochlear micromechanical motions, bring to a critical mass the evidence in favor of the hypothesis that movements of the cochlear partition in multiple resonant modes play important roles in the operation of the intact, normal cochlea.

ACKNOWLEDGMENTS

We thank Dr. M. C. Liberman and Dr. C. A. Shera for comments on the manuscript. Supported by Research Grant No. 5 RO1 DC 00235 and training Grant No. T32 DC00038, from the National Institute of Deafness and Other Communication Disorders.

¹The model developed by Goldstein (1990, 1991, 1995) and Lin and Goldstein (1995) with two excitation drives at 180° phase difference was successful in incorporating into a single framework a wide range of difficult to reconcile observations, but was rejected because it predicted click responses that appeared to be too complex considering existing reports. Instead, to match both tone-response-reversal and classic-click data, it was hypothesized that the two excitation drives changed their phase relationship in the transition from transient response to steady-state response (Lin and Goldstein, 1995). Since our data show the nonclassic click-response patterns predicted by the model with two drives at 180° phase, the hypothesized phase adaptation is no longer needed.

Allen, J. B. (1983). "Magnitude and phase-frequency response to single tones in the auditory nerve," *J. Acoust. Soc. Am.* **73**, 2071–2092.

Allen, J. B., and Neely, S. T. (1992). "Micromechanical models of the cochlea," *Phys. Today* July 40–47.

Antoli-Candela, F., Jr., and Kiang, N. Y. S. (1978). "Unit activity underlying the N_1 potential," in *Evoked Electrical Activity in the Auditory Nervous System*, edited by R. F. Naunton and C. Fernández (Academic, New York), pp. 165–191.

Carney, L. H., McDuffy, M. J., and Shekhter, I. (1999). "Frequency glides in the impulse responses of auditory-nerve fibers," *J. Acoust. Soc. Am.* **105**, 2384–2391.

Cheatham, M. A., and Dallos, P. (1989). "Two-tone suppression in inner hair cell responses," *Hearing Res.* **40**, 187–196.

Cheatham, M. A., and Dallos, P. (1998). "The level dependence of response phase: Observations from cochlear hair cells," *J. Acoust. Soc. Am.* **104**, 356–369.

Cody, A. R., and Mountain, D. C. (1989). "Low-frequency responses of inner hair cells: Evidence for a mechanical origin of peak splitting," *Hearing Res.* **41**, 89–99.

Cooper, N. P. (1999). "Radial variation in the vibrations of the cochlear partition," in *Recent Developments in Auditory Mechanics*, edited by H. Wada and T. Takasaka (World Scientific, Singapore).

Cooper, N. P., and Rhode, W. S. (1995). "Nonlinear mechanics at the apex of the guinea-pig cochlea," *Hearing Res.* **82**, 225–243.

Dallos, P. J., Popper, A. N., and Fay, R. R. (1996). *The Cochlea* (Springer-Verlag, New York).

Dallos, P., and Cheatham, M. A. (1989). "Nonlinearities in cochlear receptor potentials and their origins," *J. Acoust. Soc. Am.* **86**, 1790–1798.

de Boer, E. (1993). "The sulcus connection. On a mode of participation of outer hair cells in cochlear mechanics," *J. Acoust. Soc. Am.* **93**, 2845–2859.

de Boer, E., and Nuttall, A. L. (1997). "The mechanical waveform of the basilar membrane. I. Frequency modulations ('glides') in impulse responses and cross-correlation functions," *J. Acoust. Soc. Am.* **101**, 3583–3592.

Freeman, D. M., and Weiss, T. F. (1990a). "Hydrodynamic forces on hair bundles at high frequencies," *Hearing Res.* **48**, 31–36.

Freeman, D. M., and Weiss, T. F. (1990b). "Hydrodynamic analysis of a two-dimensional model for micromechanical resonance of free-standing hair bundles," *Hearing Res.* **48**, 37–68.

Gifford, M. L., and Guinan, Jr., J. J. (1983). "Effects of crossed-olivocochlear-bundle stimulation on cat auditory nerve fiber responses to tones," *J. Acoust. Soc. Am.* **74**, 115–123.

Goldstein, J. L. (1990). "Modeling rapid waveform compression on the basilar membrane as multiple-bandpass-nonlinearity filtering," *Hearing Res.* **49**, 39–60.

Goldstein, J. L. (1991). "Modeling the nonlinear cochlear mechanical basis of psychophysical tuning," *J. Acoust. Soc. Am.* **90**, 2267–2268.

Goldstein, J. L. (1995). "Relations among compression, suppression and combination tones in mechanical responses of the basilar membrane: Data and MBPNL model," *Hearing Res.* **89**, 52–68.

Gray, P. R. (1967). "Conditional probability analyses of the spike activity of single neurons," *Biophys. J.* **7**, 759–777.

Guinan, Jr., J. J. (1996). "The physiology of olivocochlear efferents," in *The Cochlea*, edited by P. J. Dallos, A. N. Popper, and R. R. Fay (Springer-Verlag, New York), pp. 435–502.

Guinan, Jr., J. J., and Gifford, M. L. (1988). "Effects of electrical stimulation of efferent olivocochlear neurons on cat auditory-nerve fibers. II. Spontaneous rate," *Hearing Res.* **33**, 115–128.

Guinan, Jr., J. J., Guinan, S. S., and Norris, B. E. (1972). "Single auditory units in the superior olivary complex. I: Responses to sounds and classifications based on physiological properties," *Int. J. Neurosci.* **4**, 101–120.

Gummer, A. W., Hemmert, W., and Zenner, H. P. (1996). "Resonant tectorial membrane motion in the inner ear: Its crucial role in frequency tuning," *Proc. Natl. Acad. Sci. USA* **93**, 8727–8732.

Hubbard, A. E., Shatz, L., Yang, Z., and Mountain, D. C. (1999). "Multi-mode cochlear models," in *Recent Developments in Auditory Mechanics*, edited by H. Wada and T. Takasaka (World Scientific, Singapore).

Karavita, K. D., and Mountain, D. C. (1998). "Estimating electrically-evoked motion in the excised cochlea using video microscopy data," *Assoc. Res. Otolaryngol. Abstr.* **21**, 180.

Kemp, D. T. (1978). "Stimulated acoustic emissions from within the human auditory system," *J. Acoust. Soc. Am.* **64**, 1386–1391.

Khanna, S. M., and Hao, L. F. (1999). "Reticular lamina vibrations in the apical turn of a living guinea pig cochlea," *Hearing Res.* **132**, 15–33.

Kiang, N. Y. S., Watanabe, T., Thomas, E. C., and Clark, L. F. (1965). *Discharge Patterns of Single Fibers in the Cat's Auditory Nerve* (MIT, Cambridge, MA).

Kiang, N. Y. S. (1984). "Peripheral neural processing of auditory information," in *Handbook of Physiology. The Nervous System III* (American Physiological Society, Bethesda, MD), pp. 639–674.

Kiang, N. Y. S. (1990). "Curious oddments of auditory-nerve studies," *Hearing Res.* **49**, 1–16.

Kiang, N. Y. S., Baer, T., Marr, E. M., and Demont, D. (1969). "Discharge rates of single auditory nerve fibers as functions of tone level," *J. Acoust. Soc. Am.* **46**, 106.

Kiang, N. Y. S., and Moxon, E. C. (1972). "Physiological considerations in artificial stimulation of the inner ear," *Ann. Otol. Rhinol. Laryngol.* **81**, 714–731.

Liberman, M. C. (1978). "Auditory-nerve response from cats raised in a low-noise chamber," *J. Acoust. Soc. Am.* **63**, 442–455.

Liberman, M. C. (1982). "Single-neuron labeling in the cat auditory nerve," *Science* **216**, 1239–1241.

Liberman, M. C., and Kiang, N. Y. S. (1978). "Acoustic trauma in cats. Cochlear pathology and auditory-nerve activity," *Acta Oto-Laryngol. Suppl.* **358**, 1–63.

Liberman, M. C., and Kiang, N. Y. S. (1984). "Single-neuron labeling and chronic cochlear pathology. IV. Stereocilia damage and alterations in rate- and phase-level functions," *Hearing Res.* **16**, 75–90.

Lin, T., and Goldstein, J. L. (1995). "Quantifying 2-factor phase relations in non-linear responses from low characteristic-frequency auditory-nerve fibers," *Hearing Res.* **90**, 126–138.

Møller, A. R., and Nilsson, H. G. (1979). "Inner ear impulse response and basilar membrane modeling," *Acustica* **41**, 258–262.

Mountain, D. C. (1989). "Measurement of low-frequency receptor potentials in inner hair cells: A theoretical analysis," *Hearing Res.* **41**, 101–106.

Mountain, D. C. (1998). "Modal analysis: A new paradigm for cochlear mechanics," *Assoc. Res. Otolaryngol. Abstr.* **21**, 64.

Mountain, D. C., and Cody, A. R. (1999). "Multiple modes of inner hair cell stimulation," *Hearing Res.* **132**, 1–14.

Nilsen, K. E., and Russell, I. J. (1999). "Timing of cochlear feedback: spatial and temporal representation of a tone across the basilar membrane," *Nat. Neurosci.* **2**, 642–648.

Nuttall, A. L., Guo, M., and Ren, T. (1999). "The radial pattern of basilar membrane motion evoked by electric stimulation of the cochlea," *Hearing Res.* **131**, 39–46.

Patuzzi, R., Johnstone, B. M., and Sellick, P. M. (1984). "The alteration of the vibration of the basilar membrane produced by loud sound," *Hearing Res.* **13**, 99–100.

Pfeiffer, R. R., and Molnar, C. E. (1970). "Cochlear nerve fiber discharge patterns: Relationship to the cochlear microphonic," *Science* **167**, 1614–1616.

Pfeiffer, R. R., and Kim, D. O. (1972). "Response patterns of single cochlear nerve fibers to click stimuli: Descriptions for cat," *J. Acoust. Soc. Am.* **52**, 1669–1677.

- Recio, A., Rich, N. C., Narayan, S. S., and Ruggero, M. A. (1998). "Basilar-membrane responses to clicks at the base of the chinchilla cochlea," *J. Acoust. Soc. Am.* **103**, 1972–1989.
- Richter, C. P., and Dallos, P. (1998). "Basilar membrane micro-mechanics measured in the gerbil inner ear," *Assoc. Res. Otolaryngol. Abstr.* **21**, 181.
- Ruggero, M. A., and Rich, N. C. (1983). "Chinchilla auditory-nerve responses to low-frequency tones," *J. Acoust. Soc. Am.* **73**, 2096–2018.
- Ruggero, M. A., Rich, N. C., and Recio, A. (1996a). "The effect of intense acoustic stimulation on basilar-membrane vibrations," *Aud. Neurosci.* **2**, 329–345.
- Ruggero, M. A., Rich, N. C., Shivapuja, B. G., and Temchin, A. N. (1996b). "Auditory-nerve responses to low-frequency tones: Intensity dependence," *Aud. Neurosci.* **2**, 159–185.
- Ruggero, M. A., Robles, L., and Rich, N. C. (1986). "Basilar membrane mechanics at the base of the chinchilla cochlea. II. Responses to low-frequency tones and relationship to microphonic and spike initiation in the VIII nerve," *J. Acoust. Soc. Am.* **80**, 1375–1383.
- Russell, I. J., and Sellick, P. M. (1983). "Low-frequency characteristics of intracellularly recorded receptor potentials in guinea-pig cochlear hair cells," *J. Physiol. (London)* **338**, 179–206.
- Ryugo, D. K. (1992). "The auditory nerve: Peripheral innervation, cell body morphology, and central projections," in *The Mammalian Auditory Pathway: Neuroanatomy*, edited by D. B. Webster, A. N. Popper, and R. R. Fay (Springer-Verlag, New York), Vol. 2, pp. 23–65.
- Schoonhoven, R., Keijzer, J., Versnel, H., and Prijs, V. F. (1994). "A dual filter model describing single-fiber responses to clicks in the normal and noise-damaged cochlea," *J. Acoust. Soc. Am.* **95**, 2104–2121.
- Sellick, P. M., Patuzzi, R., and Johnstone, B. M. (1982). "Modulation of responses of spiral ganglion cells in guinea pig cochlea by low frequency sound," *Hearing Res.* **7**, 199–221.
- Sewell, W. F. (1984a). "Furosemide selectively reduces one component in rate-level functions from auditory-nerve fibers," *Hearing Res.* **15**, 69–72.
- Sewell, W. F. (1984b). "The effects of furosemide on the endocochlear potential and auditory-nerve fiber tuning curves in cats," *Hearing Res.* **14**, 305–314.
- Shera, C. A., and Guinan, Jr., J. J. (1999). "Frequency dependence of stimulus-frequency-emission phase: Implications for cochlear mechanics," in *Recent Developments in Auditory Mechanics*, edited by H. Wada and T. Takasaka (World Scientific, Singapore).
- Sokolich, W. G., Hamernik, R. P., Zwislocki, J. J., and Schmiedt, R. A. (1976). "Inferred response polarities of cochlear hair cells," *J. Acoust. Soc. Am.* **59**, 963–974.
- Stankovic, K. M., and Guinan, Jr., J. J. (1999). "Medial efferent effects on auditory-nerve responses to tail-frequency tones I: Rate reduction," *J. Acoust. Soc. Am.* **106**, 857–869.
- Stankovic, K., M., and Guinan, J. J., Jr.. "Medial efferent effects on auditory-nerve responses to tail-frequency tones II: Alteration of phase," *J. Acoust. Soc. Am.* (submitted).
- Versnel, H., Schoonhoven, R., and Prijs, V. F. (1992). "Single-fibre and whole-nerve responses to clicks as a function of sound intensity in the guinea pig," *Hearing Res.* **59**, 138–156.
- Versnel, H., Prijs, V. F., and Schoonhoven, R. (1997). "Auditory-nerve fiber responses to clicks in guinea pigs with a damaged cochlea," *J. Acoust. Soc. Am.* **101**, 993–1009.
- Wong, J. C., Miller, R. L., Calhoun, B. M., Sachs, M. B., and Young, E. D. (1998). "Effects of high sound levels on responses to the vowel 'eh' in cat auditory nerve," *Hearing Res.* **123**, 61–77.
- Xue, S., Mountain, D. C., and Hubbard, A. E. (1993). "Direct measurement of electrically-evoked basilar membrane motion," in *Biophysics of Hair Cell Sensory Systems*, edited by H. Duifhuis, J. W. Horst, P. van Dijk, and S. M. Netten (World Scientific, Singapore), pp. 361–368.
- Zhang, L., Mountain, D. C., and Hubbard, A. E. (1997). "Shape and stiffness changes of the organ of Corti from base to apex cannot predict characteristic frequency changes: Are multiple modes the answer?," in *Diversity in Auditory Mechanics*, edited by E. R. Lewis, G. R. Long, R. F. Lyon, P. M. Narins, C. R. Steele, and E. Hecht-Poinar (World Scientific, Singapore), pp. 472–478.
- Zwislocki, J. J. (1980). "Five decades of research on cochlear mechanics," *J. Acoust. Soc. Am.* **67**, 1679–1685.

The case of the missing pitch templates: How harmonic templates emerge in the early auditory system

Shihab Shamma and David Klein

*Center for Auditory and Acoustics Research, Institute for Systems Research,
Electrical Engineering Department, University of Maryland, College Park, Maryland 20742*

(Received 9 April 1999; revised 6 October 1999; accepted 18 January 2000)

Periodicity pitch is the most salient and important of all pitch percepts. Psychoacoustical models of this percept have long postulated the existence of internalized harmonic templates against which incoming resolved spectra can be compared, and pitch determined according to the best matching templates [J. Goldstein, *J. Acoust. Soc. Am.* **54**, 1496–1516 (1973)]. However, it has been a mystery where and how such harmonic templates can come about. We present here a biologically plausible model for how such templates can form in the early stages of the auditory system. The model demonstrates that *any* broadband stimulus, including noise and random click trains, suffices for generating the templates, and that there is no need for any delay lines, oscillators, or other neural temporal structures. The model consists of two key stages: cochlear filtering followed by coincidence detection. The cochlear stage provides responses analogous to those recorded in the auditory nerve and cochlear nucleus. Specifically, it performs moderately sharp frequency analysis via a filterbank with tonotopically ordered center frequencies (CFs); the rectified and phase-locked filter responses are further enhanced temporally to resemble the synchronized responses of cells in the cochlear nucleus. The second stage is a matrix of coincidence detectors that compute the average pairwise instantaneous correlation (or product) between responses from all CFs across the channels. Model simulations show that for any broadband stimulus, a degree of high coincidence occurs among cochlear channels that are spaced precisely at harmonic intervals. Accumulating coincidences over time results in the formation of harmonic templates for all fundamental frequencies in the phase-locking frequency range. The model accounts for the critical role played by three subtle but important factors in cochlear function: the nonlinear transformations following the filtering stage, the rapid phase shifts of the traveling wave near its resonance, and the spectral resolution of the cochlear filters. Finally, we discuss the physiological correlates and location of such a process and its resulting templates. © 2000 Acoustical Society of America. [S0001-4966(00)04804-9]

PACS numbers: 43.66.Ba, 43.66.Jh [RVS]

INTRODUCTION AND BACKGROUND

More than any other auditory percept in the last century, pitch has been a potent source of inspiration and controversy in auditory research. Its importance stems from its role in perceiving the prosody of speech, melody of music, and in organizing the acoustic environment into different sources (Summerfield and Assmann, 1990; de Cheveigne *et al.*, 1995). It is generally appreciated that the term “pitch” refers to many distinct percepts (de Cheveigne, 1998; Moore, 1989): They include “spectral pitch” evoked by sinusoidal signals, “residue pitch” (Schouten, 1940; de Boer, 1976) associated with unresolved (high) harmonics, very slow click trains, or the envelope of amplitude modulated noise and sinusoids, and “periodicity pitch” (also known as virtual and missing fundamental pitch) evoked by low order, spectrally resolved harmonic tone complexes. The focus of this paper is on “periodicity pitch,” the pitch usually associated with musical intervals and melodies, and with speakers voices and speech prosody.

There is general agreement on the perceptual properties and acoustic parameters that give rise to periodicity pitch in humans (and presumably in other mammals and birds) (Langner, 1992; Moore, 1989; Plomp, 1976). For instance,

the most salient pitch is evoked by harmonically related tone complexes that are (at least partially) spectrally resolved; the pitch heard is normally that of the fundamental frequency of these harmonics regardless of the energy in that fundamental component; the pitch is roughly in the range 50–2000 Hz. The most effective (or dominant) harmonics are the low order harmonics (the 2nd–5th harmonics). The salience of the pitch increases proportional to the number of resolved harmonics. Multiple pitches are often perceived if there are only a few harmonics in the complex, or if the tones form an inharmonic sequence.

Numerous theories have been proposed to account for periodicity pitch percepts. Most successful among them are the so-called “spectral pitch theories,” best exemplified by the “central pattern recognition” theories (Goldstein, 1973b; Terhardt, 1974; Bilsen, 1977; Wightman, 1973), and the variations and implementations proposed since then (Duifhuis *et al.*, 1982; Cohen *et al.*, 1995). The two operations common to all are: (1) the pitch value is derived (centrally) from a spectral profile defined along the tonotopic axis of the cochlea (regardless of how this profile is computed); and (2) the input spectrum is compared to internally stored spectral templates, consisting of the harmonic series of all possible

fundamentals.¹ These theories have been enormously successful in explaining and predicting the pitches of complex tones, and consequently have provided the dominant view of pitch perception.

Spectral pitch theories, however, suffer two criticisms. The first is the lack thus far of convincing biological evidence for the existence of these templates or for how they might be generated. "Learning" the harmonic templates has usually been assumed to be a straightforward consequence of frequent exposure during early development to speech or natural sounds which tend to be rich in harmonic structure (Terhardt, 1974). However, there are several difficulties with this scenario. Infants are thought to be born with an innate sense of musical pitch (Clarkson and Rogers, 1995; Montgomery and Clarkson, 1997), presumably long before any serious exposure to speech (sounds in the womb are predominantly noiselike due to the heart and other internal organs). Another difficulty is that voiced speech usually has a relatively weaker fundamental component, raising the question of why learned templates consisting of prominent higher harmonics are perceived at (or are linked to) the pitch of the fundamental and not any other arbitrary frequency. A second criticism of spectral pitch theories is their inability to account for other weaker pitch percepts such as "residue pitch," which apparently operate in different parameter ranges, and may require different mechanisms.

To address these criticisms, alternative theories have been proposed to explain how the pitch percept might be computed without the need for stored harmonic templates. These theories can be described as "temporal" in that they postulate mechanisms that extract a pitch value from the temporal response in each auditory channel (independent of other channels), and then combine the results from across all channels to get the final estimate. As such, "temporal" theories unlike "spectral" theories, make no use of an ordered tonotopic axis, i.e., their computations are unaffected by a shuffling of the tonotopic axis (Lyon and Shamma, 1996).²

"Temporal" models vary enormously in the nature of the cues they utilize from each channel, e.g., first or higher order intervals (Evans, 1978; Cariani and Delgutte, 1996a, b; Rhode, 1995; Moore, 1986), autocorrelations of the responses (Slaney and Lyon, 1993; Licklider, 1951; Meddis and Hewitt, 1991; de Cheveigne, 1998), or synchronization measures and oscillators (Patterson and Holdsworth, 1991; Langner and Schreiner, 1988); they also differ in the mechanisms to measure them, e.g., delay lines and coincidence detectors, or intrinsic oscillators. One often stated advantage of these theories is that most can account for both residue pitch, as well as periodicity pitch, with the same mechanisms.

However, just as with the spectral models, the temporal models suffer from certain shortcomings. For instance, the physiological basis of these models is also uncertain. Thus while many central auditory responses can be interpreted as exhibiting delays or appropriate oscillatory patterns, the anatomical and physiological data do not yet coalesce as a whole into a compelling picture (Langner, 1992). Furthermore, most physiological pitch data tend to be in frequency ranges and from units with best frequencies that are relevant

for residue pitch (30–300 Hz) or slow temporal modulations (<30 Hz) (Schreiner and Urbas, 1988; Schreiner and Langner, 1988; Schwartz and Tomlinson, 1990) rather than periodicity pitch. Finally, recent psychoacoustical findings have been interpreted in favor of a dual (rather than a unitary) model of pitch perception (Carlyon, 1998a).

In summary, it is fair to say that spectral pitch theories would be more palatable to many: (1) if there is a biologically compelling mechanism for how harmonic templates might come about, and evidence for their existence; and (2) if the models could be extended to take into account "residue pitch" percepts and their properties. This paper addresses primarily the first issue, and provides ideas for what physiological mechanisms and anatomical substrates are potentially involved, and where to search for them. An important goal of this paper is to demonstrate that harmonic templates may emerge as a consequence of basic properties of early auditory processing, and not of exposure to any special sound stimuli such as harmonically rich speech or music. To emphasize this point, we shall use broadband noise and irregular click trains (i.e., sounds that lack any harmonic character) to produce the harmonic templates. We shall also briefly touch upon the problem of residue pitch, and discuss potential candidate mechanisms for unifying the estimation of both periodicity and residue pitch percepts without resort to organized correlation delay lines and other purely temporal structures.

The model we describe here explains how harmonic templates could emerge as a simple consequence of coincidence detection among channels representing the outputs of a cochlearlike filter bank. Once formed, the templates can be used to estimate the pitch as in the many variants of the spectral-matching pitch algorithms. Our focus in this paper is on the template-formation phase. Our goal is to illustrate how biologically plausible processes, response patterns, and connectivity in the early auditory nuclei can give rise to ordered harmonic templates without the need for any specially tailored inputs (such as clean harmonic complex tones), or supervised constraints (such as labeled and ordered inputs and outputs).

In the following, we shall first illustrate the essential mathematical structure of the model (Sec. I) and then discuss why the templates emerge (Sec. II). Next we discuss the potential biological structures and pathways that underlie the model (Sec. III). The implications of this model to the encoding of residue pitch are discussed in Sec. IV. We finally discuss the wider implications of our findings to models of auditory processing and to neural processing strategies in general (Sec. V).

I. A MATHEMATICAL MODEL FOR HARMONIC TEMPLATE GENERATION

The two basic stages of the model are illustrated in Fig. 1: An analysis stage consists of filter bank followed by temporal and spectral sharpening analogous to the processing seen in the cochlea and cochlear nucleus. The second stage is a matrix of coincidence detectors that computes the pairwise instantaneous correlation among all filter outputs.

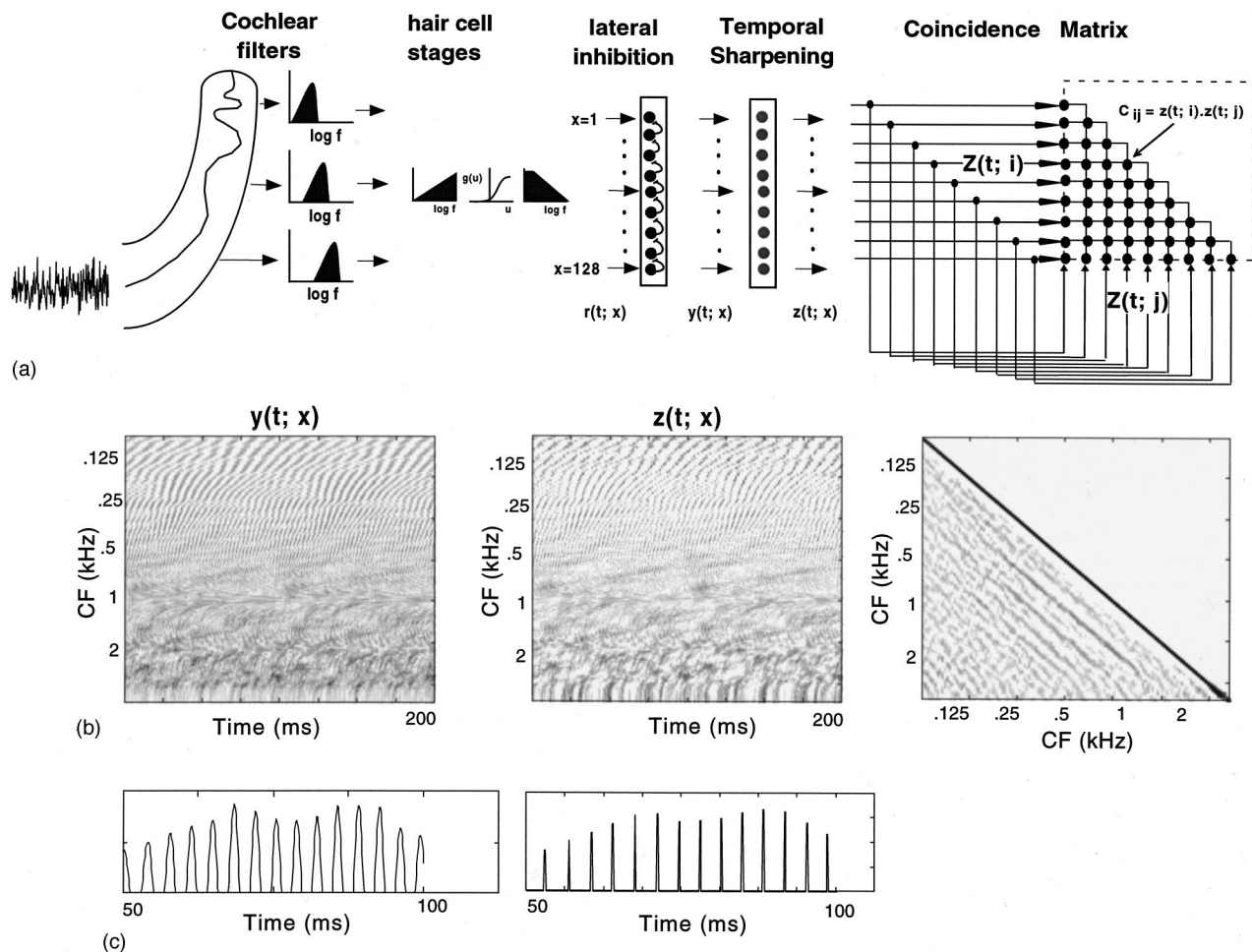


FIG. 1. Schematic model of early auditory stages. (a) Sound is analyzed by a bank of 128 tonotopically ordered cochlear filters spanning CFs between 100 and 4000 Hz. The output waveform from each filter is passed through a hair cell model ($r(t; x)$), followed by a first difference across the channel array simulating the action of a lateral inhibitory network (LIN) ($y(t; x)$). The responses are then temporally sharpened, becoming more synchronized within each channel ($z(t; x)$). The final stage is a matrix of coincidence detectors that compares the responses from all pairs of channels across the array. (b) The spatio-temporal responses of the channel array at different stages of the model: (left-to-right)—The responses at the LIN output ($y(t; x)$); the synchronized responses ($z(t; x)$); the output of the coincidence matrix (C) after one iteration. (c) The waveform transformation at the synchronization stage. (Left)—the waveform at CF ≈ 140 Hz ($y(t; x=12)$). (Right)—the waveform after temporal sharpening.

A. The analysis stage

This stage consists of a simplified minimal model of early auditory processing. It consists of a cochlear filter bank, followed by hair cell rectification and central spectro-temporal sharpening. These operations are depicted in Fig. 1, and described below in detail.

1. Cochlear filter bank

We employ a bank of 128 bandpass filters, equally spaced along a logarithmic frequency axis, x with center frequencies (CF) spanning a range of 5.3 octaves. The filters are moderately tuned and significantly asymmetric, with a steep roll-off on the high-frequency sides, as illustrated in Fig. 2(a) (Wang and Shamma, 1994; Yang *et al.*, 1992). They have constant Q 's, and hence their bandwidths gradually broaden (on a linear scale) toward the higher CFs. They are also related to each other by a simple dilation of their impulse responses. Given a discrete-time signal $s(t)$, and cochlear filter impulse responses $h(t; x)$, $x = 1, \dots, 128$ and $t = 0, 1, \dots, n$, any filter's response is computed as

$$u(t; x) = s(t) * h(t; x), \quad (1)$$

where $*$ denotes convolution with respect to time.

2. Hair cell filtering and rectification

Hair cells convert the filter outputs into electrical activity along the tonotopically ordered auditory-nerve array. This biophysical process is usually modeled by a three-step process (Shamma *et al.*, 1986; Shamma and Morrish, 1986): a high pass filter accounting for the velocity coupling of the hair cell cilia; a sigmoid function that describes nonlinear hair cell transducer channels; and a low pass filter representing the leakage in hair cell currents that gradually attenuates phase-locked responses beyond 800 Hz.

Here we shall simplify the analysis by incorporating the first temporal derivative into the cochlear filters. Next, the hair cell nonlinearity $g(\cdot)$ is modeled as a simple half-wave rectifier:

$$r(t; x) = g(u(t; x)) = g(s(t) * h(t; x)), \quad (2)$$

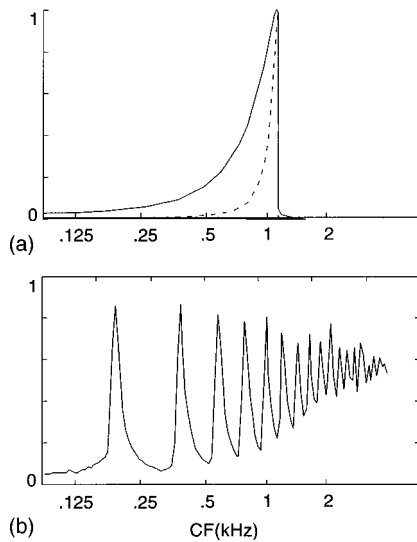


FIG. 2. Details of the model filters. (a) (solid line) Magnitude transfer function of the filter at CF=1 kHz; (dotted line) The effective magnitude transfer function *after* the LIN stage (see text). (b) The integrated output of the LIN ($\sum_i |y(t;x)|$) reflecting the spectrum of a harmonic series stimulus consisting of 20 harmonics of a 200-Hz fundamental [see Wang and Shamma (1994) for details].

where $g(u)=0$ for $u<0$, and $g(u)=u$ otherwise. Note that $g(\cdot)$ can be redefined as a sigmodal function to account for more complex nonlinear effects such as saturation or wider dynamic ranges. The effects of these added modifications is small for reasons discussed later. The hair cell low pass filter is bundled into the following stage as we describe next. The model outputs at this stage are depicted in Fig. 1(b)–(c) for a broadband noise stimulus.

3. Spectral and temporal sharpening of the filter outputs

This stage is helpful in enhancing the representation of the harmonics in the templates as we shall discuss later. Spectral sharpening mimics the effect of lateral inhibition (Shamma, 1985a, b), and is modeled by a simple derivative across the channel array (or a first-difference operation between the filter outputs) (Wang and Shamma, 1994):

$$y(t;x) = r(t;x) - r(t;x-1), \quad (3)$$

for $x=2, \dots, 128$, and $y(t;1)=0$. It can be shown that this step effectively sharpens the cochlear filters (Wang and Shamma, 1994; Lyon and Shamma, 1996), and is in principle unnecessary if the cochlear filters used are sharp enough to resolve approximately up to eight harmonics. Figure 2(b) illustrates that our frequency analysis at this stage can partially resolve approximately 8–10 harmonics of a 20-harmonic series stimulus (with a fundamental at 200 Hz).

The next stage performs temporal sharpening which enhances the synchrony of the phase-locked responses. This process mimics transformations such as those seen between the auditory-nerve and the onset units of the cochlear nucleus (Oertel *et al.*, 1990; Palmer *et al.*, 1995; Rhode, 1995). It is approximated by sampling the *positive peaks* of $y(t;x)$:

$$z(t;x) = \sum_{t_p} \delta(t-t_p) \cdot y(t_p;x), \quad (4)$$

where t_p =locations of the positive peaks in time, and $\delta(\cdot)$ is the discrete Dirac delta-function [$\delta(0)=1$ and $\delta(\cdot)=0$ otherwise]. $z(t;x)$ then becomes a spectrally sharpened and highly temporally synchronized version of the filter responses as illustrated in Fig. 1(b)–(c). A simple way to include the effects of the diminishing phase-locked responses with increasing frequency is to replace $\delta(\cdot)$ with a pulse of variable width $\Pi_m(\cdot)$, starting at the zero-crossing point, i.e., $\Pi_m(k)=1$, $0 \leq k \leq m$; the larger m is, the smaller is the frequency range of phase-locking and synchrony [Fig. 1(c)]:

$$z(t;x) = \sum_{t_p} \Pi_m(t-t_p) \cdot y(t_p). \quad (5)$$

B. The coincidence matching stage

This stage performs an instantaneous match between the responses of all pairs of channels in the array, and integrates all results over time to produce its final output. From a mathematical perspective, the network is a matrix of coincidence detectors, each multiplying the responses from a pair of channels as depicted in Fig. 1(a):

$$C_{ij}(t) = z(t;i) \cdot z(t;j), \quad (6)$$

for all $i, j=1, \dots, 128$ such that $j < i$; for $i=j$, $C_{ii}(t) = z(t;i)$. The absolute values of $C_{ij}(t)$ are then accumulated over time until an adequately smoothed output T_{ij} is obtained:

$$T_{ij} = \sum_{N,t} |C_{ij}(t)|, \quad (7)$$

for N realizations of a random stimulus. Note there are no neural delays anywhere in this model. Instead, coincidences are computed from simultaneous outputs of the filter bank, and the results are then integrated over time. Note also that in the equation above, it is the *absolute* value of the coincidence that is integrated, and that the average value of $z(\cdot)$ can be removed because it is about the same for all channels, and hence contributes only a uniform constant at all locations.

C. Model simulations

The coincidence network above is capable of producing the harmonic templates as its final averaged output regardless of the exact nature of its input signal, provided it is broadband conveying energy at all frequencies < 3 kHz. We illustrate in Fig. 3 the templates generated with broadband noise and random click train input signals $s(t)$. In Fig. 3(a), the 200-ms stimulus consists of equally spaced, random-phase tones (with 10-Hz separation, in the range between 10 Hz and 4000 Hz) with random phases. Usually many examples of $s(t)$ are generated with different random phases. The final output of the network (T_{ij}) is the average over all these stimulus iterations [$N=300$ in Fig. 3(a)]. Figure 3(b) shows the average output T_{ij} for a random click train stimulus with random widths ($N=300$).

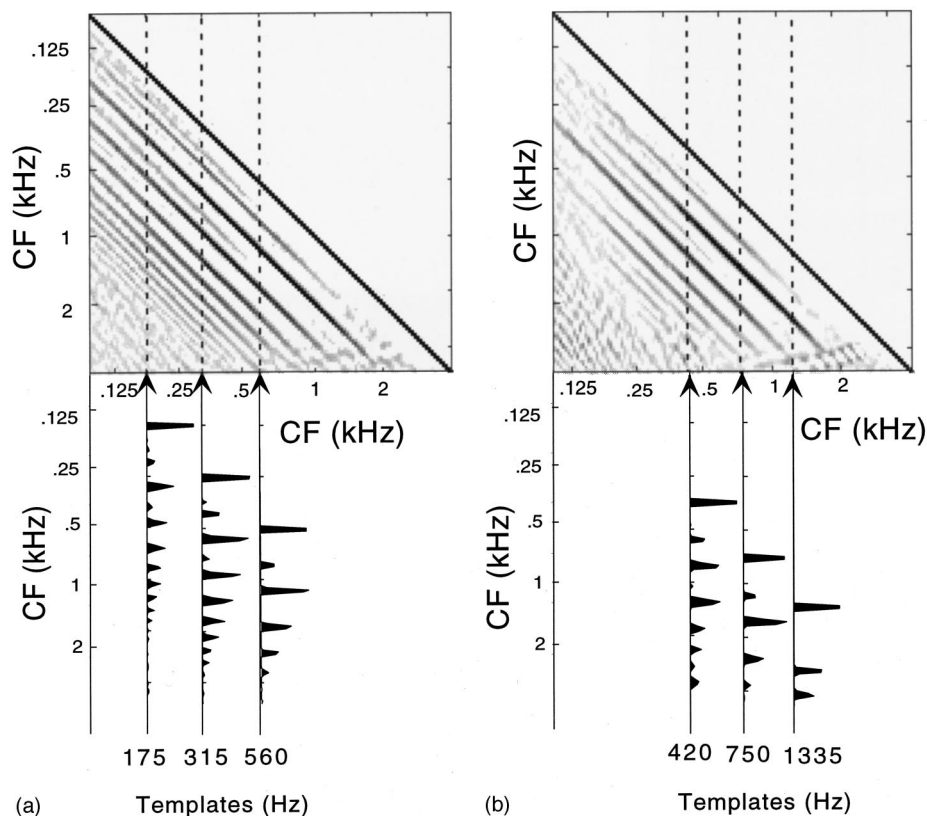


FIG. 3. The harmonic templates in the integrated output of the coincidence matrix. Templates emerge as regions of high coincidence that run parallel to the main diagonal, and are exactly spaced at harmonically related CF distances. (a) The templates generated by a broadband noise stimulus. Three templates are shown individually by the cross sections (fundamentals at 175, 315, 560 Hz). For each, the pattern shows prominent peaks at harmonically related CFs, that gradually decrease in amplitude for higher order harmonics. (b) The templates generated by a random click train with random widths. Cross sections for the three templates are shown below the figure (fundamentals at 420, 750, 1335 Hz).

The simulations show strongly correlated outputs from channels that are separated exactly by harmonic distances from each other. These strong coincidences form a pattern of multiple diagonals that are spaced at exactly harmonic intervals apart. For instance, consider the pattern of strong coincidences for the channel at CF=175 Hz displayed below the coincidence matrix outputs in Fig. 3(a). The pattern shows prominent peaks at CFs that are integral multiples of 175 Hz. This pattern is interpreted as the “harmonic template” of the 175-Hz series. Similarly, the templates for all other harmonic series can be found across the diagonals of the network output (cf. the harmonic series templates for several other fundamentals in Fig. 3). Note also that the number of harmonics represented in each template decreases with increasing fundamentals as phase-locking diminishes gradually beyond 1 kHz.

D. Final comments

The mathematical structure of the network and simulations described above are but one example of many variants that can be used. The two key operations are a cochlearlike filtering stage followed by coincidence detection. Relaxing the degree of spectral and temporal sharpening in the model only gradually reduces the clarity of the templates by either diminishing the height of the harmonic peaks or reducing their number. Similarly, replacing the “product” in the coincidence operation [Eq. (6)] with a squared sum or other “matching” operations does not alter the locations of the harmonic peaks. The reasons behind this robustness are discussed in the next section.

II. WHY DO THE HARMONIC TEMPLATES EMERGE?

In this section, we examine the reasons why the coincidences in Fig. 3 occur at harmonic intervals between the cochlear channels despite the lack of any harmonic structure in the input stimuli. We shall specifically discuss the critical role played by three subtle but important factors in the model: the nonlinear transformations following the filtering stage; the rapid phase shifts of the traveling wave near its resonance; and the spectral resolution of the cochlea.

A. Nonlinear transformations of the filter outputs

In the model outputs, the harmonic-template lines emerge as a consequence of the strong coincidences between responses of harmonically related cochlear filters. To understand why this is so, consider a sharply tuned filter bank driven by a broadband noise stimulus. Each filter in this bank produces a phase-locked response waveform that is quasi-periodic and reflects predominantly its CF (Ruggero, 1973). This is exemplified in Fig. 4 by the quasi-sinusoidal responses for a CF \approx 250 [Fig. 4(a)]. If the filter outputs are not half-wave rectified or otherwise nonlinearly distorted, then the outputs from any such pair of filters will be orthogonal (or entirely uncorrelated) since each would contain Fourier coefficients only near its CF.

However, the situation is drastically different if the filter responses are half-wave rectified, because this creates “distortion” components and the waveform can be thought of as composed of a fundamental frequency (the CF of the filter) and its harmonics [Fig. 4(b)]. Consequently, the rectified waveform from any filter can now partially coincide with

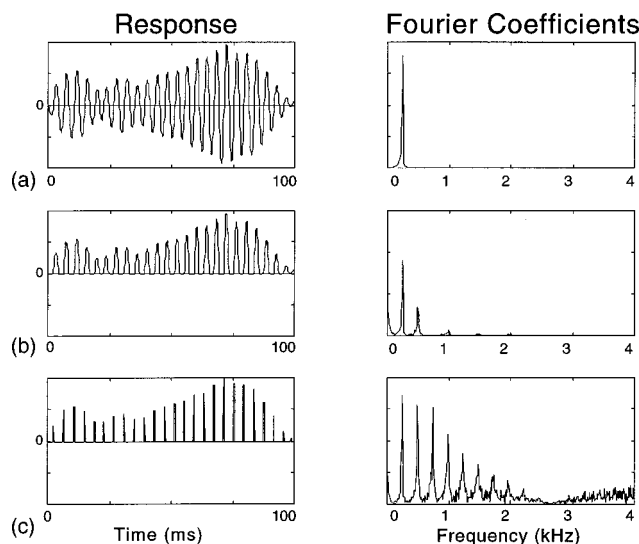


FIG. 4. The effects of nonlinear deformations and temporal sharpening. The response waveforms within each channel implicitly convey harmonic-distortion components with varying strength. The left column illustrates response waveforms with increasing nonlinear distortion and synchrony; the right column illustrates the Fourier coefficients corresponding to each waveform. The number and amplitude of the harmonic distortion components increase with increasing synchrony and nonlinear deformation of the response waveform. (a) Linear filter response at $CF \approx 250$ Hz; (b) half-wave rectified response; (c) the synchronized impulse train corresponding to the 250-Hz response.

outputs of other filters that are at harmonically related CFs. For instance, the rectified waveform from the filter at $CF = 250$ Hz contains harmonics of 250 Hz with gradually decreasing intensity [Fig. 4(b)], and hence may coincide strongly with filter outputs at $CF = 500, 750, \dots$ Hz. The important role played by the half-wave rectification is not unique to this operation; rather, it is a common consequence of many instantaneous nonlinear distortions of the filter outputs. For example, similar harmonic coincidence patterns emerge if the filter waveforms are distorted by a saturating nonlinearity, a limited dynamic range, or are converted to a series of synchronized impulses as is done in the model; Eq. (4) [Fig. 4(c)].

It is in this context that one can appreciate the role of enhanced temporal synchrony in the model. The synchronization of the filter response waveforms is a highly nonlinear operation that ensures that the impulse train from each filter contains within it the fundamental frequency (at the CF) and, prominently, many of its harmonics [Fig. 4(c)]. That is why the pulse train from a filter at $CF = 250$ Hz will correlate well with pulse trains produced by filters at harmonically related CFs up to a relatively high order.

B. The phase of the cochlear traveling wave

How is it possible that the highly synchronized waveform at a given CF (e.g., 250 Hz) be in just the right phase to coincide with outputs from other CFs (e.g., the response at 500 Hz)? The answer highlights the role of the cochlear traveling wave, specifically its phase delays, in the formation of the templates.

Figure 5 illustrates the typical features of two traveling waves evoked by two tones, say at 250 and 450 Hz. Near the

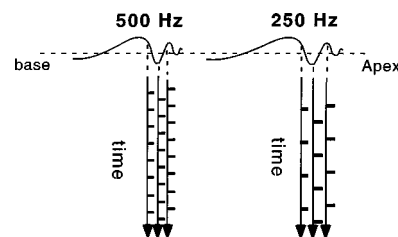


FIG. 5. Traveling wave phase shifts near the resonance of a traveling wave. The schematic illustrates that the response patterns near the resonance of the traveling waves can be significantly phase shifted relative to each other over very short distances.

resonance of each wave ($CF = 250$ and 500 Hz), the travel velocity decreases rapidly, and the wave as a result accumulates phase delays at an accelerated pace (Lyon and Shamma, 1996; Shamma, 1985a). Consequently, near the CF, one may find responses of widely different (even opposite) phases in closely spaced locations (or channels). That is, each of the CF regions of 250 and 500 Hz contains synchronized responses to these frequencies at various phases, and hence it is likely that at least a pair of channels will coincide and positively correlate. This argument still applies when the stimulus contains many tones (as with a harmonic complex or broadband noise) because these phase delays are characteristic of the cochlear filters and not of the stimulus. Thus as long as the responses at a given CF are determined by a relatively sharply tuned cochlear filter, they will necessarily exhibit these rapid phase shifts near the CF, as can be seen in Fig. 1(b) where the synchronized responses to the noise stimulus are similar in adjacent channels except for a rapid phase delay toward the lower CFs. Finally, note that the formation of the templates requires only that the local phase shifts be relatively rapid, and is insensitive to changes in absolute phase values or in the detailed shape of the rapid phase functions. Consequently, nonlinear phase changes such as those induced by increasing sound levels and other manipulations are unimportant as long as they leave the travelling wave phase functions relatively rapid near the resonance.

C. The sharpness of frequency analysis

Cochlear frequency analysis and subsequent spectral sharpening of the filter outputs [by lateral inhibition [Shamma, 1985b; Eq. (3)] enhance the features of the harmonic templates. This is because sharp filters (by definition) respond only to frequencies near their center frequencies, and hence usually produce more regular (periodic) synchronized responses regardless of the nature of the input stimulus. This point is illustrated in Fig. 6 where we examine the effect of broadening the cochlear filters on the synchronized responses to a broadband noise stimulus. Figure 6(a) shows the synchronized responses (left plot) and their corresponding Fourier series coefficients (right plot) using our regular filters. Here, the response at each CF contains well-defined components at CF and its harmonic distortions as is evident by the well separated Fourier peaks. If the filters are made significantly broader (for instance, by removing the lateral inhibition stage), the synchronized responses from each filter

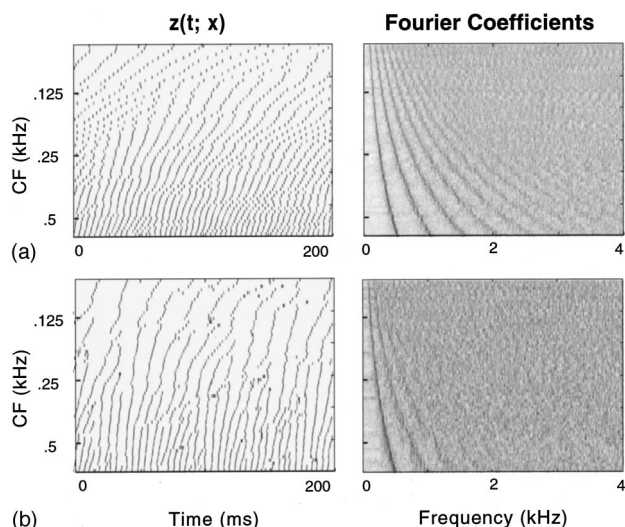


FIG. 6. The effects of spectral resolution on the templates. (Left) The synchronized responses to a broadband noise stimulus [as in Fig. 1(b)]. (Right) The corresponding Fourier series coefficients for all channels (each labeled by its CF along the ordinate). (a) The responses due to the regular model filters (as in Fig. 2). (b) The responses using broader filters (by removing the LIN stage in Fig. 1).

become considerably more jittered due to the increased interference within each channel [Fig. 6(b), left plot]. This in turn smears considerably the Fourier representation of the higher order distortion harmonics [Fig. 6(b), right plot]. Therefore, cochlear frequency selectivity is critical for the formation of the harmonic templates: Sharper filters result in clearer high order harmonic peaks in the templates.

D. Summary

The harmonic templates arise from two basic processing stages: cochlear filtering, followed by a matrix of coincidence detectors. The precise shape of these templates, the clarity of their peaks, and the order of their highest harmonics is influenced by the details of these two operations. The following list summarizes these factors:

- (1) Phase-locking of the filter responses is a critical factor in the template formation. All templates are ultimately derived from the fine-time structure of the filter responses. Thus the gradual loss of phase-locking (or synchrony) to higher frequencies (approximately $>2-3$ kHz) is the reason why they are not represented in the templates, and hence play little or no role in the perception of periodicity pitch. In the model, the degree of phase-locking can be simulated by changing the width of the pulse function $p(t)$: the sharper the pulse, the better is the phase-locking to higher frequencies.

- (2) Nonlinear transformation of the filter responses is essential in generating the (distortion) harmonics that ultimately form the templates. Half-wave rectification and increasing temporal synchrony are two such transformations. Thus increasing temporal synchrony improves the representation of the higher harmonics.

- (3) High spectral resolution improves the representation of the harmonic peaks in the templates. In the model, the lateral inhibitory stage increases the effective tuning of the

filters; removing this stage therefore reduces the number and sharpness of the harmonic peaks in the templates.

- (4) Phase delays of the traveling wave provide locally phase shifted copies of the responses at each CF. While such phase shifts are typical near the resonance of any bandpass filter, they are especially large in the cochlear filters because of their steep high-frequency roll-off just above the CF. Note that it is important in the model to provide sufficiently dense sampling of the CF axis (number of channels/octave) in order to capture these phase shifts; the sparser the sampling, the weaker are the coincidence peaks in the templates.

- (5) The formation of the harmonic templates and their parameters are solely determined by the intrinsic properties of the cochlear filters and coincidences and not of the stimulus. That is, given enough time, the same templates will emerge for any broadband stimulus whether it is noise, harmonic sequences, or impulses.

Note that the combined effect of all these factors give rise to templates (Fig. 3) with features that resemble closely those suggested by some of the algorithmic implementations of the spectral pitch theories [e.g., as in Duifhuis *et al.* (1982); Cohen *et al.* (1995)]. For example, in these implementations, the ideal harmonic templates with their equal amplitude spectral lines (as in Goldstein, 1973b) are modified in two ways: Harmonic peaks are gradually decreased in amplitude and/or increased in width with increasing order. These features arise in our templates due to the various factors discussed above.

Finally, we observe that a close examination of the harmonic templates in Fig. 3 reveals substantially smaller peaks that are interspersed among the harmonic peaks. These peaks are due to sub-harmonic interval correlations. For instance, the 175-Hz fundamental has moderate positive correlation with *approximately* 262 Hz, because these two frequencies are integer multiple harmonics (2nd and 3rd) of *approximately* 87 Hz; other peaks between the first and second harmonics of this template may sometimes be visible at the $\frac{3}{4}, \frac{4}{5}, \dots$ harmonic ratios. Most other sub-harmonic peaks are much smaller and are rarely evident in our simulations.

The above argument suggests that the templates formed by the coincidence matrix (Fig. 3) are not exactly the ideal harmonic templates hypothesized by classic central pattern matching models (Goldstein, 1973b). Rather, our templates should be more accurately described as equivalent to the “auto-correlation” of those ideal templates. However, the computational errors in using our templates in the same manner as an ideal harmonic template are insignificant because the “nonharmonic” peaks in our templates are relatively small.

III. PHYSIOLOGICAL CORRELATES OF THE MODEL

We discuss here the biological plausibility of the model and the correspondence between its stages and known physiological responses in the early auditory pathways. Some elements of the model have clear biological underpinnings, while others are speculative. For instance, the frequency analysis, phase shifts around the CF, half-wave rectification, and the phase-locking of the responses are all well known analogs of basilar membrane and hair-cell function.

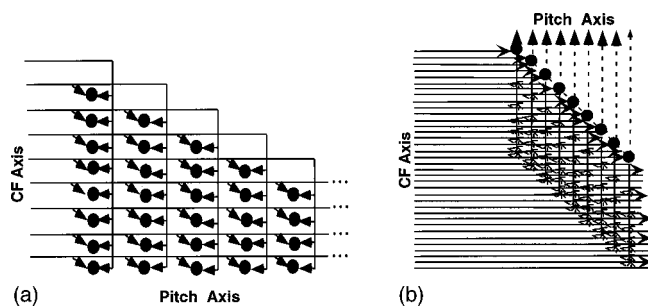


FIG. 7. Biological realizations of the coincidence detectors matrix. (a) The inputs from the auditory channel array are compared pairwise by the network of coincidence detectors. Cells in each column have a common CF input from one side, and a progressively increasing CF input from the other side. The templates emerge along the columns (as illustrated earlier in Fig. 3) when coincidence detectors at harmonic CF distances are strengthened, while others drop out. (b) A different realization where pairwise coincidences are measured and reinforced in the dendrites rather than in separate cells.

More speculative, however, is the anatomy and location of the coincidence matrix, and the identity of its immediate input pathway. Since phase-locking up to relatively high frequencies (at least 2 kHz) is necessary at the input of the matrix, this places it at, or prior to, the inferior colliculus. Furthermore, the synchronized responses at the input of the coincidence matrix are highly reminiscent of the responses of the variety of onset cells in the cochlear nucleus. While these observations suggest certain scenarios as depicted in Fig. 7, the early auditory system is clearly complex and mysterious enough to support many other variant, or even drastically different substrates.

Figure 7 shows two examples of possible “neural” realizations of the coincidence matrix. Figure 7(a) is a more literal interpretation of the mathematical model. The matrix consists of tonotopically organized coincidence detectors, where all cells in a column have the same CF, and are also driven by inputs from higher CFs. Thus in the fully formed matrix, each cell ends up driven by a pair of CF inputs: one at its primary CF, and another from a higher, harmonically related CF. In the alternative realization of Fig. 7(b), each cell is driven by its primary CF, but it also has an extensive dendritic tree which spans higher CFs. Initially the dendrites are devoid of synapses. They begin to form during the learning phase at CF locations where the responses correlate well with the primary CF input. In the end, each coincidence cell will be driven by many CF inputs, and hence will appear very broadly tuned. Clearly, a mix of these two scenarios is also possible.

But where is the input pathway to the coincidence matrix? The candidate pathway must be spectrally well resolved and phase-locked as in the auditory nerve. In the cochlear nucleus, many cell types exhibit the appropriate spectrally and temporally sharp responses, especially the onset and primarylike cells in the low CF regions (Rhode, 1994, 1995; Smith and Rhode, 1989; Evans and Zhao, 1998). These cells may project to the coincidence matrix in the limniscus nuclei or the IC. Alternatively, Fig. 7(b) resembles closely the anatomical features of the Octopus cells (the presumed onset-I cells) (Oertel *et al.*, 1990; Palmer *et al.*, 1995), suggesting that they may serve themselves as the coincidence matrix.

Unfortunately, most data available at present from various onset cells and other appropriate cell types in the cochlear nucleus are from units with relatively high CFs (> 3 kHz), and hence one cannot be certain of their role in periodicity pitch (Palmer *et al.*, 1995; Rhode, 1995; Evans and Zhao, 1998). For instance, the strong dependence of onset cells (especially onset-I) on the phases of the components a complex tone stimulus observed in high CF cells may not occur in low CF cells (Evans and Zhao, 1998).

There are numerous pitch phenomena that are closely related to periodicity pitch, and derived exclusively from binaural stimuli (such as the Huggins pitch³). These results suggest that the coincidence detectors may be located at or post binaural convergence nuclei. For instance, it is conceivable that the MSO can serve both its traditional binaural coincidence role (Jeffress, 1948; Shamma *et al.*, 1989), and a monaural coincidence role for the encoding periodicity pitch. Clearly, there is little solid support at present to indicate the existence of such structures in the IC or other central nuclei, and the only definite conclusion that can be made at this time is that much more physiological data are needed to disprove any of these hypotheses.

IV. RESIDUE PITCH

Humans perceive a clear “residue” pitch from tone complexes of unresolved components that is equal to the period of the waveform envelope. Unlike periodicity pitch, this percept is sensitive to the phase of the components and is weakest when they are in random phase. It also has different psychoacoustical properties, e.g., bigger *dl*'s, and a different dependence on tone duration (Carlyon, 1998b). This pitch is not related to any harmonicity in the stimulus. There is, therefore, little reason to assume that this pitch is derived from the harmonic templates; instead, it may have a different origin and neural mechanisms, a conclusion also supported psychoacoustically (Carlyon, 1998a). Nevertheless, the need to unify these two pitch percepts in a single mechanism has been a strong motivation for the development of the “temporal” models of pitch alluded to earlier.

A. Representing residue pitch in the coincidence matrix

It is possible to show that a simple scheme based on the coincidence matrix is also capable of measuring residue pitch. The basic idea is illustrated in Fig. 8 for a harmonic series stimulus consisting of in-phase high (unresolved) harmonics (10th–110th harmonics of a 70-Hz fundamental). Figure 8(a) shows the synchronized responses evoked by this stimulus (generated by the cochlear model described in Sec. I). A necessary additional ingredient for this scheme to work is a simple monotonic relative increase in response latency from high-to-low CFs at a rate of a few milliseconds per octave [e.g., 6 ms per octave in Fig. 8(b)], similar to that found or postulated in the IC and cortex (Langner and Schreiner, 1988; Greenberg *et al.*, 1998; Hattori and Suga, 1997). This latency-shift effectively delays the response waves in the low CF channels relative to the high CF as shown in Fig. 8(b) causing them to overlap and coincide across some channels. The coincidence matrix indicates the

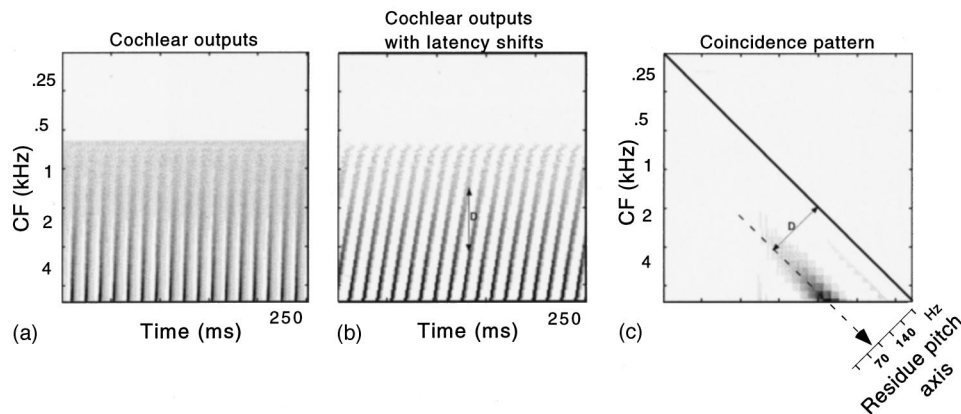


FIG. 8. Measuring residue pitch with the coincidence matrix. (a) The cochlear responses to a stimulus consisting of 10th–110th harmonics of 70 Hz. (b) The responses are systematically delayed by a gradual increase in latency from high to low CF channels. (c) The average output of the coincidence matrix shows a strong line of coincidences parallel to the diagonal, and at a distance (D) that reflects indirectly the period of the stimulus [as indicated in (b)]. Faster rates cause this distance (or line of coincidences) to move gradually closer to the main diagonal as indicated by the pitch axis.

repetition period of the responses in terms of the distance (D) separating the coincident channels in the input array. Thus different periods evoke different coincidence patterns, with faster rates (e.g., 140 Hz) causing coincidences closer to the center diagonal as illustrated in Fig. 8(c).

If the stimulus contains both resolved and unresolved harmonics, both will be simultaneously represented in the coincidence matrix outputs. For example, Fig. 9(a) (left panel) shows the outputs evoked by a stimulus composed of the 1st–31st harmonics of a 250-Hz fundamental. Two distinct coincidence patterns emerge. The first is the diagonal stripes in the high (unresolved) harmonics region (highlighted within the dashed circle). The other is the harmonically spaced “patches” in the low (resolved) harmonic region (outside of the dashed circle). Note that the borders of these two regions depend on the fundamental frequency of the stimulus. For instance, for a fundamental frequency of 70 Hz, the coincidence matrix output is dominated by the “residue pitch” diagonal patterns (center panel). The opposite is true for a 500-Hz fundamental stimulus (right panel) where only the pattern of resolved harmonic patches is evident.

Many of the well-studied properties of these two types of pitch percepts can be readily seen in the coincidence pat-

terns. For instance, consider the sensitivity of residue pitch to the phases of the unresolved harmonic components. If the phases are completely randomized, the synchronization of the cochlear responses to unresolved harmonics is severely disrupted, leading to the loss of the striped patterns in the output of the coincidence matrix as illustrated in the highlighted region of Fig. 9(b) (left panel) and the entire pattern in Fig. 9(b) (center panel). These figures also illustrate that the patterns in the resolved harmonics region are, as expected, insensitive to phase randomization [Fig. 9(b), left and right panels].

B. Amplitude modulated and rippled noise

Figure 10 illustrates how the coincidence matrix output represents the pitch percepts of two different broadband noise stimuli. In Fig. 10(a) the stimulus is an amplitude-modulated white noise with a modulation rate of 100 Hz (waveform below left panel). The cochlear responses are reminiscent of those due to unresolved harmonics (left panel), and so is the striped output pattern of the coincidence matrix (right panel). The second stimulus is the “iterated ripple noise” [Fig. 10(b)]. It is generated by adding (or sub-

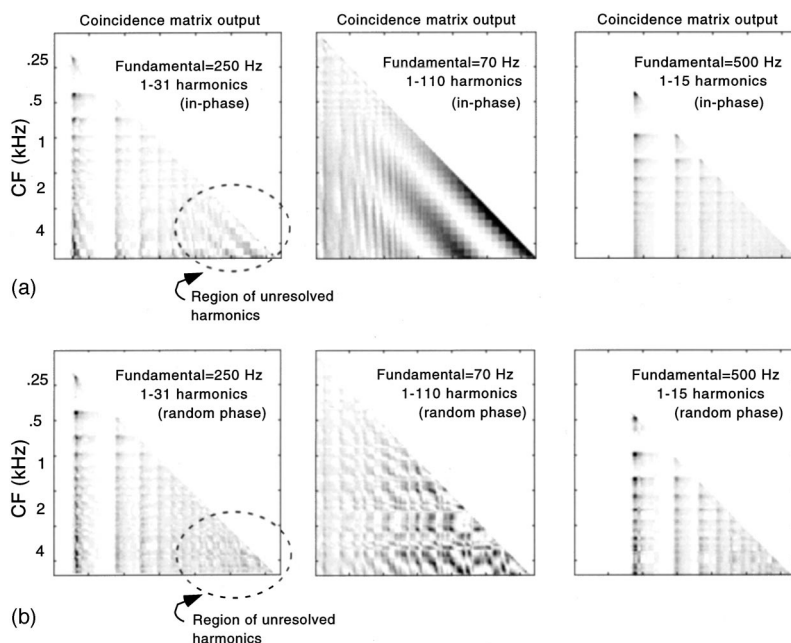


FIG. 9. The representation of resolved and unresolved harmonics in the coincidence matrix outputs. (a) The coincidence matrix outputs for in-phase harmonic series of different fundamental frequencies. (Left panel) A partially resolved harmonic series of 1st–31st harmonics of 250 Hz. The high-order unresolved harmonics give rise to a striped pattern which is highlighted by the dashed circle in the figure. (Center panel) A mostly unresolved harmonic series of 1st–110th harmonics of 70 Hz. The striped pattern dominates the output. (Right panel) A mostly resolved harmonic series of 1st–15th harmonics of 500 Hz. (b) The coincidence matrix outputs for the same harmonic series stimuli as above, but with randomized phases. The striped pattern due to the unresolved components disappears in left and center panels.

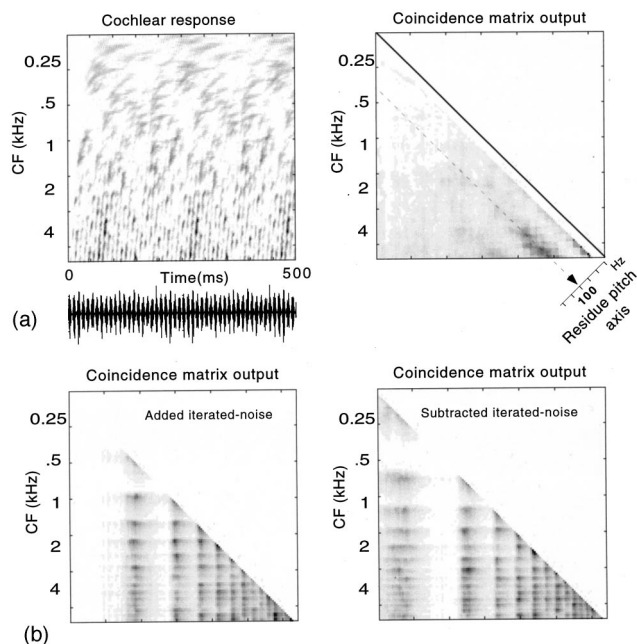


FIG. 10. The representation of pitch due to broadband noise stimuli. (a) Residue pitch evoked by amplitude-modulated broadband noise. (Left panel) Cochlear responses to a 100-Hz amplitude modulated noise (waveform shown below the panel). The responses in the high-CF regions are reminiscent of the responses to unresolved harmonics in Fig. 8. (Right panel) The coincidence matrix outputs indicating the location (at 100 Hz) of the coincidence peak along the same residue pitch axis as in Fig. 8. (b) The coincidence matrix outputs for iterated ripple noise constructed by delaying the noise (2 ms) and then adding (left panel) or subtracting (right panel) the noise to itself. The patterns resemble those of resolved components.

tracting) a delayed version of a white noise to itself several times (Yost and Hill, 1979). The stimulus in Fig. 10(b) is delayed by 2 ms, and added (or subtracted) to itself 16 times. The spectrum of such a stimulus has equally spaced peaks that are $1/d$ Hz apart (Yost and Hill, 1979). The coincidence matrix output exhibits a patchy appearance, and hence we interpret it as evoking *only* a percept of periodicity pitch (and not of residue pitch). Note that the added-noise stimulus (left panel) evokes a percept and a coincidence pattern similar to that of a resolved harmonic series of fundamental = 500 Hz [see also Fig. 9(c)]. By comparison, the subtracted-noise stimulus (right panel) resembles the *inharmonic* series (250, 750, 1250, ... Hz) and evokes a correspondingly different coincidence pattern.

C. Computing pitch values

The coincidence matrix can potentially be used to compute both “periodicity” and “residue” pitches depending on how the output is integrated. For periodicity pitch, the output is summed using the learned harmonic templates shown in Fig. 3 [or more graphically as in Fig. 7(b)]. For residue pitch, the output is summed along the diagonals illustrated by the dashed line in Figs. 8(c) and 10(a). Clearly, each of these two types of outputs contributes optimally only to one of the summation methods. Thus the striped pattern contributes little information if summed according to the harmonic templates. Similarly, the harmonic intervals between the patches are irrelevant along the diagonals.

D. Summary

The illustrations in Figs. 8–10 suggest that the coincidence matrix may serve as a common computational mechanism for both periodicity and residue pitch provided that two different summation strategies are employed: the harmonic templates for periodicity pitch and the diagonals for residue pitch. It is also possible that these two computations are segregated in the auditory system into two independent coincidence networks, where the representation of each pitch percept is independently optimized and only one summation strategy is used. The results from these two coincidence networks could subsequently be combined and registered relative to each other (Carlyon, 1998a).

V. DISCUSSION

We have described a model for how harmonic templates might arise during early development of the auditory system. The model demonstrates that the templates are a natural consequence of basic properties of processing in the early stages of the auditory system. Most important among these properties are cochlear filtering, phase-locked representation of its outputs, enhanced temporal synchrony, and, finally, coincidences across the channel array. We have discussed the contributions of each of these properties to the clarity of the template peaks and the highest harmonic order represented.

An important conclusion from this model is that the harmonic templates are robust and reflect fundamental features of peripheral auditory function. Thus for the model to work at all, we must have cochlear frequency analysis; we must have rapid traveling wave delays near the wave’s resonance; and we must have phase-locking and half-wave rectification on the auditory nerve. Beyond these fundamental features, all other details, such as enhanced temporal synchrony and spectral sharpness, are helpful in improving the templates in a graded fashion.

Another important conclusion is that template formation is largely independent of the stimulus as long as there is energy available at all frequencies (<3 kHz) over a period of time. That is, harmonic templates will appear if we had used harmonic sounds, impulses, or any other broadband stimulus provided that all frequencies are represented over the ensemble. However, even if the stimulus energy is not well balanced due, for instance, to partial threshold elevation or a notch in the audiogram, the templates will still arise, but with reduced contributions from these frequencies. For example, if the channel at CF=400 Hz is removed at the outset (e.g., due to a localized hair cell death at that location), then the model predicts that the 400-Hz template will not be learned, and that this pitch will not be heard from a complex of higher order harmonics (e.g., 800, 1200, 1600 Hz). All other templates will form, but with contribution from the 400 Hz missing. For instance, the 200-Hz template will have all its peaks intact except for the 400 Hz. Note that this prediction is contrary to that obtained from a “temporal” model such as the correlogram (Slaney and Lyon, 1993), where the perceptual contribution to the 400-Hz pitch comes from all CF channels regardless of what is happening at the CF=400 Hz channel.

Finally, the model suggests a simple answer to the question of why harmonic templates postulated in psychoacoustic studies have a prominent fundamental when natural harmonic sounds (e.g., speech) often have little or no energy at the fundamental? Equivalently, why does a partial set of upper harmonics evoke a pitch at the fundamental and not at any other arbitrary frequency, thus implying that the learned templates must be linked to the fundamental? The answer is that harmonic templates are formed from exposure to broadband noise, clicks, and other stimuli where all frequencies are available, and not simply from examples of harmonic sounds such as voices which may not have the fundamental.

A. Where to search for physiological evidence

What physiological or anatomical evidence should we look for to confirm the presence of the harmonic templates? Two sets of data are needed to shed light on the model. The first concerns the inputs to the coincidence matrix, and the second deals with the coincidence cells themselves. The input pathway must be sharply tuned in its *synchronous responses*. This is an important consideration which is often ignored when reporting on the tuning or iso-intensity response curves of these cochlear units. To establish the relevance of any cells for the encoding of periodicity pitch, it is essential that the units studied receive phase-locked auditory-nerve inputs, and hence must have low CFs (< 3 kHz). It is also best if their tuning properties are measured with reference to their phase locked, and not their average rate, inputs. For example, an onset cell may appear very broadly tuned due to its relatively high threshold and very limited dynamic range (Rhode, 1995; Evans and Zhao, 1998). However, the unit may be sharply tuned if one considers how well it is synchronized to one of several closely spaced stimulus components (Greenberg *et al.*, 1986). Alternatively, onset cells may constitute the coincidence matrix themselves, and hence receive input from several CFs (e.g., a pair), and appear broadly tuned.

Coincidence detectors, wherever they may reside, should exhibit distinctive response patterns to harmonic series stimuli such as click trains. For instance, coincidence cells as in Fig. 7(b) should be selective to the rate of a click train (tuned to the fundamental of the cell's template); they must also be insensitive to the phase of the harmonics in the stimulus; and finally, they should be broadly tuned, or at least broadly facilitated (Palmer *et al.*, 1995). The response patterns are different if the coincidence cells have pairwise inputs as in Fig. 7(a). The cells should be doubly tuned or facilitated, and must exhibit predictable tunings to multiple click rates in a phase-insensitive manner. None of these response properties have been reliably demonstrated in the IC or lower auditory nuclei, and it remains to be seen if more controlled recordings in the low CF regions can shed light on these questions.

B. The principle of coincidence detection

In biologically inspired models of various auditory tasks, it has been common to postulate neural delays (prior to coincidence detection) so as to affect various correlation operations. These delays are explicit in some algorithms, e.g.,

in the binaural processing of interaural time delays (Jeffress, 1948; Colburn and Durlach, 1978), or in computing the correlograms for pitch (Slaney and Lyon, 1993). In other algorithms, these delays are implicit only within purely temporal operations that must use them, e.g., in the ALSR and dominant frequency algorithms for spectral shape extraction from auditory-nerve responses (Young and Sachs, 1979; Lyon and Shamma, 1996), or in the use of intrinsic oscillations or firing intervals of variable rates for pitch estimation (Langner, 1992; Hewitt and Meddis, 1994; Winter *et al.*, 1999). As mentioned earlier, the need for neural delays stems almost entirely from the need to make interval measurements on single channels independent of other channels.

In previous reports, we have demonstrated that simple coincidence measurements of responses *across* the auditory channels can extract the same kinds of information robustly, without need for functional neural delays. Thus lateral inhibition across the outputs of the auditory-nerve fiber array (which is essentially a form of coincidence detection) can extract a highly resolved spectrum of a broadband complex stimulus over wide stimulus levels (Shamma, 1985a, b). Similarly, a coincidence matrix identical to the one discussed here (Fig. 1) for auditory channels from the two ears [the stereausis network (Shamma *et al.*, 1989)] potentially can explain binaural phenomena accounted for by traditional cross-correlation models. The algorithm described in this paper repeats the same theme discussed above. That is, coincidences across the fiber array carry sufficient information to generate the harmonic templates, and hence obviate the need for the neural delay lines invoked in many of the current pitch models.

To summarize, the need to invoke neural delay lines stems from a common view of auditory processing as primarily temporal in the sense defined earlier. This view may be partially a consequence of the experimental difficulty of measuring the distribution of auditory responses across the tonotopic axis, and hence of appreciating the richness and subtlety of the *spatiotemporal* cues created by the cochlea. Decoding such cues often requires simple coincidence detection to measure time differences between channels rather than absolute time intervals within a channel. Temporal models, instead, essentially recreate cochlealike frequency analysis centrally by postulating additional ordered delay lines, correlators, and narrowly tuned filters with fine (microsecond) accuracy to detect and measure response periodicities and stimulus parameters.

C. Using the model to compute pitch

This paper has primarily dealt with the question of how the harmonic templates associated with periodicity pitch emerge as a natural consequence of cochlear function and subsequent coincidence networks. We have also discussed in passing the representation of residue pitch in the coincidence network (Sec. IV), and hence the potential role of coincidence as a unifying mechanism underlying both types of pitch percepts.

It is possible to formulate a complete and detailed computational model of pitch perception based on these mechanisms. The model would conceptually consist of two parts:

(1) computation of periodicity pitch based on the learned harmonic templates; and (2) a computation of residue pitch based on the diagonal patterns in the coincidence matrix outputs. As eluded to earlier, the periodicity pitch portion of this model has existed for decades in the literature as template matching algorithms of various forms (Duifhuis *et al.*, 1982; Cohen *et al.*, 1995). For instance, Cohen and Grossberg (1995) presented a detailed account of the psychoacoustical data that could be accounted for by the harmonic template match model, including the ambiguous pitches and phase insensitivity. Undoubtedly, further refinements on this part of the model can be made in the future. For example, implementing the matching operation between the input spectra and the harmonic templates as essentially a simple cross correlation often produces multiple additional estimates at the octaves and other intervals (Duifhuis *et al.*, 1982; Cohen *et al.*, 1995; Goldstein, 1973a). While subjects often report these pitches under experimental scrutiny, pitch percepts in casual listening conditions are usually more unitary (or less analytical), a property that may reflect subsequent integrative processing at higher auditory centers.

The residue pitch portion of our model (Sec. IV) complements the template matching algorithm. It accounts for the phase sensitivity and other properties of pitch percepts that *do not* involve any resolved harmonics such as the pitch of high order harmonics and amplitude modulated noise. However, two important issues remain to be addressed in the future. The first is whether the coincidence algorithm for residue pitch can withstand a critical quantitative scrutiny of its properties, similar to that done earlier for the template matching algorithm (Cohen *et al.*, 1995; Goldstein, 1973a). The second issue concerns the exact nature of the integration of the two pitch percepts. The relevance of this issue stems from the fact that natural stimuli commonly contain simultaneous cues for both pitch percepts (e.g., resolved and unresolved harmonics). Consequently, the final unitary pitch percept must be derived from both available cues.

VI. SUMMARY AND CONCLUSIONS

We have presented a biologically plausible model for forming harmonic templates in the early stages of the auditory system *with broadband noise stimulation, and without need for neural delay lines and other temporal structures*. The model consists of two key operations: a cochlear filtering stage followed by coincidence detection. The cochlear stage provides responses analogous to those seen in the auditory nerve and cochlear nucleus. The second stage is a matrix of coincidence detectors that compute the long-term average of pairwise instantaneous correlation (or products) between responses from all CFs across the channels. Model simulations show that for any broadband stimulus, high coincidences occur between cochlear channels that are exactly harmonic intervals apart. Accumulating coincidences over time results in the formation of harmonic templates for all fundamental frequencies in the phase-locking frequency range. The model explains the critical role played by such important factors in cochlear function as the nonlinear transformations following the filtering stage, the rapid phase-shifts of the traveling wave near its resonance, and the spec-

tral resolution of the cochlear filters. More specifically, the following items summarize the major findings of the model:

(1) *Phase-locking of the filter responses* is a critical factor in the template formation. All templates are ultimately derived from the fine-time structure of the filter responses. Thus the gradual loss of phase locking (or synchrony) to higher frequencies (approximately $>2-3$ kHz) is partially the reason why they are not represented in the templates, and hence play little or no role in the perception of periodicity pitch.

(2) *Nonlinear transformation of the filter responses* is essential in generating the “distortion” harmonics that ultimately form the templates. Half-wave rectification and increasing temporal synchrony are two such transformations.

(3) *High spectral resolution* improves the representation of the harmonic peaks in the templates. Broadening the analysis filters smears the representation of the higher order harmonics considerably.

(4) *Phase delays of the traveling wave* provide locally phase shifted copies of the responses at each CF, which in turn insures there is always a pair of channels at harmonic CFs that can be highly correlated regardless of the phase of the stimulus harmonics.

(5) *The exact nature of the sound stimulus is immaterial* for the formation of the harmonic templates and their parameters. Instead they are solely determined by the intrinsic properties of the cochlear filters and coincidences. That is, given enough time, the same templates will emerge for any broadband stimulus whether it is noise, harmonic sequences, or impulses.

ACKNOWLEDGMENTS

We are grateful to Dr. Steve Greenberg and Dr. Ray Meddis for extensive reviews and advice. This work was supported in part by a grant from the Office of Naval Research under the ODDR&E MURI97 Program to the Center for Auditory and Acoustic Research, and the National Science Foundation under the Learning and Intelligent Systems Initiative Grant No. CMS9720334.

¹The use of harmonic templates in Wightman's model is somewhat less obvious than in others. Wightman's model utilizes a linear frequency axis to represent the spectrum, and then performs a cepstral analysis (Oppenheim and Schaffer, 1976) to compute the inter-harmonic distances (and hence the pitch). Therefore, the harmonic templates used are (implicitly) the Fourier basis functions used in the cepstral analysis.

²The distinction we make here between “spectral” and “temporal” pitch theories is not universally accepted. Our use of the terms here is rather specific. We define “spectral pitch theories” as those that take as their starting point a spectral pattern from which the pitch is computed *regardless* of the nature of the preceding cues and mechanisms used to extract this spectral pattern. For example, it is immaterial to our definition whether the spectrum is extracted from the average auditory-nerve firing rates (Sachs and Young, 1979), time-intervals (Seneff, 1988; Ghitza, 1988), or other cues (Young and Sachs, 1979; Shamma, 1985b). “Temporal pitch theories” do not make use of a spectral pattern; instead, they utilize temporal response features from each auditory-nerve fiber independently of other channels. Therefore, unlike in “spectral” theories, tonotopic order plays no role in “temporal” theories; i.e., shuffling around the auditory-nerve fiber array destroys the spectral pattern, but makes no difference to the temporal patterns on each channel.

³I am indebted to Dr. Steve Greenberg for bringing to my attention an interesting link between the rapid phase shifts of the cochlear traveling

wave (discussed earlier in Sec. II and Fig. 5) and the *Huggins pitch* (Cramer and Huggins, 1958). The *Huggins pitch* is a "tonal" percept created by local phase shifts in the spectrum of otherwise uncorrelated noise presented to the two ears. The saliency of the percept is crucially dependent on the spectral interval over which the phase shifts occur—it is best when they occur over an interval of 3–6 percent of the center frequency of the relevant filter. This estimate may reflect the region over which the rapid phase shifts occur, and indirectly measures the spectral resolution expected from such percepts (i.e., the smallest distance between two locally imposed phase shifts). Bilsen (1977) has demonstrated that multiple phase shifts create a multi-tonal percept which, if harmonically related, evoke a missing fundamental pitch percept. As Bilsen points out (Bilsen, 1977), these results are consistent with a template matching pitch algorithm in which the input spectrum is derived centrally from the binaural inputs which displays peaks corresponding to the tonal Huggins pitch percepts. How this spectrum is derived is beyond the scope of this paper, but traditional binaural coincidence algorithms can readily perform this task.

- Bilsen, F. (1977). "Pitch of noise signals: Evidence for a central spectrum," *J. Acoust. Soc. Am.* **61**, 150–161.
- Cariani, P., and Delgutte, B. (1996a). "Neural correlates of the pitch of complex tones. i: Pitch and pitch salience," *J. Neurophysiol.* **76**, 1698–1716.
- Cariani, P., 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.
- Carlyon, R. (1998a). "Comments on a unitary model of pitch perception," *J. Acoust. Soc. Am.* **104**, 1118–1121.
- Carlyon, R. (1998b). "The effects of resolvability on the encoding of fundamental frequency by the auditory system," in *Psychophysical and Physiological Advances in Hearing*, edited by A. R. Palmer, A. Rees, A. Q. Summerfield, and R. Meddis (Whurr, London).
- Clarkson, M., and Rogers, E. (1995). "Infants require low-frequency energy to hear the pitch of the missing fundamental," *J. Acoust. Soc. Am.* **98**, 148–154.
- Cohen, M., Grossberg, S., and Wyse, L. (1995). "A spectral network model of pitch perception," *J. Acoust. Soc. Am.* **98**, 862–879.
- Colburn, S., and Durlach, N. (1978). "Models of binaural interactions," in *Handbook of Perception*, edited by E. Carterette and M. Friedman (Academic, New York), Vol. IV.
- Cramer, E., and Huggins, W. (1958). "Reaction of pitch through binaural interactions," *J. Acoust. Soc. Am.* **30**, 413–417.
- de Boer, E. (1976). "On the residue in hearing and auditory pitch perception," in *Handbook of Sensory Physiology*, edited by W. Keidel and D. Neff (Springer-Verlag, Berlin), Vol. III, pp. 479–583.
- de Cheveigne, A. (1998). "Cancellation model of pitch perception," *J. Acoust. Soc. Am.* **103**, 1261–1271.
- de Cheveigne, A., McAdams, S., and Marin, C. (1995). "Concurrent vowel identification. ii: Effect of phase, harmonicity, and task," *J. Acoust. Soc. Am.* **101**, 2848–2856.
- Duifhuis, H., Willems, L., and Sluyter, R. (1982). "Measurement of pitch in speech: An implementation of Goldstein's theory of pitch perception," *J. Acoust. Soc. Am.* **71**, 1568–1580.
- Evans, E. (1978). "Place and time coding in the peripheral auditory system: Some physiological pros and cons," *Audiology* **17**, 369–420.
- Evans, E., and Zhao, W. (1998). "Periodicity coding of the fundamental frequency of harmonic complexes. Physiological and pharmacological study of onset units in the ventral cochlear nucleus," in *Psychophysical and Physiological Advances in Hearing. Proceedings of the 11th International Symposium on Hearing*, edited by A. R. Palmer, A. Rees, A. Q. Summerfield, and R. Meddis (Whurr Publishers, London), pp. 186–194.
- Ghitza, O. (1988). "Temporal non-place information in the auditory-nerve firing patterns as a front-end for speech recognition in a noisy environment," *J. Phonetics* **16**, 109–204.
- Goldstein, J. (1973a). "An optimum processor theory for the central formation of pitch of complex tones," *J. Acoust. Soc. Am.* **54**, 1496–1516.
- Goldstein, J. (1973b). "An optimum processor theory for the central formation of the pitch of complex tones," *J. Acoust. Soc. Am.* **54**, 1496–1516.
- Greenberg, S., Geisler, C., and Deng, L. (1986). "Frequency selectivity of single cochlear nerve fibers based on the temporal response patterns of two-tone signals," *J. Acoust. Soc. Am.* **79**, 1010–1019.
- Greenberg, S., Poeppel, D., and Roberts, T. (1998). "A space-time theory of pitch and timbre based on cortical expansion of the cochlear travelling-wave delay," in *Psychophysical and Physiological Advances in Hearing. Proceedings of the 11th International Symposium on Hearing*, edited by A. R. Palmer, A. Rees, A. Q. Summerfield, and R. Meddis (Whurr Publishers, London).
- Hattori, T., and Suga, N. (1997). "The inferior colliculus of the mustached bat has the frequency-vs-latency coordinates," *J. Comp. Physiol. A* **180**, 271–284.
- Hewitt, J., and Meddis, R. (1994). "A computer model of amplitude-modulation sensitivity of single units in the inferior colliculus," *J. Acoust. Soc. Am.* **95**, 2145–2159.
- Jeffress, A. (1948). "A place theory of sound localization," *J. Comp. Physiol. Psychol.* **61**, 468–486.
- Langner, G. (1992). "Periodicity coding in the auditory system," *Hear. Res.* **6**, 115–142.
- Langner, G., and Schreiner, C. (1988). "Periodicity coding in the inferior colliculus of the cat," *J. Neurophysiol.* **60**, 1805–1822.
- Licklider, J. (1951). "A duplex theory of pitch perception," *Experientia* **7**, 128–133.
- Lyon, R., and Shamma, S. (1996). "Auditory representation of timbre and pitch," in *Auditory Computations*, edited by H. Hawkins, E. T. McMullen, A. Popper, and R. Fay (Springer-Verlag, Berlin), pp. 221–270.
- Meddis, R., and Hewitt, J. (1991). "Virtual pitch and phase sensitivity of a computer model of the auditory periphery. i: Pitch identification," *J. Acoust. Soc. Am.* **89**, 2866–2882.
- Montgomery, C., and Clarkson, M. (1997). "Infants' pitch perception: Masking by low- and high-frequency noises," *J. Acoust. Soc. Am.* **102**, 3665–3672.
- Moore, B. (1989). *An Introduction of the Psychology of Hearing*, 3rd ed. (Academic, London).
- Moore, B. C. J. (1986). *Frequency Selectivity in Hearing* (Academic, London), Chap. 5.
- Oertel, D., Wu, S., and Dizack, C. (1990). "Morphology and physiology of cells in slice preparation of posterovental cochlear nucleus of mice," *J. Comp. Neurol.* **295**, 136–154.
- Oppenheim, A., and Schaffer, R. (1976). *Digital Signal Processing* (Prentice-Hall, New Jersey).
- Palmer, A., Winter, I., Jiang, D., and James, N. (1995). "Across frequency integration by neurons in the ventral cochlear nucleus," in *Advances in Hearing Research*, edited by J. Manley, G. Klump, C. Kopple, H. Fastl, and H. Oeckinghaus (World Scientific, Singapore).
- Patterson, R., and Holdsworth, J. (1991). "A functional model of neural activity patterns and auditory images," in *Advances in Speech, Hearing and Language Processing*, edited by W. A. Ainsworth (JAI Press, London), Vol. 3.
- Plomp, R. (1976). *Aspects of Tone Sensation* (Academic, New York).
- Rhode, W. (1994). "Lateral suppression and inhibition in the cochlear nucleus of the cat," *J. Neurophysiol.* **71**, 493–519.
- Rhode, W. (1995). "Interspike intervals as a correlate of periodicity pitch in cat cochlear nucleus," *J. Acoust. Soc. Am.* **97**, 2414–2429.
- Ruggero, M. (1973). "Response to noise in auditory nerve fibers in squirrel monkey," *J. Neurophysiol.* **36**, 569–587.
- Sachs, M. B., and Young, E. D. (1979). "Encoding of steady state vowels in the auditory-nerve: Representation in terms of discharge rate," *J. Acoust. Soc. Am.* **66**, 470–479.
- Schouten, J. (1940). "The residue and the mechanism of hearing," *Proc. K. Ned. Akad. Wet.* **43**, 991–999.
- Schreiner, C., and Langner, G. (1988). "Periodicity coding in the inferior colliculus of the cat. ii. topographical organization," *J. Neurophysiol.* **60**, 1823–1840.
- Schreiner, C., and Urbas, J. (1988). "Representation of amplitude modulation in the auditory cortex of the cat. i: The anterior field," *Hear. Res.* **21**, 227–241.
- Schwartz, D., and Tomlinson, R. (1990). "Spectral response patterns of auditory cortex neurons to harmonic complex tones in alert monkey (*macaca mulatta*)," *J. Neurophysiol.* **64**, 282–299.
- Seneff, S. (1988). "A joint synchrony/mean-rate model of auditory processing," *J. Phonetics* **85**, 55–76.
- Shamma, S. (1985a). "Speech processing in the auditory system: I. representation of speech sounds in the responses of the auditory nerve," *J. Acoust. Soc. Am.* **78**, 1612–1621.
- Shamma, S. (1985b). "Speech processing in the auditory system: II. lateral inhibition and the central processing of speech evoked activity in the auditory nerve," *J. Acoust. Soc. Am.* **78**, 1622–1632.

- Shamma, S., Chadwick, R., Wilbur, J., Morrish, K., and Rinzel, J. (1986). "A biophysical model of cochlear processing: Intensity dependence of pure tone responses," *J. Acoust. Soc. Am.* **80**, 133–145.
- Shamma, S., and Morrish, K. (1986). "Synchrony suppression in complex stimulus responses of a biophysical model of the cochlea," *J. Acoust. Soc. Am.* **81**, 1486–1498.
- Shamma, S., Shen, N., and Gopalaswamy, P. (1989). "Stereausis: Binaural processing without neural delays," *J. Acoust. Soc. Am.* **86**, 989–1006.
- Slaney, M., and Lyon, R. (1993). "On the importance of time—A temporal representation of sound," in *Visual Representations of Speech Signals*, edited by M. Cooke, S. Beet, and M. Crawford (Wiley, New York).
- Smith, P., and Rhode, W. (1989). "Structural and functional properties distinguish two types of multipolar cells in the cat ventral cochlear nucleus," *J. Acoust. Soc. Am.* **282**, 595–616.
- Summerfield, A., and Assmann, P. (1990). "Modelling the perception of concurrent vowels: Vowels with different fundamental frequencies," *J. Acoust. Soc. Am.* **88**, 680–697.
- Terhardt, E. (1974). "Pitch consonance and harmony," *J. Acoust. Soc. Am.* **55**, 1061–1069.
- Wang, K., and Shamma, S. A. (1994). "Self-normalization and noise-robustness in early auditory representations," *IEEE Trans. Speech Audio Process.* **2**, 421–435.
- Wightman, F. (1973). "A pattern transformation model of pitch," *J. Acoust. Soc. Am.* **54**, 397–406.
- Winter, I., Wiegube, L., and Patterson, R. (1999). "Encoding iterated ripple noise and harmonic complexes in the ventral cochlear nucleus," in *Abstr. 552, Assn. Res. Otol. Annual Meeting*.
- Yang, X., Wang, K., and Shamma, S. A. (1992). "Auditory representations of acoustic signals," *IEEE Trans. Inf. Theory, Special Issue on Wavelet Transforms and Multiresolution Signal Analysis* **38**, 824–839.
- Yost, W., and Hill, R. (1979). "Model of the pitch and pitch strength of ripple-noise," *J. Acoust. Soc. Am.* **66**, 400–410.
- Young, E., and Sachs, M. (1979). "Representation of steady-state vowels in the temporal aspects of the discharge patterns of populations of auditory-nerve fibers," *J. Acoust. Soc. Am.* **66**, 1381–1403.

A model of the perceptual asymmetry between peaks and troughs of frequency modulation

Alain de Cheveigné

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 619-02, Japan

(Received 15 January 1998; revised 3 March 1999; accepted 14 January 2000)

Pitch discrimination at peaks of frequency modulation is better than at troughs [L. Demany and K. I. McAnally, *J. Acoust. Soc. Am.* **96**, 706–715 (1989)]. A similar asymmetry emerges within a time-domain pitch perception model based on autocorrelation. The model requires the following assumptions: (a) The neural discharge patterns must be temporally sharpened to a single narrow pulse per period (possibly by neural convergence within the cochlear nucleus). (b) Autocorrelation must be implemented as a cross correlation between the neural pulse train and a delayed pulse train convolved with a short kernel function. This kernel function must be asymmetric in time. (c) Pitch discrimination must rely on higher-order modes of the autocorrelation function. This particular implementation of the autocorrelation model produces modes that are sharper for peaks than for troughs, and thus accounts for the pitch discrimination asymmetry observed experimentally. As a by-product it can account for “hyperacute” discrimination observed at peaks of triangular modulation. © 2000 Acoustical Society of America. [S0001-4966(00)05104-3]

PACS numbers: 43.66.Ba, 43.66.Hg, 43.64.Bt [JWH]

INTRODUCTION

Demany and McAnally (1994) discovered an interesting perceptual asymmetry between peaks (local maxima) and troughs (local minima) of frequency modulation: the *perceptual salience* of peaks is greater than that of troughs. In a first experiment, subjects were presented with a frequency-modulated pure tone with a modulating waveform that had as many peaks as troughs. When requested to report the sequence of notes that they heard, subjects consistently reported maxima more often than minima. The asymmetry could not be ascribed to a mere advantage of high frequencies over low: in some conditions certain peaks were *lower* in frequency than certain troughs, but nevertheless more salient.

In a second experiment it was found that increased salience of peaks over troughs is reflected also by better *frequency discrimination*. Demany and his colleagues subsequently carried out an extensive set of experiments that showed that the peak/trough asymmetry is quite general. The many conditions that they explored allowed a wide range of hypotheses to be tested and, for many, rejected. These data are reviewed in detail in the next section. The ubiquity and strength of the peak/trough asymmetry (threshold ratios averaged over subjects as large as 5 in some experiments) suggest that it might reflect some important property of the pitch perception system. The later sections of this paper propose a model of this effect.

I. REVIEW OF THE EXPERIMENTAL DATA

In their second experiment, Demany and McAnally (1994) presented subjects with tones modulated by a wave-

form shaped either as a single cycle of a cosine (“trough”), or as the opposite of the same (“peak”). The modulator, defined on a logarithmic scale, had an amplitude of $\frac{1}{2}$ oct. For peaks, the initial frequency was 707 Hz, and for troughs 1414 Hz, but the vertex frequency, at the stimulus center, was about 1000 Hz in both cases [Fig. 1(a)]. In the present paper, “vertex” is used to designate the extremum of frequency modulation, either maximum or minimum, upon which the pitch discrimination task is focused. The subjects were requested to discriminate these stimuli from stimuli that started and stopped at the same frequency, but had a deeper frequency modulation, and thus a slightly different frequency at the vertex. Discrimination was clearly better for peaks than for troughs: thresholds for troughs were about twice those of peaks.

Demany and his colleagues have shown the phenomenon to be quite general. The asymmetry has been observed for modulated pure tones with vertex frequencies of 250, 500, and 1000 Hz (Demany and Clément, 1995a, 1998). It also exists at 4000 Hz, but the trough/peak threshold ratio is smaller (Demany and Clément, 1995a). It exists for harmonic complex tones with vertex frequencies of 200 or 500 Hz (Demany and Clément, 1995b, 1998), and in particular for Shepard tones that consist of an infinite series of octave-spaced partials shaped by a fixed band-limited spectral window (Shepard, 1964). The case of Shepard tones is interesting because it allows an ambiguity in the interpretation of the asymmetry to be resolved. Peak and trough stimuli have similar frequencies at the vertex, but the frequency at the beginning of the stimulus is an octave higher for the trough than for the peak (for a modulation amplitude of 0.5 oct), and the same is of course true at the end. Perceptual asymmetry

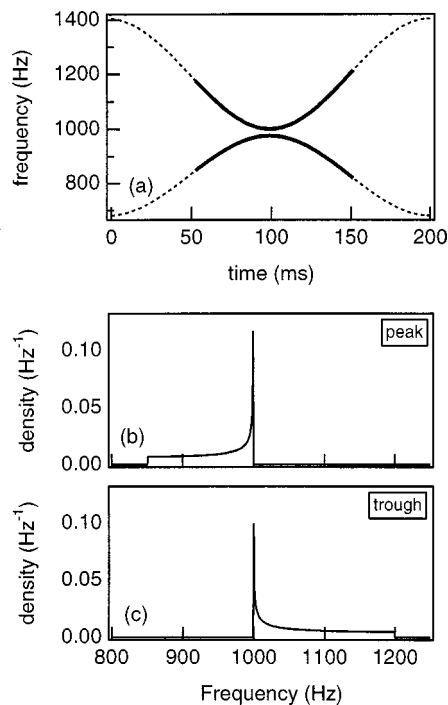


FIG. 1. (a) Top: modulation waveform of a trough. Bottom: modulation waveform of a peak. (b) Instantaneous frequency distribution of a peak over the portion marked as a full line in (a). (c) Same, for a trough.

could conceivably result from that aspect of the stimuli. Shepard tones spaced an octave apart are identical, so this interpretation can be ruled out in their case (and probably in general). The asymmetry should be attributed to other cues such as the shape of the vertex or the direction of the FM glides that flank it.

The asymmetry has been observed for modulation amplitudes of 0.5, 1, or 1.5 oct (in the latter case for unidirectional glides, Demany and Clément, 1997). When the vertex frequency was roved within a 0.5-oct range the trough/peak threshold ratio was only slightly reduced at 250 or 1000 Hz, but practically abolished at 4000 Hz. The asymmetry is not affected by training despite a general decrease in thresholds (Demany and Clément, 1995a), but it seems to differ somewhat between subjects. It has been investigated mainly with stimuli 200 or 400 ms in duration, but it exists for stimuli as long as 800 ms or as short as 50 ms but not 25 ms (Demany, 1997). The effect persisted when an amplitude modulation was superimposed on the FM, even when the AM was designed to favor troughs over peaks (Demany and McAnally, 1994; Demany and Clément, 1998).

The effect has been studied principally with cosine-shaped modulation waveforms [Fig. 1(a)], but it exists also with modulation waveforms that have extrema that are either sharper (almost triangular) or duller (with flat portions as long as 100 ms). The trough/peak ratio tends, however, to be greater for sharp than for blunt extrema (Demany and Clément, 1997). The effect was also observed for ramps consisting of half a cosine modulation waveform, either the first half or the second half (Demany and Clément, 1997).

This impressive sum of data suggests that the perceptual asymmetry between peaks and troughs might reflect fundamental characteristics of the pitch perception process, even if

it has so far resisted explanation in such terms. Several explanations that seem plausible *a priori* can be rejected based on the experimental results. One might, for example, imagine that the general improvement of discrimination with frequency up to 2 kHz, where relative thresholds are smallest (Moore, 1973), might somehow favor peaks over troughs. By showing that the asymmetry was not reversed at 4000 Hz, Demany and Clément (1995a) ruled out this explanation. Demany and McAnally (1994) had already shown that the effect could not be due to a mere change of amplitude or loudness with frequency (FM-induced AM).

The effect is not likely to be due to the asymmetry of excitation patterns on the basilar membrane. These should, if anything, favor troughs over peaks, as there is an upward spread of masking for pure tones at 70 dB (Demany and McAnally, 1994). The argument was strengthened by Demany and Clément (1995a) when they replicated the experiment at 35 dB, at which level the asymmetry of the excitation pattern should be opposite of that at 70 dB. It was further strengthened by the observation of the same asymmetry for complex harmonic tones (in particular tones shaped by a fixed spectral window). The excitation pattern of a complex tone is more complex and spread out than that of a pure tone, and less likely to differ between peaks and troughs in a way that could favor the former over the latter.

The fact that the asymmetry is observed for amplitude-modulated high-pass noise, that lacks spectral cues, suggests that it stems from a process operating in the time domain (Demany and Clément, 1995b). This is congruent with the fact that it is weaker at 4000 Hz, or even nonexistent for roving stimuli at that frequency (Demany and Clément, 1995a). On the other hand, Demany and Clément (1997) have also argued *against* a time-domain process, on the basis of the proactive and retroactive effects that they observed with ramps. For those authors a retroactive effect suggests a central mechanism, whereas high-resolution time-domain information is not usually observed beyond the periphery and neural relays below the inferior colliculus.

The fact that the asymmetry was observed with Shepard tones rules out any role of the initial (or final) frequency difference between peak and trough stimuli, as Shepard tones differing by an octave are identical. It also rules out a simple effect of the temporal pattern of extreme frequencies (low-high-low versus high-low-high). The asymmetry must therefore depend on dynamic frequency changes, for example, the frequency ramps preceding or following the temporally centered vertex. Demany and Clément (1995b, 1997) argued that discrimination cannot be based on differences in FM *slope* between standard and target, at least in the case of peaks: the differences are too small compared to FM slope thresholds measured by Dooley and Moore (1988). Overall, it is unlikely that the effect depends on a perceptual asymmetry of the *ramps* by themselves, but rather on some interaction between ramp and peak.

A priori, one can imagine either a *proactive* effect of the ramp preceding the vertex, or a *retroactive* effect of the ramp following it, or both. By cutting the cosine-modulated stimulus in two (either the first half or the second), Demany and Clément (1997) were able to demonstrate that the effect of a

ramp on the discrimination of a vertex can be *both* proactive and retroactive. Is the effect beneficial to peaks, or detrimental to valleys, or both? Demany and Clément (1995b) suggested a detrimental effect on valleys, at least for amplitude-modulated noise. However, Demany and Clément (1997) found, for triangular-shaped modulation, “hyperacute” perception at the peak. This latter effect is surprising: shift detection thresholds of practically durationless peaks were about 10 cents or 0.6%. This is of the same order as the threshold of a 6.25-ms steady-state pure tone at 1 kHz (Moore, 1973). Frequency would vary by about 18 cents within a 6-ms window centered on the peak (and by the double for a window situated on either side), and it is difficult to understand how “moving targets” of the sort can be discriminated with such acuity.

Clément (1996) proposed a model to explain the asymmetry, based on a formal neural network. The input layer received a tonotopic representation of the instantaneous frequency of the stimulus (derived by an infinite-resolution frequency extraction process that was not described). The cells in this layer were interconnected laterally by asymmetric inhibitory connections with temporal integration properties. This layer favored the representation of frequencies preceded by lower frequencies, at the expense of those preceded by higher frequencies. The second layer served to integrate temporally this modified frequency representation. Asymmetric lateral inhibition within the first layer accounts for the perceptual advantage of peaks over troughs, while the temporal integration of the second layer accounts for the fact that effects can be both “proactive” and “retroactive.”

Clément’s explanation of the asymmetry is plausible, but not entirely satisfying. In a sense, it is a direct translation of the phrase “the auditory system favors peaks over troughs” by which one might summarize the experimental results. It offers a plausible account of *how* such an asymmetry might be implemented, but not *why* things should occur in that way. An argument for the functional advantage of the asymmetric lateral inhibition, or an explanation in terms of *emergent* properties of the pitch perception process, would be more satisfying (though not necessarily more correct). The purpose of the present paper is to propose an explanation of the latter sort.

II. THE MODEL

A. Structure

The model is designed to produce a quantity (“pitch estimate”) that allows the pitch of two stimuli to be compared, for example, between two intervals of a discrimination task. Pitch asymmetries can be explained if the pitch estimate is less accurate for valleys than for peaks. The overall structure of the model is illustrated in Fig. 2. It is similar to the autocorrelation model of Licklider (1956) or Meddis and Hewitt (1991a, b), but differs in three important ways. First, the discharge probability at the output of the basilar-membrane/haircell model is sharpened before further processing. Second, the standard autocorrelation function (ACF) of Licklider’s model is replaced by an “asymmetric autocorrelation function” (AACF, defined presently). Third, the

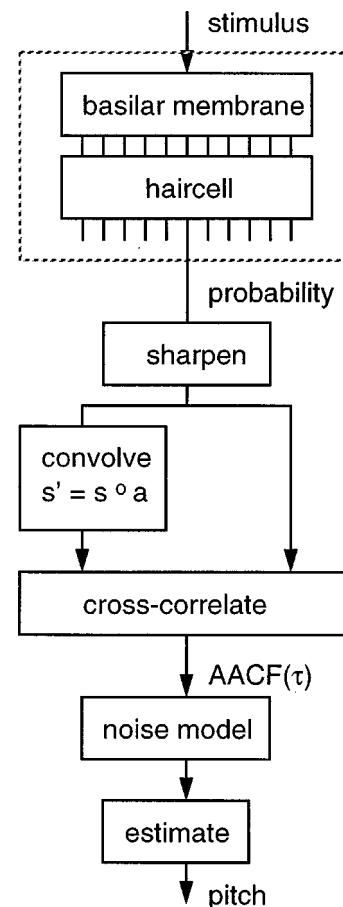


FIG. 2. Overall structure of the model. The part within dotted lines is implemented simply as a half-wave rectifier. Its output is first sharpened (by raising it to a power) and then cross correlated with a version of itself convolved with an asymmetric function, $a(t)$, to obtain an “asymmetric autocorrelation function” (AACF). After application of a noise model, the period is estimated from a high-order mode of the AACF.

pitch estimate involves higher-order modes of that function, rather than just the first-order “pitch peak.” These assumptions are detailed in the next paragraphs, after the presentation of a “lemma” that embodies the principle on which the model is based. All assumptions are examined critically in Sec. IV.

B. Lemma: Convolution of asymmetric peaked functions

The purpose of this section is to point out that the sharpness of the *convolution* of two functions that are sharp and asymmetric depends on whether the asymmetries agree or not. The “lemma” is the root of the model. No formal proof is attempted; the lemma is simply illustrated for two windowed exponential functions (Fig. 3). On the left, the asymmetries of the functions to be convolved agree. Their convolution has a peak that is somewhat blunt. On the right, one function has been time reversed, and their asymmetries are thus opposed. Their convolution has a sharp peak, related to the existence and orientation of the discontinuities in the functions. If the position of the maximum had to be determined precisely, for example, in the presence of a small amount of noise, it is clear that accuracy would be greater on the right than on the left.

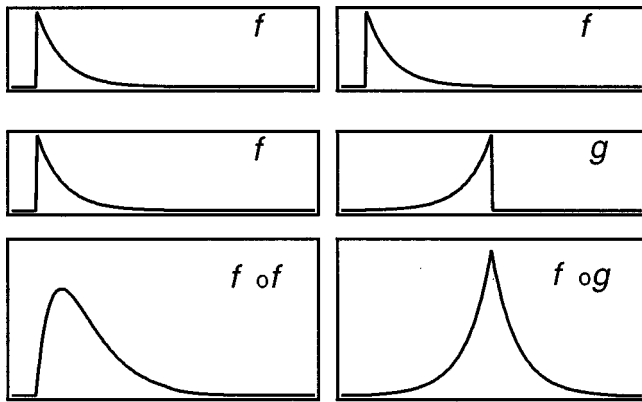


FIG. 3. Top and middle left: two asymmetric peaked functions with similar orientations. Bottom left: their convolution. Top and middle right: two asymmetric peaked functions with opposite orientations. Bottom right: their convolution. Note the difference in “peakiness” between the two convolutions.

The strategy followed by this paper relies on this lemma: (a) identify appropriate asymmetric functions within the auditory process, (b) show that the asymmetry of one of these functions is reversed for peaks and valleys of modulation, and (c) argue that pitch depends upon their convolution. One asymmetric function can be found in the distribution of instantaneous period (inverse of instantaneous frequency) in the region of the vertex [Fig. 1(b) and (c)]. The asymmetry of this distribution is opposite for peaks and valleys. The second asymmetric function will be introduced by assumption 2, within the autocorrelation calculation. Assumptions 1 and 3 are of less fundamental importance, but they make the model easier to illustrate, and allow it to handle a wider range of conditions.

C. Assumptions

Assumption 1: Temporal sharpening of neural responses. The discharge probability of an auditory nerve fiber generally resembles the half-wave rectified motion of the basilar membrane. In response to a pure tone, the probability function consists of peaks that occupy a large proportion of the period, and are therefore quite wide (Ruggero, 1992) [Fig. 4(a)]. Peaks of the autocorrelation function (ACF) are also wide [Fig. 4(b)], and the auditory system might have difficulty deriving an accurate pitch estimate on this basis. Probability functions may, however, be sharpened, for example, within onset cells in the cochlear nucleus (Palmer and Winter, 1992), as illustrated in Fig. 4(c). As a result of this sharpening, peaks of the ACF are also much sharper [Fig. 4(d)]. *Assumption 1* is that a sharpening stage is involved in the pitch perception process. A full discussion of the necessity and plausibility of this assumption is deferred to Sec. IV A.

Assumption 2: Asymmetric implementation of autocorrelation. The autocorrelation function (ACF) is, by definition, the cross correlation of a signal with itself. The product of two terms, the signal $s(t)$ and the delayed version $s(t - \tau)$, is summed over a window $w(t)$:

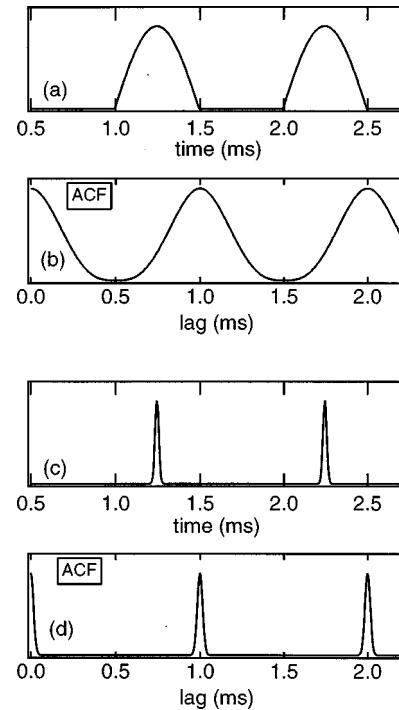


FIG. 4. (a) Discharge probability in response to a 1-kHz pure tone. (b) Autocorrelation function (ACF) of discharge probability. (c) Sharpened discharge probability function, obtained by raising the waveform of (a) to the power 200. (d) ACF of sharpened discharge probability function.

$$\text{ACF}_t(\tau) = \int_{-\infty}^t w(\theta - t) s(\theta) s(\theta - \tau) d\theta, \quad (1)$$

where θ is a temporal integration variable. This is actually the definition of a running autocorrelation function indexed by t . Suppose now that the delayed function $s(t - \tau)$ is convolved with a kernel function $a(t)$ before multiplication:

$$s' = s \circ a. \quad (2)$$

The kernel function $a(t)$ is supposed to be short and asymmetric: skewed with a tail towards large t . Let us call “asymmetric autocorrelation function” (AACF) the running cross-correlation function between s and s' :

$$\text{AACF}_t(\tau) = \int_{-\infty}^t w(\theta - t) s(\theta) s'(\theta - \tau) d\theta. \quad (3)$$

Figure 5 illustrates this process. In (a) is plotted the discharge probability function, sharpened as argued in the previous paragraph. In (b) is shown its convolution with a short kernel function shaped as a decaying exponential of time constant t_a . In (c) is shown the cross correlation of waveforms in (a) and (b), called AACF. Its modes are asymmetric and skewed in the direction of shorter lags. Other than that, the AACF is not too different from the ACF, hence the name “asymmetric autocorrelation function.” *Assumption 2* is that the pitch perception process involves such an asymmetric autocorrelation function, rather than the standard symmetric autocorrelation that is usually postulated by autocorrelation-based models of pitch (Licklider, 1956; Meddis and Hewitt, 1991a, b). This assumption is justified in Sec. IV A.

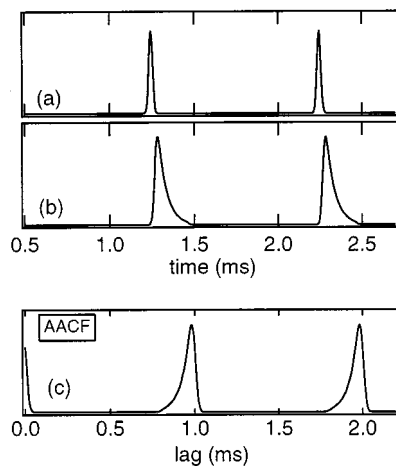


FIG. 5. (a) Sharpened discharge probability function. (b) Same, after convolution with a window shaped as a decaying exponential of time constant 0.25 ms. (c) “Asymmetric autocorrelation function” (AACF) of the sharpened probability function [cross correlation of (a) and (b)].

Assumption 3: Pitch is derived from higher-order modes. In the autocorrelation model of pitch, it is usually assumed that pitch is derived from the first mode (“period peak”) of the ACF. However, it has been argued that the auditory system might take into account higher-order modes (Srulovicz and Goldstein, 1983; de Cheveigné, 1989; Slaney, 1990). Higher-order modes have a smaller *relative* width, and this helps account for the small size of discrimination thresholds. The maximum order is limited by the stimulus duration, so the use of higher-order modes might explain the approximate inverse dependency on stimulus duration observed by Moore (1973) below 1–2 kHz. *Assumption 3* is that higher-order modes of the ACF (or AACF) are exploited in the task of discriminating the vertex frequency of peaks or troughs.

D. Modeling conventions

The following conventions are made in the interest of clarity. Basilar membrane filtering is ignored and the properties of hair-cell transduction are reduced to a simple half-wave rectification. The integration window of the running ACF and AACF is chosen to be square (rather than exponential as in the models of Licklider, 1956, or Meddis and Hewitt, 1991a, b):

$$\text{AACF}_t(\tau) = \int_{t-D}^t s(\theta)s'(\theta - \tau) d\theta, \quad (4)$$

where D is the duration of the window. The square shape allows the position of the window to be defined more precisely, so the effects of window size and position can be better understood. Note that, for nonzero values of τ , the calculation includes samples that lie *outside* the window. In signal processing the signal is often set to zero outside the integration window, to avoid the contribution of such samples. That custom is not followed here.

In the simulations, the duration D is chosen long relative to the period in order to avoid rapid fluctuations of the period estimate, rather than on the basis of physiological or psycho-

physical estimates. This is necessary for the model to play its role. Note that it implies some circularity, as the period is not known before estimation. This consideration determines the *lower* limit of the window size. The upper limit is treated as a free parameter, as is the position of the window within the stimulus.

The convolution kernel function $a(t)$ used in the AACF is taken to be shaped like a decaying exponential. Its time constant t_a must be small relative to the period, but otherwise it is treated as a free parameter.

Pitch is supposed to depend on the position of the *maximum* of one or more modes of the ACF (or AACF), rather than on, say, their center of gravity or positions of maximum slope. The accuracy with which this position can be determined is assumed to depend on the “sharpness” of the mode (inverse of the width at some proportion of the height). This assumption is reasonable if noise is superimposed on the function.

III. BEHAVIOR OF THE MODEL

To set the stage, the simple case of an unmodulated tone is considered first. Then the model is applied to the frequency-modulated tones used by Demany and colleagues. In every case, two versions of the model are illustrated: standard autocorrelation (ACF) and asymmetric autocorrelation (AACF).

A. Unmodulated tone

Figure 6(a) represents the discharge probability evoked by a 1-kHz pure tone of duration 10 ms. The representation is idealized (no transitory effects of filtering or haircell transduction). Figure 6(b) shows the ACF of the probability function in (a). The integration window size was 5 ms, and it was placed to cover the end of the stimulus. Had it been chosen as large as (or larger than) the stimulus duration, the envelope of the ACF would have been triangular. Had it been chosen shorter, for example, equal to 1 ms, the envelope would have been flatter (had it been chosen shorter still, ACF_t would have fluctuated strongly with t). Whatever the window size, the ACF is zero for lags greater than the stimulus duration.

Figure 6(c) shows the AACF in response to the same tone. The time constant of the convolution kernel function was $t_a = 0.25$ ms. The integration window was the same as for the ACF. The AACF has the same envelope as the ACF, and can be calculated for the same range of lags. However, each peak is skewed, as in Fig. 5(c).

The purpose of this example was to provide a simple reference to help understand the more complex patterns obtained with modulated stimuli. The stimulus was chosen relatively short to illustrate the fact that the running autocorrelation is limited to lags shorter than the stimulus duration. It also shows how the envelope of the ACF or AACF depends on the shapes of the stimulus envelope and integration window.

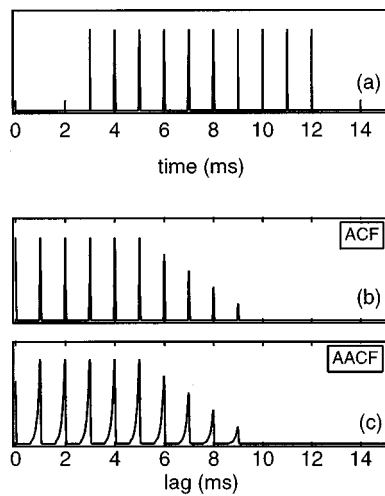


FIG. 6. (a) Idealized discharge probability evoked by a 10-ms pure tone of frequency 1 kHz. Filtering and transitory effects are ignored in this simulation. (b) The ACF of the discharge, for positive lags. The integration window covered the last 5 ms of the stimulus. If it had covered the entire stimulus, the ACF would have had a triangular envelope. (c) The AACF for the same stimulus, with the same integration window. The time constant of the convolution function in the AACF was 0.25 ms.

B. Cosine-shaped peaks and troughs of modulation

The stimulus is a 200-ms frequency-modulated pure-tone, such as used by Demany and McAnally (1994). The modulation waveform is shaped as one period of a cosine function (valley) or its opposite (peak) [Fig. 1(a)]. Its amplitude is 0.5 oct and the frequency at the vertex is 1 kHz.

Figure 7(a) shows the standard ACF for a peak. It was calculated with a window of duration $D=40$ ms placed before the modulation vertex (i.e., covering the last 40 ms of the ramp preceding the vertex). The effect of window position is discussed further on. The modes are widened as a result of modulation, and their width increases with mode order. They are also highly skewed, with a sharp peak and a tail towards longer lags. This reflects the nonuniform instantaneous frequency distribution a consequence of the rounded shape of the peak of the modulation function.

Figure 7(b) shows the standard ACF for a valley. The integration window was placed as in (a). The modes are skewed as in (a) but in the opposite direction. They are, however, just as sharp: if pitch were derived from the position of the modes of the ACF, there would be no reason for discrimination to be worse for valleys than for peaks of modulation.

Figure 7(c) shows the AACF for a peak. It was calculated using the same integration window as for Fig. 7(a). The time constant of the convolution kernel function was $t_a=0.25$ ms. The modes are wider at their base than those of the ACF (an expected effect of the convolution), but they remain quite sharp at their tips. Low-order modes are skewed to the left, and high-order modes to the right, while intermediate modes are more or less symmetrical: this reflects the interplay of the fixed width of the convolution kernel function with the increasing width of the ACF modes due to modulation.

Figure 7(d) shows the AACF for a valley. The modes

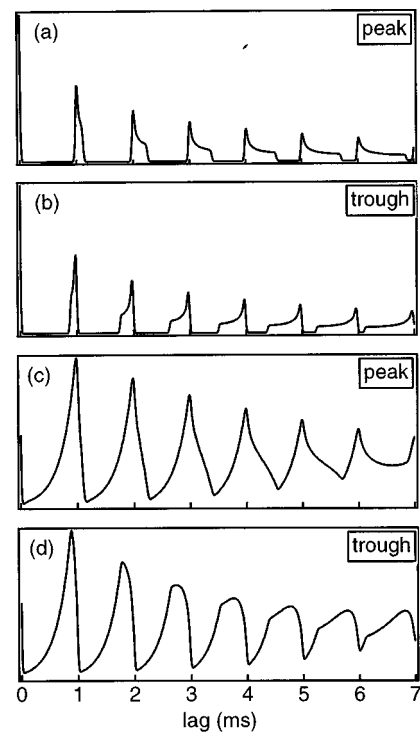


FIG. 7. (a) The ACF for a cosine-shaped modulation peak, calculated over a 40-ms window placed before the vertex. (b) The ACF for a cosine-shaped trough, calculated with a similarly placed window. Peaks are skewed in opposite directions for peaks and valleys, but their sharpness is the same. (c) The AACF for a cosine-shaped peak. (d) The AACF for a cosine-shaped trough. Peaks are rounder for valleys than for peaks, particularly peaks of higher order.

are less sharp than for the peak, especially higher-order modes. This difference can be understood from the “lemma” of Sec. II B. The two functions that, convolved together, produce the mode shape are the *frequency distribution*, on one hand (skewed in different directions for peaks and troughs) and the *exponential-shaped kernel function* of the AACF, on the other. This difference explains, according to the present model, the asymmetry found by Demany and colleagues for cosine-shaped stimuli. As the explanation involves only samples preceding the vertex, it applies also to the asymmetries that they found between end-of-rise and end-of-fall of half cosine-shaped ramps (Demany and Cl  ment, 1997).

1. The effect of window position

The shape of the AACF (and of the ACF) is quite sensitive to window position. Figure 8 (thin line) shows the AACF for a 40-ms window placed *after* a modulation peak (i.e., covering the first 40 ms of the decreasing ramp). The thick line represents the envelope of the AACF for the previous window position (before the apex).

Higher-order modes are sharper and more prominent when the window is placed after the peak rather than before. The difference may seem paradoxical, as windows placed before and after the apex cover similar portions of the waveform. The explanation is simple: the ACF and AACF calculations include samples that *precede* the window. For a window placed before the vertex, these samples belong to the

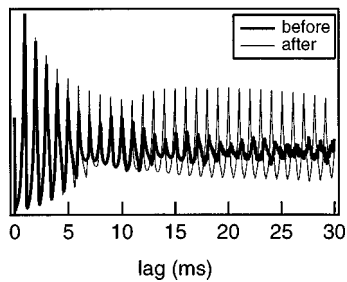


FIG. 8. Thick line: AACF for a cosine-shaped peak, calculated over a 40-ms window placed before the vertex. Thin line: same for a window placed after the vertex.

same slope of the modulation waveform as the window itself, whereas for a window placed after the vertex they belong to the opposite slope. A mode of order k can be understood as a histogram of the durations of the population of interpulse intervals of order k . For a window placed after the vertex, many of these intervals span the vertex, that is, the region where the sweep rate is smallest. In addition, one end of a k th order interval follows a slope of increasing frequency (decreasing period) while the other follows the opposite slope. For both reasons, the spread of k th-order interval durations is reduced when the window follows the vertex. This effect is greatest for large k . Aside from this difference, one expects an asymmetry similar to that observed with a window placed before the apex, because modes in response to a trough (not shown) are less sharp than for a modulation peak.

2. Noise model

The difference in peak shapes between Fig. 7(c) and (d) accounts qualitatively for the peak/trough asymmetry. To derive quantitative predictions of discrimination thresholds, the AACF can be assumed to be corrupted by addition of Gaussian noise with a standard deviation proportional, for each sample of the AACF, to the square root of the value of that sample. This noise model approximates the random accumulation of counts in a histogram bin, such as might occur if the AACF were calculated from neural spike trains within a coincidence network:

$$\text{AACF}'(\tau) = \text{AACF}(\tau) + N_{\sigma} \text{AACF}(\tau)^{0.5}, \quad (5)$$

where AACF' is the noisy pattern, AACF is the noiseless pattern [normalized by division by $\text{AACF}(0)$], and N_{σ} is a Gaussian noise of standard deviation σ and mean 0. It is further assumed that pitch is derived from the *position of the maximum* of the noisy pattern near its mode of order k . Figure 9 (top) shows the simulated standard-deviation-to-mean ratio of this statistic as a function of σ , for $k=3$.

Variability is greater for troughs than for peaks, especially at small values of σ for which it is determined mainly by the shape of the modes near their peaks. For larger values of σ , gross aspects of the mode shape affect the variability which becomes more similar for peaks and troughs. The ratio of variabilities between troughs and peaks therefore decreases as a function of σ , as plotted in Fig. 9 (bottom). It

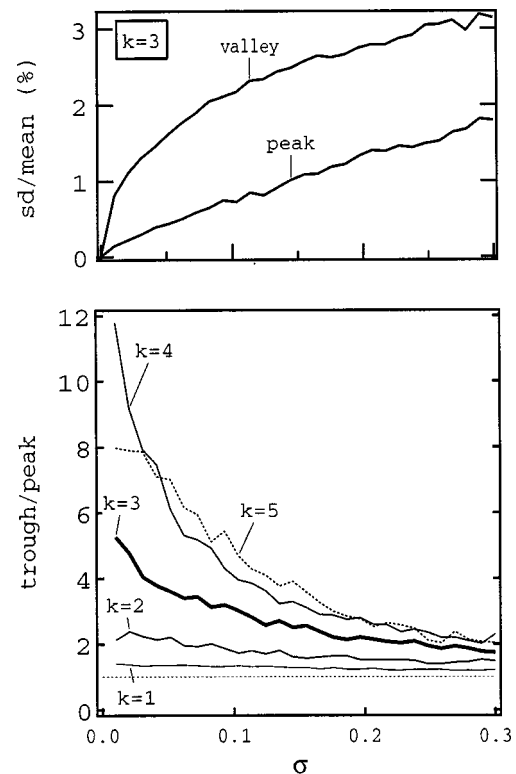


FIG. 9. Top: ratio of standard deviation to mean of pitch estimate as a function of noise standard deviation (σ) for peak and trough. The pitch estimate was derived from the position of the mode of order 3 of the AACF. Bottom: peak/trough ratio of pitch discrimination thresholds for various values of the mode order k .

also depends on the order of the mode: it increases up to $k=4$ or 5 , and decreases slightly beyond that value (not shown).

The simulation used a sampling rate of 100 kHz, and variability was estimated over 5000 repetitions. The vertex frequency was 0.5 kHz. Probability functions were sharpened by raising the output of the haircell model (scaled to 0–1) to the 1000th power. The AACF was calculated with an exponentially shaped kernel of time constant $t_a=0.4$ ms, over a square integration window of duration $D=40$ ms placed before the vertex.

3. Comparison with experimental data

The model has many parameters (degree of sharpening, kernel time constant t_a , window size and position, mode order k , noise magnitude σ , etc.) that affect thresholds. To obtain specific figures, let us arbitrarily choose $\sigma=0.1$ and $k=3$. The variability of the estimate of the mode position can be read from Fig. 9 (top). Supposing that this factor governs pitch discrimination accuracy, relative thresholds are expected to be larger by a factor of $\sqrt{2}$ (Hartmann, 1997). For $\sigma=0.1$, thresholds for peaks should be about 1% (17 cents), and for troughs 3% (50 cents), with a trough/peak threshold ratio of 3. These figures are in rough agreement with those of Demany and McAnally (1994) or Demany and Clément (1995a) for cosine-modulated pure tones.

However, those authors also observed a wide range of interindividual differences. To some extent they can be ac-

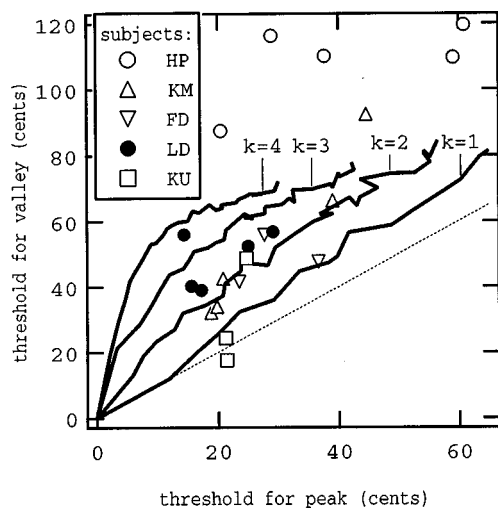


FIG. 10. Points: plot of thresholds for valleys (ordinate) versus peaks (abscissa), for different subjects and sessions of Demany and McAnally (1994). Lines: thresholds predicted by the model for various values of the mode order k .

counted for by assuming different parameter values. Figure 10 shows a scatter plot of experimental data points obtained by Demany and McAnally (1994), together with thresholds predicted by the model for several values of k , and σ ranging between 0 and 0.3. Most data points can be approximated by model predictions. However, a few data points (subject HP) lie beyond the range of model predictions. They could possibly be accounted for by assuming an additional source of variability, but asymmetry is likely to be reduced if this variability affects peaks and troughs equally. It is not clear how this issue can be resolved.

To summarize, the model produces pitch thresholds and peak/trough ratios that are in approximate agreement with experimental data obtained for cosine-modulated pure tones. In principle the explanation is also valid for complex tones, although this was not verified formally. The model would produce similar predictions at other frequencies if one supposed a uniform scaling of its temporal parameters (time constant, window size, etc.). As there is no compelling reason to expect perfectly uniform scaling, the differences across frequency noted by Demany and Clément (1995a) are not unexpected. The next section examines whether the model can be generalized to two other forms of modulation used by Demany and colleagues: triangular-shaped modulation and half-cosine ramps.

C. Triangular peaks and troughs of modulation

So far the reasoning relied on the fact that the frequency distribution is not uniform near the vertex of cosine-shaped modulation. This is no longer true if the modulation is triangle shaped, because the linear sweeps on either side of a sharp vertex imply an almost uniform distribution. The model should fail in that case, yet we know that the perceptual asymmetry is, if anything, *greater* for sharp than for flattened vertices (Demany and Clément, 1997). How to ac-

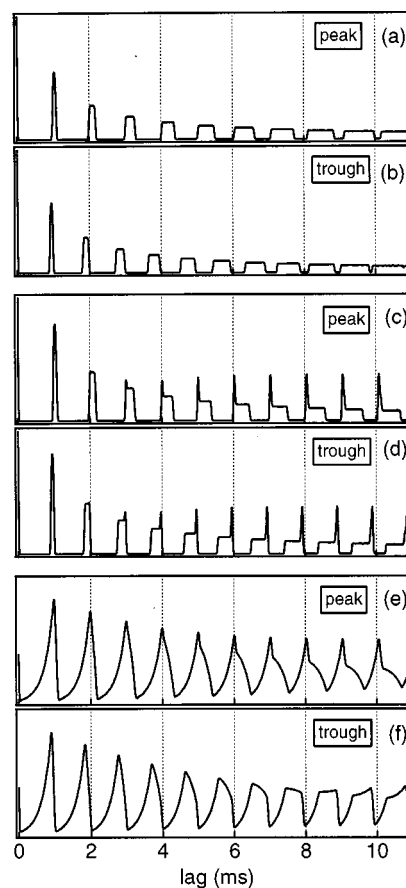


FIG. 11. (a) The ACF for a triangle-shaped peak, calculated over a 40-ms window placed before the vertex. (b) Same for a trough. (c) The ACF for a triangle-shaped peak, calculated over a 40-ms window placed after the vertex. (d) Same for a trough. (e) The AACF for a triangle-shaped window placed after the vertex. (f) Same for a trough.

count for the asymmetry, and also for the “hyperacute” frequency discrimination that Demany and Clément observed for peaks of triangular modulation?

If the integration window is placed *before* the vertex, indeed no asymmetry is observed. Figure 11(a) and (b) show the ACFs for a peak and a trough, respectively. Modes are blunt and low, all the more so as their order is high, and it does not seem that they would support accurate discrimination. If the window is placed before the vertex, one can account for neither perceptual asymmetry nor hyperacute discrimination.

If instead the integration window is placed *after* the vertex, each mode consists of two parts: a rectangular pedestal and a sharp “spire,” as shown in Fig. 11(c) for a peak of modulation. The spire represents intervals that span the peak, while the pedestal represents intervals that do not. The spire can be attributed to the effect noted in Sec. III B: if the pulse that ends an interval is displaced along one slope of the vertex, the pulse that begins it moves along the other. Any FM-induced change of period at one end is compensated by the opposite change at the other. The compensation is yet more effective than for cosine modulation, because the slopes on either side are more nearly constant and equal. For a valley of modulation the modes are similarly skewed, but in the opposite direction [Fig. 11(d)]. The spires are equally sharp

for peaks and valleys, so one cannot expect an asymmetry if discrimination is based on the ACF.

For the AACF, on the other hand, modes are sharper for a peak than for a trough [Fig. 11(e) and (f)]. This difference explains the asymmetry that Demany and Clément (1997) observed for *triangular* modulation, according to the present model. The explanation is the same as for cosine modulation, although the origin of the nonuniform interval distribution is different. *En passant*, the model accounts for the “hyperacute” discrimination observed for triangular peaks, a consequence of the mechanism that produces sharp spires in the higher-order modes. A similar hyperacuity would hold for troughs if it were not for the deleterious effect of the mechanism that produces the asymmetry.

D. Ramps

1. Cosine-shaped ramps

As pointed out earlier, the reasoning of Sec. III B for windows placed before the vertex applies directly to *ends* of ramps shaped like half-cosine periods. The model predicts an asymmetry in favor of rising ramps, as was observed by Demany and Clément (1997). For *beginnings* of ramps, one can apply the results of Sec. III B for windows placed after the vertex, with the important difference that samples to the left of the window now fall outside the signal (rather than on the opposite slope of a vertex). As a result, high-order modes are less well defined than in the case of full cosines, and discrimination should be less accurate, as Demany and Clément observed. One nevertheless still expects a peak/trough asymmetry which was indeed found (not shown).

2. Linear-shaped ramps

Contrary to all the conditions considered so far, the model predicts no asymmetry in the case of *linear* ramps. The frequency distribution is uniform, and the “compensation” mechanism that produced sharp spires is ineffective as it depends on the existence of ramps on both side of a vertex. The perceptual asymmetry should thus not be observed for linear ramps. Hyperacuity of pitch discrimination also should not be observed. Discrimination might thus actually be *worse* for the end of a ramp than for a peak flanked by linear ramps on both sides, despite the absence of the potentially interfering portion following the peak. Unfortunately linear ramps were not included in the stimuli investigated by Demany and Clément (1997), so we cannot test this prediction against their data.

IV. DISCUSSION

This paper offered an account for the perceptual asymmetry found by Demany and colleagues. It requires several strong assumptions that are reviewed here.

A. Importance and plausibility of the assumptions

Assumption 1. Assumption 1 (narrowing of the pulses of the spike probability function s) allows the shape of pulses of the convolution $s' = s \circ a$ to be determined mainly by the shape of the asymmetric kernel function $a(t)$. Such narrowing is not critical to the principle of the model, but if the

pulses of s were wider, their shape would dominate the convolution and the asymmetry of the pulses of s' would be reduced, and it is uncertain whether the model could explain peak/trough ratios as large as those observed. Certainly, the model could accommodate *several* sharp, well-spaced pulses per period instead of one.

Sharp pitch-following responses have been found in the cochlear nucleus. Temporal sharpening might be the result of the convergence of fibers on coincidence counting neurons (Kim *et al.*, 1986, 1988; Palmer and Winter, 1996), inhibition (Møller, 1983), refractory effects, etc. Standard deviations of first spike latency as small as 100 μ s have been reported by Rhode and Greenberg (1992), and there is evidence that time constants of spike integration can indeed be short, and possibly also tunable by varying the resting potential of neuron membranes (Banks and Smith, 1992). It has been proposed that such sharp responses might be a good substrate for pitch, but the proposition has also been questioned, because sharp pulses tend to be quite phase sensitive, and disappear for stimuli that lack a deeply modulated temporal envelope (Palmer and Winter, 1992). Other arguments against sharpening were put forward by Shofner (1999) and de Cheveigné (1999).

Sharpening is not a common assumption in pitch models based on autocorrelation (Licklider, 1956; Meddis and Hewitt, 1991a, b) or cancellation (de Cheveigné, 1998), but those models could certainly accommodate it. The strobed temporal integration (STI) of Patterson *et al.* (1992) is equivalent to a cross correlation between the discharge probability function and a strobe function with a single pulse per period. That strobe function is similar to the sharpened probability function assumed here, although the roles they play in the respective models are quite different.

Assumption 2. Assumption 2 (asymmetric autocorrelation) introduces the “peaked asymmetric function” that, convolved with the asymmetric interval distribution, produces a difference in accuracy between the frequency representations of peaks and troughs. This assumption is essential. It should be stressed that the opposite Kernel shape would produce the *opposite* perceptual asymmetry. It is therefore essential to justify the orientation of the asymmetry of the convolution kernel function involved in the AACF.

One can imagine at least two physiological accounts for smoothing of the delayed probability function (but not the undelayed function) by convolution with a short, rightward-skewed kernel function. (a) A first hypothetical account is that a jitter arises in the neural delay line, for example by reduction of conduction velocity following a spike. Spikes that follow closely after another spike would be delayed, adding a “tail” to the distribution of spike times on the side of greater delays. (b) A second hypothetical account is that the “coincidence counting neuron” of Licklider’s neural network is actually a *gating neuron*. The transmission of undelayed spikes is controlled by the density of delayed spikes, and the measure of this density naturally involves temporal smoothing. Each delayed spike contributes to the state of the neuron by an amount that decays exponentially, hence the exponential shape of the smoothing kernel function. Supposing that the gate controls the probability of

transmission of undelayed spikes without affecting their latency, the *undelayed* probability function remains unsmoothed. Spike probability at the output of the gate is thus the product of the undelayed input probability by the delayed probability convolved by an exponentially shaped smoothing kernel function. The spike count, across an array of such neurons indexed by delay, should resemble the AACF.

The second interpretation requires an assumption (gating neuron) that may seem unnecessary in the context of autocorrelation. However, it is congruent with the notion of “neural cancellation filter” that was used for pitch estimation in a model very similar to autocorrelation (de Cheveigné, 1998). The nuance between “gating neurons” and “coincidence counters” is significant if one imagines that the neural filters are one stage within a time-domain neural signal-processing network (Delgutte, 1984). A gating neuron can feed further processing, whereas a simple coincidence counter is a “dead end” as far as further time-domain processing is concerned. Note that if the first interpretation were true (jitter in the delay line), the convolution kernel function might be temporally wider for longer delays, rather than fixed as was assumed here.

Incidentally, the asymmetry of the AACF makes it sensitive to time reversal, and this might constitute yet another explanation of the ramp/damp asymmetries observed by Patterson (1994a, b; de Cheveigné, 1998).

Assumption 3. Assumption 3 (use of higher-order modes as pitch cues) is not essential for producing a peak/trough asymmetry for cosine modulation, although it makes things easier by “magnifying” the skew. It is required for producing an asymmetry for triangular modulation (the skew arises mainly for modes of order >1). It is also useful for explaining the hyperacuity of frequency discrimination at a triangular peak.

As pointed out earlier, higher-order mode cues are a useful hypothesis if one is to explain very accurate pitch discrimination, or the inverse dependency of thresholds on duration. They are indispensable in any model based on interval histograms derived from *single fibers*, as the first mode of both ISI and AC histograms disappears between 500 Hz and 1 kHz due to refractory effects. However, this constraint does not apply to interval statistics within *groups* of fibers, which are anyway a more reasonable assumption as they make much better use of the available temporal information (de Cheveigné, 1993, 1998).

Yost (1996) found that iterated-rippled-noise stimuli that differed by (waveform) autocorrelation modes other than the first had similar pitch strengths and were not readily discriminable. It is not clear whether that result can be generalized to imply that higher-order modes of autocorrelation histograms cannot be used in *any* task, in particular frequency discrimination with the stimuli of Demany and colleagues.

An obstacle to the use of higher-order modes is that they require long delay lines. The lack of physiological support for long delay lines is a weakness of time-domain models such as autocorrelation, and the higher-mode assumption rather makes it worse. High-order modes are also problematic in that they are multiple, and therefore ambiguous, and the auditory system would need a mechanism to keep track

of the order of the particular mode being used for a pitch match. The “narrowed autocorrelation function” (Brown and Puckett, 1989) has been proposed as a synthetic representation of all modes up to a certain order. It combines the accuracy of the higher-order modes with the lack of ambiguity of the lower-order modes (de Cheveigné, 1989; Slaney, 1990). Alternatively one can imagine an algorithmic process by which modes are registered in succession, starting at the lowest-order mode.

Other assumptions. Integration windows were chosen square for clarity, and it is not expected that other shapes would lead to significantly different results. Exponential-shaped windows are a natural choice, although estimates of temporal integration windows in other contexts (Plack and Moore, 1990) have lead to more rounded shapes.

Integration windows were chosen relatively large (40 ms) to produce legible effects in the plots, but asymmetries also occur for shorter windows. As argued above, windows should not be smaller than the stimulus period. The value of 2.5 ms suggested by Licklider (1956) and retained by Meddis and Hewitt (1991a, b) would be adequate at 1 kHz, but much too short at lower frequencies. It certainly should not be taken as a norm. One can argue that the auditory system has some latitude in choosing window sizes, and can within limits choose the best size for the task at hand (Hartmann and Klein, 1980).

The behavior of the model is evidently dependent on the range of mode *orders* that are allowed to enter the pitch estimate, and the rule by which evidence from modes of different orders is integrated. The range illustrated here was chosen arbitrarily. The asymmetry is also critically dependent on the use of the *position of the maximum* of a mode as the statistic. The matched filter of Srulovicz and Goldstein (1983) would probably not show the same asymmetry. A statistic based on the position of maximum *slope* might show the opposite asymmetry, etc.

B. Significance of the model

The perceptual asymmetry found by Demany and colleagues appears to be robust and general over a wide range of stimuli. This paper offered an account on the basis of emergent properties of a sensory mechanism. One cannot exclude that the asymmetry is due instead to a cognitive process that somehow “favors peaks,” or a mechanism such as that proposed by Clément (1996) at some intermediate stage. At this point, it is probably unwise to cling dogmatically to one hypothesis over the others. The contribution of the present model is twofold: to demonstrate that a sensory-based explanation is possible, but also a *contrario* to reveal its cost in terms of assumptions.

This model should not be seen as a general-purpose model of pitch perception, or as an “improvement” over previous models. It is rather a description of the assumptions that must be added to a standard model, autocorrelation, possibly reflecting detailed properties of that model’s implementation, in order to account for this particular phenomenon. It is not worth burdening the general-purpose model with assumptions needed only to explain a restricted class of phe-

nomena. It is, however, worth keeping track of any evidence, behavioral or physiological, in favor or against them.

The number and strength of the assumptions “raise the stakes:” either the model is accepted and the assumptions constitute a set of strong constraints on auditory processes, or else it is rejected precisely for that reason, as it requires the conjunction of hypotheses that are too many and too strong. This choice depends on the plausibility of the hypotheses, which has been discussed. It also depends on the existence and plausibility of alternative explanations for both asymmetry and hyperacuity.

En passant, the model gave an account of hyperacute frequency discrimination at peaks of triangular modulation. This depends only on assumption 3, and not on the other assumptions, and so this model of hyperacute discrimination might be retained even if the complete model of pitch asymmetry is not.

Alternative explanations of the asymmetry might be built following the same principles but differing in particulars. For example, the “lemma” of Sec. II B might be applied to another representation of the frequency distribution, for example, spectral. Convolution with a kernel function of appropriate shape (skewed with a tail towards higher frequencies) might produce the appropriate form of asymmetry. Interestingly, such a shape is found in pure-tone excitation patterns at high amplitudes: convolution with the frequency distribution should yield a slightly sharper excitation pattern at peaks than at troughs. However, it was seen in the Introduction that there are several reasons to doubt an explanation based on the asymmetry of excitation patterns, and to prefer an account based on temporal cues, of which this paper is an attempt. Whether the model is a success or not, the phenomenon that was revealed by Demany and colleagues is sufficiently ubiquitous to justify serious efforts to elucidate it.

V. CONCLUSION

A model was proposed to explain the perceptual asymmetry between peaks and troughs of frequency modulation. It is based on a particular implementation of the autocorrelation model of pitch perception, in which the delayed signal is “smeared” before multiplication by the undelayed signal. The interaction between the asymmetric function that describes the smearing, and the asymmetric distribution of interspike intervals, produces cues that are more accurate for peaks than for troughs. To summarize, we have the following.

- (1) The perceptual asymmetry between peaks and troughs of frequency modulation can be explained within the context of a time-domain model of pitch perception based on a modified form of autocorrelation.
- (2) The delayed term that enters the calculation of the autocorrelation function is assumed to be convolved with a short kernel function before multiplication with the undelayed term. This kernel function is asymmetrical and skewed towards positive values of time.
- (3) The difference in pitch estimation accuracy between peaks and troughs arises from an interaction between the asymmetric convolution kernel function and the nonuni-

form distribution of interpulse intervals near the vertex of modulation. This interaction produces cues that are sharp in the case of peaks, and blunt in the case of troughs.

- (4) The nonuniform interval distribution results from two factors. The first is the rounded shape of the instantaneous frequency distribution near the peak or trough (for cosine-shaped modulation). The second results from the existence of two populations of interpulse intervals, one comprising intervals that span the vertex, the other intervals between pulses on the same slope of the vertex. For linear modulation (triangular shaped) the second factor contributes alone to the nonuniform distribution of intervals.
- (5) The relative invariance of interpulse intervals that span the vertex accounts also for “hyperacute” discrimination of triangular-shaped modulation peaks. Discrimination of troughs benefits less because of the smoothing mechanism that causes the perceptual asymmetry.
- (6) The model accounts for perceptual asymmetry between peaks and troughs of modulation, whether cosine shaped or triangular. It also accounts for perceptual asymmetry of half-cosine-shaped ramps. However, for *linear* ramps, neither perceptual asymmetry nor hyperacute discrimination are to be expected. No experimental evidence is yet available to test this prediction.
- (7) The model requires several strong assumptions in addition to the asymmetric form of the autocorrelation. Discharge probability functions originating in the periphery must be “sharpened” to narrow pulses before pitch estimation. Discrimination must rely on the high-order modes of the (asymmetric) autocorrelation histogram, rather than just the first mode.

ACKNOWLEDGMENTS

Part of this work was carried out at ATR Human Information Processing Research Laboratories, within a research agreement with the Centre 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 Laurent Demany for many discussions on the subject of this perceptual asymmetry, and to him, Malcolm Slaney, and one anonymous reviewer for comments on earlier versions of the manuscript.

- Banks, M. I., and Smith, P. H. (1992). “Intracellular recordings from neurobiotin-labeled cells in brain slices of the rat medial nucleus of the trapezoid body,” *J. Neurosci.* **12**, 2819–2837.
- Brown, J. C., and Puckette, M. S. (1989). “Calculation of a ‘narrowed’ autocorrelation function,” *J. Acoust. Soc. Am.* **85**, 1595–1601.
- Clément, S. (1996). “Modélisation de la perception auditive des modulations de fréquence,” Master’s thesis in Cognitive Science, Université Bordeaux 2, France.
- de Cheveigné, A. (1989). “Pitch and the narrowed autocoincidence histogram,” *Proc. ICMPC*, Kyoto, pp. 67–70.
- de Cheveigné, A. (1993). “Separation of concurrent harmonic sounds: Fundamental frequency estimation and a time-domain cancellation model of auditory processing,” *J. Acoust. Soc. Am.* **93**, 3271–3290.
- de Cheveigné, A. (1998). “Cancellation model of pitch perception,” *J. Acoust. Soc. Am.* **103**, 1261–1271.
- de Cheveigné, A. (1999). “Pitch shifts of mistuned partials: a time-domain model,” *J. Acoust. Soc. Am.* **106**, 887–897.

- Delgutte, B. (1984). "Speech coding in the auditory nerve: II. Processing schemes for vowel-like sounds," J. Acoust. Soc. Am. **75**, 879–886.
- Demany, L. (1997). Personal communication.
- Demany, L., and McAnally, K. I. (1994). "The perception of frequency peaks and troughs in wide frequency modulations," J. Acoust. Soc. Am. **96**, 706–715.
- Demany, L., and Clément, S. (1995a). "The perception of frequency peaks and troughs in wide frequency modulations. II. Effects of frequency register, stimulus uncertainty, and intensity," J. Acoust. Soc. Am. **97**, 2454–2459.
- Demany, L., and Clément, S. (1995b). "The perception of frequency peaks and troughs in wide frequency modulations. III. Complex carriers," J. Acoust. Soc. Am. **98**, 2515–2523.
- Demany, L., and Clément, S. (1997). "The perception of frequency peaks and troughs in wide frequency modulations. IV. Effects of modulation waveform," J. Acoust. Soc. Am. **102**, 2935–2944.
- Demany, L., and Clément, S. (1998). "The perceptual asymmetry of frequency modulation," in *Psychophysical and Physiological Advances in Hearing*, edited by A. R. Palmer, A. Rees, A. Q. Summerfield, and R. Meddis (Whurr, London).
- Dooley, G. L., and Moore, B. C. J. (1988). "Duration discrimination of steady and gliding tones: a new method for estimating sensitivity to rate of change," J. Acoust. Soc. Am. **84**, 1332–1337.
- Hartmann, W. H. (1997). *Signals, Sound and Sensation* (AIP, Woodbury, NY).
- Hartmann, W. M., and Klein, M. A. (1980). "Theory of frequency modulation detection for low modulation frequencies," J. Acoust. Soc. Am. **67**, 935–946.
- Kim, D. O., Rhode, W. S., and Greenberg, S. R. (1986). "Responses of cochlear nucleus neurons to speech signals: neural encoding of pitch, intensity and other parameters," in *Auditory Frequency Selectivity*, edited by B. C. J. Moore and R. D. Patterson (Plenum, New York), pp. 281–282.
- Kim, D. O., and Leonard, G. (1988). "Pitch-period following response of cat cochlear nucleus neurons to speech sounds," in *Basic Issues in Hearing*, edited by H. Duifhuis, J. W. Horst, and H. P. Wit (Academic, London), pp. 252–260.
- Licklider, J. C. R. (1956). "Auditory frequency analysis," in *Information Theory*, edited by C. Cherry (Butterworth, London), pp. 253–268.
- 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.
- Møller, A. R. (1983). *Auditory Physiology* (Academic, New York).
- Moore, B. C. J. (1973). "Frequency difference limens for short-duration tones," J. Acoust. Soc. Am. **54**, 610–619.
- Palmer, A. R., and Winter, I. M. (1992). "Cochlear nerve and cochlear nucleus responses to the fundamental frequency of voiced speech sounds and harmonic complex tones," in *Auditory Physiology and Perception*, edited by Y. Cazals, L. Demany, and K. Horner (Pergamon, Oxford), pp. 231–239.
- Palmer, A. R., and Winter, I. M. (1996). "The temporal window of two-tone facilitation in onset units of the ventral cochlear nucleus," *Audiol. Neuro-Otol.* **1**, 12–30.
- Patterson, R. D. (1994a). "The sound of a sinusoid: time-domain models," J. Acoust. Soc. Am. **96**, 1419–1428.
- Patterson, R. D. (1994b). "The sound of a sinusoid: Spectral models," J. Acoust. Soc. Am. **96**, 1409–1418.
- Patterson, R. D., Robinson, K., Holdsworth, J., McKeown, D., Zhang, C., and Allerhand, M. (1992). "Complex sounds and auditory images," in *Auditory Physiology and Perception*, edited by Y. Cazals, K. Horner, and L. Demany (Pergamon, Oxford), pp. 429–446.
- 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.
- Rhode, W. S., and Greenberg, S. (1992). "Physiology of the cochlear nuclei," in *The Mammalian Auditory Pathway: Physiology*, edited by A. N. Popper and R. R. Fay (Springer Verlag, New York), pp. 94–152.
- Ruggero, M. A. (1992). "Physiology of the auditory nerve," in *The Mammalian Auditory Pathway: Neurophysiology*, edited by A. N. Popper and R. R. Fay (Springer Verlag, New York), pp. 34–93.
- Shepard, R. N. (1964). "Circularity in judgments of relative pitch," J. Acoust. Soc. Am. **36**, 2346–2353.
- Shofner, W. P. (1999). "Responses of cochlear nucleus units in the chinchilla to iterated rippled noises: analysis of neural autocorrelograms," J. Neurophysiol. **81**, 2662–2674.
- Slaney, M. (1990). "A perceptual pitch detector," Proc. ICASSP-90, pp. 357–360.
- Srulovicz, P., and Goldstein, J. L. (1983). "A central spectrum model: a synthesis of auditory-nerve timing and place cues in monaural communication of frequency spectrum," J. Acoust. Soc. Am. **73**, 1266–1276.
- Yost, W. A. (1996). "Pitch strength of iterated rippled noise," J. Acoust. Soc. Am. **100**, 3329–3335.

Frequency discrimination in budgerigars (*Melopsittacus undulatus*): Effects of tone duration and tonal context

Micheal L. Dent, Robert J. Dooling,^{a)} and Alisa S. Pierce

Department of Psychology, University of Maryland, College Park, Maryland 20742

(Received 2 July 1999; accepted for publication 21 January 2000)

Studies of frequency resolving power in budgerigars (*Melopsittacus undulatus*) have shown that this species has excellent discrimination abilities for both simple and complex sounds falling in the region of 2 to 4 kHz—the frequency range of their contact call. In four experiments, frequency discrimination by budgerigars of short tones similar to elements found in the contact call was examined. Frequency difference limens (FDLs) for simple pure tones at 2.86 kHz were constant for tone durations above 20 ms but higher for shorter tones. Budgerigars generally showed larger FDLs for shorter duration 1-, 2-, and 4-kHz pure tones. FDLs in budgerigars for 20-ms tones embedded in a sequence of six other tones were similar to FDLs measured for tones of the same frequency presented in isolation. Moreover, there was no effect of introducing trial-by-trial variation in the location of the frequency change in the seven-tone complexes for budgerigars, a condition for which humans showed a large decrement in performance. Taken together, these results suggest budgerigars possess enhanced spectral resolving power for short duration pure tones when they are embedded in contact call-like tonal patterns. © 2000 Acoustical Society of America. [S0001-4966(00)00505-1]

PACS numbers: 43.66.Gf, 43.80.Lb [WA]

INTRODUCTION

Budgerigars (*Melopsittacus undulatus*) are small Australian parrots that show remarkable vocal plasticity and vocal learning throughout life. The dominant vocalization in their repertoire is a highly stereotyped, frequency modulated contact call about 150–200 ms in duration with most of the spectral energy concentrated in the region of 2–4 kHz (Dooling, 1986; Farabaugh and Dooling, 1996). Several experiments have described the important features that budgerigars use when listening to and discriminating among contact calls. These features are primarily spectral and include peak frequency, rate of frequency modulation, and concentration of spectral energy (Brown *et al.*, 1988; Dooling *et al.*, 1987). Budgerigars maintain nearly perfect performance on an identification task involving contact calls even when these calls are temporally or spectrally distorted (Park and Dooling, 1985, 1986), just as humans can recognize speech sounds under adverse conditions (French and Steinberg, 1947; Miller, 1981). These results on the perception and recognition of contact calls are supported by a variety of other experiments using simple sounds, which show that these birds have excellent frequency resolving power in the spectral region of 2–4 kHz (Dooling and Saunders, 1975; Okanoya and Dooling, 1987).

The experiments reported here further explored the ability of budgerigars to discriminate spectral changes in simple pure tones and contact call-like stimuli and compared these results with those from humans tested with similar procedures. In humans, a number of stimulus factors influence frequency discrimination abilities, and many of these stimulus characteristics are present in the budgerigar contact call. These include differences in durations, frequencies, and se-

quences of tones within patterns. By studying these factors in a particular frequency discrimination paradigm, we hoped to determine whether budgerigars have a particular species-specific advantage in perceiving small frequency changes in stimuli that closely resemble their contact calls. In a first experiment, frequency difference limens (FDLs) for pure tones at 2.86 kHz were measured at durations ranging from 5 to 160 ms. In a second experiment, FDLs were measured for very short tones at three additional frequencies between 1 and 4 kHz. Although FDLs for simple pure tones have been reported for budgerigars (Dooling and Saunders, 1975) and other birds (Gray and Rubel, 1987; Kuhn *et al.*, 1980; Langemann and Klump, 1992; Quine and Konishi, 1974; Sinnott *et al.*, 1980), the relation of FDL to duration has yet to be systematically examined in birds. In humans, there is an increase in FDL with decreasing tone duration for tones below 4–5 kHz (e.g., Moore, 1973). These results suggest that FDLs for low-frequency tones may be determined by temporal factors (for a review, see Moore, 1997). Since budgerigars excel in some aspects of temporal processing, they might show an enhanced ability to detect frequency changes, especially for tones of shorter duration (Amagai *et al.*, 1997; Dent *et al.*, 1999; Lohr and Dooling, 1998).

Finally, the third and fourth experiments examine auditory pattern perception by budgerigars. In an approach modeled after a series of human experiments using word-length tonal patterns (Watson *et al.*, 1975, 1976), we examined how well budgerigars could “hear out” a small frequency change in specific temporal locations of contact call-length tonal patterns. Watson *et al.* (1975) reported that human listeners’ discrimination of a small change in one tone embedded within a sequence of other tones was affected by the target tone’s frequency and temporal position within the pattern, and the listener’s knowledge of when the changes would occur. Specifically, frequency changes in targets low in fre-

^{a)}Electronic mail: dooling@psyc.umd.edu

quency relative to the surrounding components, and in targets occurring at the beginning of the tonal pattern, were harder to discriminate than other target tones. Equally as important to discrimination performance, however, was the degree of the listener's knowledge about the target and the tonal pattern. High levels of uncertainty as to when the change in the target pattern would occur resulted in poor target-tone frequency discrimination (Watson *et al.*, 1976). Espinoza-Varas and Watson (1989) discussed the importance of central processes in the auditory-perceptual representation of sounds. They suggested that, for complex patterns, peripheral receptors probably do not impose the greatest limitations on performance. Rather, central processes such as attention, learning, and memory capacity play large roles in these discrimination abilities. Such central capacities may shed some light on auditory processing that differs between humans and birds.

In the present experiments, we compared budgerigars and humans on the discrimination of call-length tonal patterns under several conditions. First, tonal patterns were constructed where each tone was 2.86 kHz (the frequency of best hearing for the budgerigar). Next, tonal patterns were constructed with component frequencies varying from 1–4 kHz assembled in a random order. Subjects were tested under conditions in which the temporal location of the frequency change was the same from trial to trial (low uncertainty) and conditions where the location of the frequency change was randomized on a trial-by-trial basis (high uncertainty). Comparing the performances for budgerigars and humans in the low uncertainty and high uncertainty conditions provides some indication of the degree to which attention-like factors contributed to the thresholds reported here.

I. GENERAL METHODS

A. Testing apparatus

The birds were tested in a wire cage (23×25×16 cm) mounted in a sound-isolation chamber (Industrial Acoustics Company, IAC-3). A response panel consisting of two microswitches with light-emitting diodes (LEDs) was mounted on the wall of the test cage just above the food hopper. The microswitch was tripped by the bird pecking the LED. The left microswitch and LED served as the observation key while the right microswitch and LED served as the report key. The behavior of the animals during test sessions was monitored by a video camera system (Sony HVM-322).

An IBM 486 computer controlled test sessions. Pure tones were generated digitally at 20 kHz and output through Tucker-Davis modules to a speaker (KEF Electronics, Holliston, MA, model 80C) mounted 36 cm above the perch in the testing cage. Tones were output at an intensity of 65 dB SPL at the location of the bird's head. Stimulus calibration was performed using a General Radio (model 1982) sound-level meter. Stimulus intensities were measured with the microphone (1/2 in attached to the sound-level meter via a 3-m extension cable) in front of the response keys in the approximate position occupied by the bird's head during testing. Stimulus intensities were calibrated several times during

these experiments to ensure that stimulus levels remained constant and the entire system was functioning appropriately.

B. Training and testing procedures

The birds were trained by a standard operant autoshaping program to peck at the left microswitch key (observation key) during a repeating background. After a random time interval of 2–7 s, a new stimulus was presented alternately with the background sound. If the bird pecked the right microswitch and LED (report key) within 2 s of this alternating pattern, the food hopper was activated for 2 s. A failure to peck the report key within 2 s of sound alternation was recorded as a miss and a new trial sequence was initiated. The dependent variable in these experiments was the percent correct responses on trials involving an alternating sound pattern. Thirty percent of all trials were “sham” trials in which the target sound was the same as the repeating background sound. A peck to the report key during a sham trial was recorded as a false alarm, and the lights in the test chamber were extinguished while the repeating background continued. The length of this time-out period was normally 5 s, but varied (up to 9 s) according to the bird's behavior. Longer time-out periods were instituted if the birds began developing higher false alarm rates. Sessions with a total false alarm rate of 16% or higher were discarded. About 10% of all sessions across birds were discarded for this reason.

Stimuli were presented according to the Method of Constant Stimuli (Dooling and Okanoya, 1995) in 5-Hz (experiments 1 and 2) or 10-Hz steps (experiments 3 and 4), with frequency values selected to bracket the presumed threshold. At the conclusion of testing, psychometric functions were constructed and thresholds were defined in several ways, including the frequency difference that the bird detected 50% of the time (unadjusted threshold), the frequency difference detected 50% of the time adjusted by the false-alarm rate (adjusted threshold; Hienz *et al.*, 1977; Sinnott *et al.*, 1980), and the frequency difference resulting in a d' of 1.5 (Dooling and Okanoya, 1995; Green and Swets, 1966; Penner, 1995). There was little difference between adjusted, unadjusted, and d' threshold values, so adjusted thresholds were used for all data analysis.

II. EXPERIMENT 1

In the first experiment, FDLs were measured for pure tones of different durations. The frequency of all tones was 2.86 kHz, which is the frequency of best hearing in budgerigars.

A. Methods

1. Subjects

Three adult budgerigars (two males and one female) were used as subjects. The birds were kept on a normal day/night cycle correlated with the season at approximately 90% of their free-feeding weights. Three humans (one male, two females) were also tested in this experiment, to allow comparison with earlier studies using humans. None of the sub-

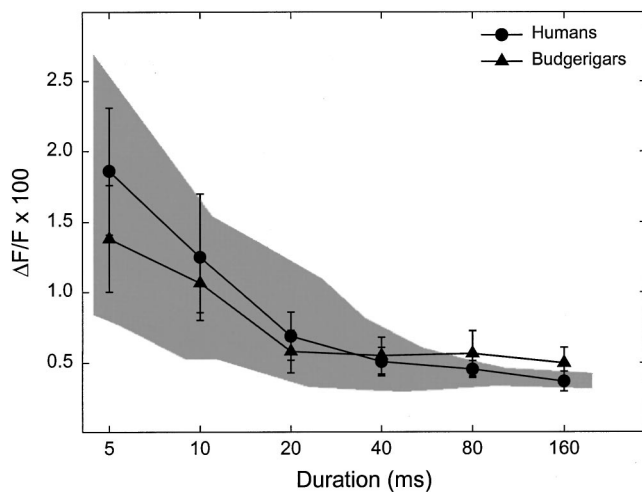


FIG. 1. Mean FDLs for three budgerigars (closed triangles) and three humans (closed circles) for 2.86-kHz pure tones of different durations. Error bars are standard errors. Each data point represents 200 trials and is plotted as Weber fraction ($\Delta F/F \times 100$) by duration (ms). The gray area represents the range of data on humans from previous experiments using slightly different stimuli and procedures (see text for references).

jects reported a history of hearing disorders, and all were researchers at the University of Maryland working in the laboratory at the time of the experiment.

2. Stimuli and procedure

All stimuli in this experiment were 2.86 kHz pure tones with 2-ms rise/fall times (shaped with a cosine function). The tones were 5, 10, 20, 40, 80, and 160 ms in duration (including the rise/fall times). FDLs were measured for a minimum of 300 trials at all durations in a random order, and a different random order was used for each bird. In a first set of tests, the tones were presented at a peak sound pressure level of 65 dB SPL as measured by the General Radio sound-level meter. In a second set of tests, the intensities of the pure tones of different durations were adjusted to be of equal energy according to the budgerigar's temporal integration function (Dooling and Searcy, 1985b).

Humans were tested with similar procedures as the birds, except that they listened to the stimuli over earphones (AKG type K-240 DF) and pressed keys on a hand-held control rather than hitting the LEDs attached to microswitches on the birds' response panel. Human subjects were tested for 100 trials at all tone durations. Tones were presented at a comfortable listening level (about 65 dB SPL) as measured by the General Radio sound-level meter.

B. Results and discussion

The data from the constant SPL and the energy-adjusted SPL experiments were not significantly different, $t(42) = 2.02$, $p > .05$, by a two-tailed paired t -test, so the results for both tests were combined for all subsequent comparisons. Figure 1 shows the mean FDLs as a function of tone duration for three budgerigars. These results are compared with the results from our humans tested using similar procedures used to test the birds and with the range of published data reported on humans tested by other investigators in other laboratories

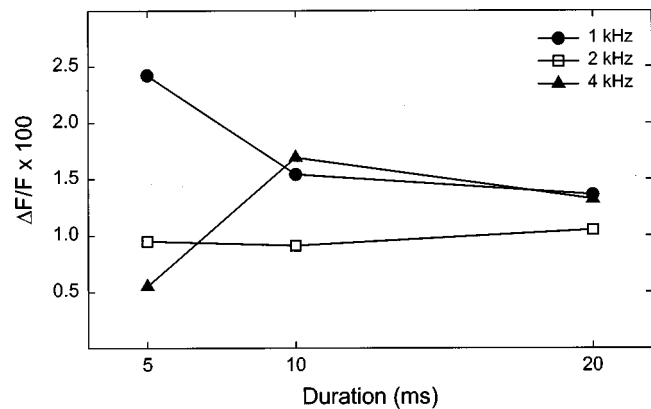


FIG. 2. Mean FDLs from budgerigars for pure tones of different durations and frequencies. Each data point represents a mean of two budgerigars and is plotted as Weber fraction ($\Delta F/F \times 100$) by duration (ms).

using different procedures (Freyman and Nelson, 1987; Hall and Wood, 1984; Hartmann *et al.*, 1985; Henning, 1970; Moore, 1973; Turnbull, 1944). The results for the humans in our experiments at all durations fall well within the range of results in other experiments on humans. For both humans and budgerigars, FDLs are relatively constant for tone durations longer than 20 ms but worsen considerably as tone duration decreases below 20 ms. A two-way repeated measures analysis of variance shows no differences between the FDLs of budgerigars and humans [$F(1,5) = 0.37$, $p > 0.05$]. For both species, there was a significant difference among durations [$F(5,25) = 27.73$, $p < 0.05$] with higher FDLs occurring at shorter tone durations. The interaction between the two factors was not significant [$F(5,25) = 1.65$, $p > 0.05$], and there were no differences between the budgerigars and the humans at any of the specific tested durations.

III. EXPERIMENT 2

The previous experiment showed that at their frequency of best hearing, budgerigars exhibit FDLs that worsen as a function of decreasing tone duration below 20 ms. Experiment 2 tested whether this result was typical of other frequencies within the budgerigar's range of hearing.

A. Methods

The specific methods and procedures were similar to those used in experiment 1. Two of the three budgerigars from experiment 1 (one male and one female) were used as subjects. The stimuli in this experiment were 1-, 2-, and 4-kHz pure tones. FDLs were measured for a minimum of 300 trials at each duration. The tones were 5, 10, and 20 ms in duration (including the 2-ms rise/fall times). FDLs were measured in a random order and a different random order was used for each bird. The tones were presented at a peak sound pressure level of 65 dB SPL as measured by the General Radio sound-level meter.

B. Results and discussion

Figure 2 shows mean FDLs for two budgerigars at three durations and three frequencies. In general, FDLs at 1 and 4 kHz are worse than FDLs at 2 kHz (and 2.86 kHz, see Fig. 1), consistent with previous results from this species (Dool-

ing and Saunders, 1975). As tone duration decreases from 20 to 10 ms, FDLs increase at all three frequencies, similar to the results of experiment 1 for 2.86 kHz. For the 5-ms tone burst duration, however, the situation is somewhat more complicated. At 1 kHz, the mean FDL for the 5-ms tone is much worse than it is for the tone duration of 10 ms. For the 2-kHz tone burst, however, the mean FDL for the 5-ms tone is similar to the mean FDL at 10 ms. Surprisingly, for the 4-kHz tone burst, the mean FDL at 5 ms is much lower than the mean FDL for the 10- or 20-ms tone burst durations. In other words, FDLs improve with increasing frequency for short duration tones.

In humans, the effect of frequency on FDLs has been the subject of considerable investigation over the years (see, for example, Moore, 1973). In general, the discrimination of pitch may involve time-based mechanisms for frequencies below 5 kHz and place-based mechanisms at higher frequencies, with some perturbations evident in the transition from time-based to frequency-based mechanisms (Moore, 1973). In addition, there is the possibility that loudness cues influence FDLs measured at high frequencies where absolute sensitivity is decreasing at a rapid rate (Henning, 1966). In budgerigars, for instance, the budgerigar audiogram shows a loss of sensitivity of about 50 dB/octave at frequencies above 4 kHz. Finally, for very short duration tones (5 ms) where the rise/fall times (2 ms) are a significant fraction of the stimulus duration, there is always the possibility that small amounts of spectral splatter could influence thresholds differentially across frequencies. Since much less is known about the psychophysics of hearing in birds compared to humans, some or all of these factors may be relevant to the differences in FDLs across frequency at 5 ms for our birds.

In general, though, the results for budgerigars parallel those reported for humans at durations above about 10 ms. These results are important for understanding the design of the following experiments, where FDLs of short tones are measured in an acoustic context of other short tones—mimicking a natural contact call produced by this species.

IV. EXPERIMENT 3

The previous experiments showed that for durations above 10 ms, budgerigars and humans show roughly similar patterns of FDLs across durations and across frequencies. In experiment 3, we explored the effects of a surrounding tonal context on discrimination of frequency change in a (24-ms) tone burst. In this experiment, the FDL was measured for short pure tones embedded in a pattern whose total duration was 198 ms. This is about the duration of naturally produced tonal contact calls of budgerigars.

A. Methods

1. Subjects

Two adult budgerigars (both males) and two humans (one male, one female) were used as subjects. Neither of the human subjects reported a history of hearing disorders, and both were working in the laboratory at the time of the experiments.

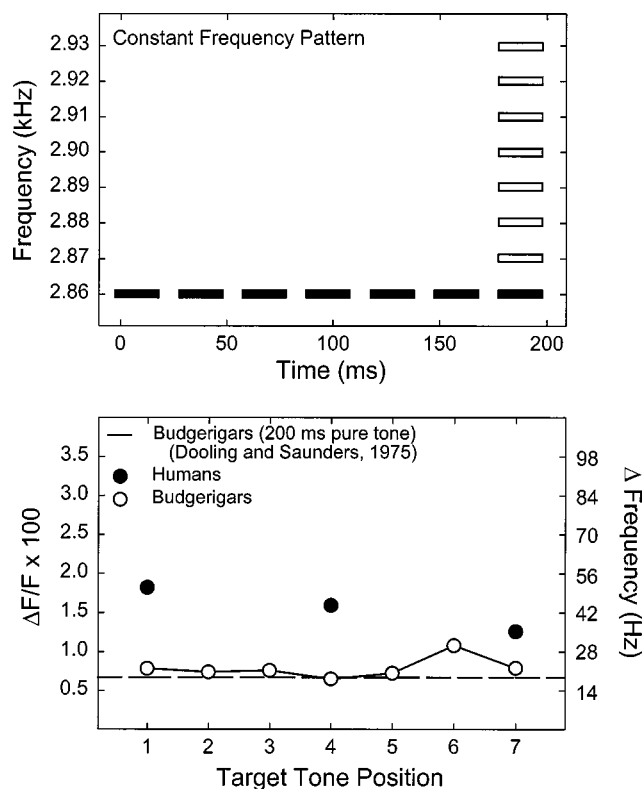


FIG. 3. The top panel is a schematic of a frequency-constant tonal pattern with examples of seventh position changes. The closed rectangles represent the constant frequency pattern. During a session measuring the FDL for the 7th position, the 7th tonal element is randomly replaced with one of the higher tonal elements (open rectangles) on any given trial. The bottom panel shows FDLs for a 2.86-kHz frequency tonal pattern. The mean of two birds (open circles) and two humans (closed circles) is represented by frequency in Hz (right axis) and Weber fractions (left axis). The dashed line represents budgerigar FDLs for a 200-ms pure tone of the same frequency in isolation (from Dooling and Saunders, 1975).

2. Stimuli and procedure

The stimuli in this experiment consisted of a sequence of seven 2.86-kHz pure tones, each 24-ms long (including 2-ms rise/fall times), with a 5-ms intertone silent interval. The birds' FDLs were measured for each tone in the seven-tone pattern. The patterns were presented at 65 dB SPL. Figure 3 (top) shows a schematic diagram of the fixed-frequency tonal pattern and an example of one tone selected as the target. In this example, seven target patterns are shown in which the frequency of the seventh component was changed in 10-Hz steps. FDLs were obtained for one component (position) in the pattern before another component was tested. This procedure was followed until all seven positions of the pattern were tested in a random order.

Humans were tested with similar procedures as the birds, except that they listened to the stimuli over earphones and pressed keys on a hand-held control rather than LEDs on the birds' response panel. The humans were tested on three positions of the fixed-frequency tonal pattern.

B. Results and discussion

Mean FDLs for the two budgerigars are shown for all seven temporal positions (Fig. 3, bottom). FDLs for these short tones embedded within a sequence of like-frequency

tones are as good as FDLs for simple tones of the same overall duration (i.e., 200 ms) (Dooling and Saunders, 1975). Weber fractions ($\Delta F/F \times 100$) for the budgerigar have a mean of 0.7 across all positions. In fact, Weber fractions for 20-ms pure tones of the same frequency presented in isolation (see experiment 1) are the same as those for the 20-ms tones embedded in the patterns used in this experiment. The results for the humans differ from those of the budgerigars, but the trends are similar to those from earlier experiments. Watson and his colleagues found that thresholds for discrimination of tones earlier in a (mixed-frequency) tonal pattern were higher than those occurring later in the pattern (Watson *et al.*, 1975). Here, humans tested on tonal patterns had Weber fractions ranging from 1.6 to 2.3, depending on the position of the change. Experiment 1 showed that FDLs for humans tested on tones of 2.86 kHz with a duration of 20 ms had Weber fractions ranging from about 0.3 to 1.2. FDLs for tones embedded in the tonal patterns were higher than those for simple pure tones for humans but not for budgerigars.

V. EXPERIMENT 4

The previous experiment showed that budgerigars were as good at discriminating frequency changes in pure tones embedded in complex tonal patterns as they were at discriminating frequency changes in tones presented in isolation. This experiment measured FDLs for pure tones embedded in tonal patterns again, but this time the tonal patterns were further modified to mimic some of the spectral characteristics of the contact calls of this species. That is, the components of the patterns were now varied in frequency, and the target tone to be discriminated was embedded within this varying-frequency pattern. Further, FDLs were measured for each component of tonal patterns where the position of the target tone varied from trial to trial to assess the role of experimental uncertainty and possible attentional factors in these experiments.

A. Methods

1. Subjects

The budgerigars and humans in this experiment were the same as those used in experiment 3.

2. Stimuli and procedure

The stimuli in this experiment consisted of a sequence of seven pure tones, each 24-ms long (including 2-ms rise/fall times), with a 5-ms intertone interval. The seven-tone patterns were constructed so that each tone was of a different frequency, ranging from 1 to 4 kHz, in 500-Hz steps, and arranged in a random order (see Fig. 4). The frequency spacing of the tonal components exceeded the critical bandwidth of the budgerigars at all frequencies (Dooling and Searcy, 1979; Saunders *et al.*, 1979). Each bird was tested on three different variable tone patterns (to account for positional and frequency effects of the tonal components), each with a different random temporal ordering of the seven tones. FDLs were measured for each tone in these patterns in random order, in both a low uncertainty and a high uncertainty con-

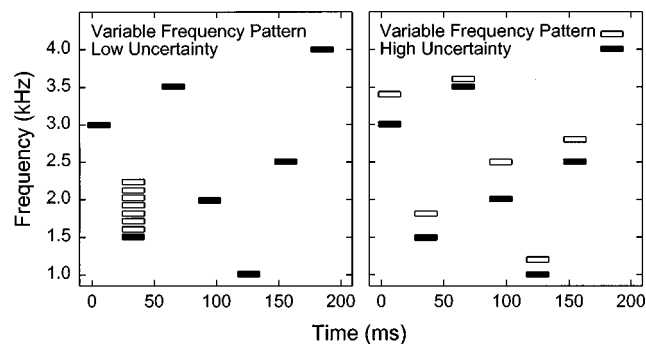


FIG. 4. Schematic of a variable frequency tonal pattern in the low uncertainty (left panel) and high uncertainty (right panel) conditions (closed rectangles). During a low uncertainty session measuring the 2nd position FDL, the second tone is replaced by a tone of higher frequency (one of the open rectangles). During the entire low uncertainty session, changes will occur only at that position. Examples of seven possible changes in a high uncertainty session are shown by the open rectangles in the right panel. Here, the change can occur in any of the positions (open rectangles, right panel) from trial to trial in the session.

dition. Thresholds for each frequency were obtained by averaging across all three variable tone patterns.

In the low uncertainty condition (Fig. 4, left), FDLs were measured for a single component within the pattern in each session. By repeatedly testing the bird with the stimulus change occurring in the same temporal and spectral location, the bird could know “where to listen” for a frequency change in the repeating tone pattern. In other words, since trial-to-trial changes occurred in the same location in the tonal pattern, the birds should be able to focus on a particular segment and a particular frequency. Humans listening in a similar low uncertainty stimulus condition are thought to use their knowledge of the target location to focus their auditory attention and reduce interference from surrounding tones (Watson *et al.*, 1976). On the other hand, in the high uncertainty condition (Fig. 4, right) the location of the frequency change in the pattern varied on a trial-by-trial basis. Watson *et al.* (1976) suggested if a listener cannot know from one trial to the next where in a multitone pattern a change is going to occur, they cannot focus attention on a specific location within the pattern, and instead must distribute attention across all components. A comparison of FDLs in low and high stimulus uncertainty conditions should reveal the role of “attention” in complex pattern perception in these birds.

Humans were tested with similar procedures as the birds, except that they listened to the stimuli over earphones and pressed keys on a hand-held control rather than LEDs on the birds’ response panel. The humans were tested on all positions of one of the variable frequency patterns under both the low and high uncertainty conditions.

B. Results and discussion

Discrimination of the variable-frequency tonal patterns was more difficult than discrimination of the fixed frequency patterns for both birds and humans, under both low and high uncertainty conditions. For birds, Weber fractions were somewhat larger below 2 kHz and relatively constant between 2.5 and 4 kHz (see Fig. 5), which is similar to the

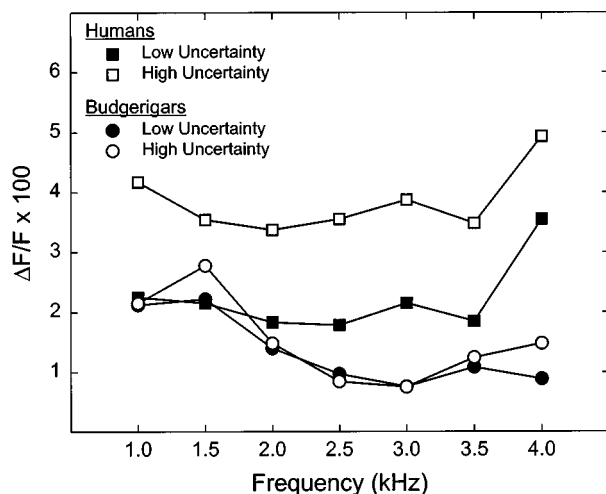


FIG. 5. Mean FDLs for the variable frequency tonal patterns are plotted as Weber fractions by frequency (kHz) for two birds and two humans. The budgerigars ran on three different variable frequency tonal patterns under both the low uncertainty (closed circles) and high uncertainty (open circles) conditions. The humans ran on one variable frequency tonal pattern under both the low uncertainty (closed squares) and high uncertainty (open squares) conditions.

results from Watson *et al.* (1975) in humans. The humans in this study had Weber thresholds that were generally larger than those of budgerigars with thresholds that were relatively constant at the lower frequencies but somewhat higher at 4 kHz (a result opposite of that reported by Watson *et al.*, 1975).

A two-way repeated measures analysis of variance compared results from humans and budgerigars at the low and high uncertainty conditions. Both the species factor [$F(1,52)=15.82$, $p<0.05$] and the level-of-uncertainty factor were significant [$F(1,52)=4.54$, $p<0.05$]. That is, budgerigars were significantly different from humans and the low uncertainty condition was significantly different from the high uncertainty condition. A subsequent two-tailed, paired sample *t*-test on the individual species at each condition showed that budgerigar FDLs between the low uncertainty and high uncertainty conditions were not significantly different [$t(13)=0.74$, $p>0.05$], while the human FDLs were significantly different between the two conditions [$t(13)=4.28$, $p<0.05$]. Thus not knowing where the change was going to occur from trial to trial affected a human's but not a budgerigar's ability to discriminate frequency changes in a single component of these complex tonal patterns.

VI. GENERAL DISCUSSION

Experiments 1 and 2 showed that the frequency discrimination abilities of budgerigars for simple pure tones were similar to those found in humans. The FDLs were relatively constant for long duration pure tones above 20 ms and increased as tone duration decreased below 20 ms. The FDLs reported here for both budgerigars and humans fell within the range of data from humans previously tested using different procedures (Freyman and Nelson, 1987; Hall and Wood, 1984; Hartmann *et al.*, 1985; Henning, 1970; Moore, 1973; Turnbull, 1944). The best frequency discrimination for

short tones was in the region of 2–3 kHz for budgerigars, and similarly in humans, the lowest FDLs were at 2 kHz (Moore, 1973).

Experiment 3 showed that as the complexity of the experimental situation increases, humans showed decrements in performance while the birds were much less affected. In this minimal uncertainty pattern perception task, the context of the components being measured did not change from trial to trial (all tones were the same duration and frequency) and neither did the temporal position of the target. In budgerigars tested under these conditions, FDLs for pure tones embedded in a sequence of like-frequency tones were the same as FDLs for simple short tones presented in isolation.

Humans, on the other hand, may have experienced some temporal masking in this situation because they were not able to discriminate changes in these tones embedded in the patterns as well as they could for the same tones presented in isolation. There is evidence that considerably less masking occurs for budgerigars in nonsimultaneous masking experiments where birds demonstrate a much greater frequency selectivity than humans (Dooling and Searcy, 1985a). Although these experiments involved detection, the same mechanisms may also be affecting discrimination abilities and this may explain why the budgerigars were not as affected as the humans were by the complexity of this frequency discrimination task.

Experiment 4 added another level of task complexity. This experiment tested birds and humans on tonal patterns consisting of tones of different frequencies. This increased the uncertainty levels but only slightly increased the thresholds compared to experiment 3. The similar thresholds in experiment 4 clearly show that the context of the tones has only a minor effect on frequency discrimination in budgerigars. When the surrounding context varies along the same dimension as the dependent variable, budgerigars have only slightly more difficulty ignoring the context. Moreover, the tasks in experiment 4 ranged from low uncertainty, where the FDL being measured was always in the same position from trial to trial, to high uncertainty, where the FDL being measured could be in any position from trial to trial. Surprisingly, the budgerigars, in contrast to humans, were completely unaffected by the manipulations of uncertainty level.

One interpretation of the present results is that budgerigars can listen in an analytic mode even under high uncertainty conditions where humans cannot. For instance, Espinoza-Varas and Watson (1986) suggest that one effect of increasing the level of uncertainty in humans is that they switch from an analytic mode of listening (focusing attention on specific components while ignoring others) to a synthetic mode of listening (a multiple-channel mode of listening where the listeners focus on how the components are related). Perhaps budgerigars do not switch listening modes in the same way as humans.

Another interpretation for the species differences in these experiments comes from results of several studies on humans, which suggest that different processes mediate frequency discrimination in low- versus high-frequency regions. Temporal coding may be more important at low frequencies, while tonotopic organization is more relevant at

higher frequencies (see Prosen *et al.*, 1989 for a review). Coding in birds is somewhat less certain than in mammals, but recent experiments suggest that some aspects of temporal processing in birds may be better than that found in humans (Dooling *et al.*, 1999). For instance, budgerigars show an enhanced ability to discriminate between two harmonic complexes with identical envelope shapes that differ only in temporal fine structure (Dent *et al.*, 1999). As another example, in discriminating silent temporal gaps in sinusoidal markers, budgerigar performance was relatively unaffected by a frequency change in the tonal marker following the gap, while humans show large decrements in performance as the separation between the markers exceeds the critical bandwidth (Amagai *et al.*, 1997). Finally, both budgerigars and zebra finches are much better than humans at detecting the mistuning of single components in a harmonic complex (Lohr and Dooling, 1998). If the transition from temporal to tonotopic coding is different in budgerigars, and they do have better temporal resolution than humans, then perhaps this may account for their superiority in discriminating complex patterns.

Yet another interpretation of the species differences in these experiments comes from auditory perception experiments in humans. Speech sounds represent extraordinarily familiar complex sounds with which humans have had extensive experience from an early age. Similar context effects have been reported in the visual domain as well. The "word superiority effect" refers to the fact that humans are better at visual letter perception when letters are presented in context of a word (for examples, see Johnston and McClelland, 1980; Reicher, 1969; Wheeler, 1970). Birds could be hearing these tonal patterns as a whole, perhaps like humans hear words. If this is so, then the species differences reported here may be due to the roles of learning strategies or differences in the focusing of auditory attention. Such factors have been shown to be important contributors to human performance in similar listening tasks (Leek and Watson, 1984, 1988). Following the lead of work on complex pattern perception in humans, these results suggest a useful future direction for experiments on the perception of species-specific vocal signals in birds. This direction would be to move beyond studies of the limits of resolution and begin to focus on attentional and memory processes that are recruited in the special processing of complex acoustic communication signals.

ACKNOWLEDGMENTS

This work was supported by NIH Grant Nos. DC-00198 and MH-00982 to R.J.D. and an Institutional NRSA from NIH (Grant No. DC-00046) to M.L.D. Special thanks to M. Leek, T. Wright, B. Lohr, and B. Brittan-Powell for comments on an earlier version of this manuscript, and to A. Nespor, E. Lichtenberg, M. Mavilia, T. Kidd, C. Moore, and M. Murphy for assistance with data collection.

- Amagai, S., Dooling, R. J., Formby, C., and Forrest, T. G. (1997). "Discrimination of silent temporal gaps in sinusoidal markers by the budgerigar (*Melopsittacus undulatus*)," *J. Acoust. Soc. Am.* **101**, 3124.
- Brown, S. D., Dooling, R. J., and O'Grady, K. (1988). "Perceptual organization of acoustic stimuli by budgerigars (*Melopsittacus undulatus*): III. Contact calls," *J. Comp. Psych.* **102**, 236–247.

- Dent, M. L., Dooling, R. J., and Leek, M. R. (1999). "Perception of harmonic complexes in budgerigars (*Melopsittacus undulatus*)," *J. Acoust. Soc. Am.* **105**, 1319.
- Dooling, R. J. (1986). "Perception of vocal signals by budgerigars (*Melopsittacus undulatus*)," *Exp. Biol.* **45**, 193–216.
- Dooling, R. J., Leek, M. R., and Dent, M. L. (1999). "Temporal resolution in birds and the perception of complex sounds," *J. Acoust. Soc. Am.* **105**, 1109.
- Dooling, R. J., and Okanoya, K. (1995). "The method of constant stimuli in testing auditory sensitivity in small birds," in *Methods in Comparative Psychoacoustics*, edited by G. M. Klump, R. J. Dooling, R. R. Fay, and W. C. Stebbins (Birkhauser Verlag, Basel, Switzerland), pp. 161–169.
- Dooling, R. J., Park, T. J., Brown, S. D., Okanoya, K., and Soli, S. D. (1987). "Perceptual organization of acoustic stimuli by budgerigars (*Melopsittacus undulatus*): II. Vocal signals," *J. Comp. Psych.* **101**, 367–381.
- Dooling, R. J., and Saunders, J. C. (1975). "Hearing in the parakeet (*Melopsittacus undulatus*): Absolute thresholds, critical ratios, frequency difference limens, and vocalizations," *J. Comp. Physiol. Psychol.* **88**, 1–20.
- Dooling, R. J., and Searcy, M. H. (1979). "The relation among critical ratios, critical bands, and intensity difference limens in the parakeet (*Melopsittacus undulatus*)," *Bull. Psychon. Soc.* **13**, 300–302.
- Dooling, R. J., and Searcy, M. H. (1985a). "Nonsimultaneous auditory masking in the budgerigar (*Melopsittacus undulatus*)," *J. Comp. Psych.* **99**, 226–230.
- Dooling, R. J., and Searcy, M. H. (1985b). "Temporal integration of acoustic signals by the budgerigar (*Melopsittacus undulatus*)," *J. Acoust. Soc. Am.* **77**, 1917–1920.
- Espinoza-Varas, B., and Watson, C. S. (1986). "Temporal discrimination for single components of nonspeech auditory patterns," *J. Acoust. Soc. Am.* **80**, 1685–1694.
- Espinoza-Varas, B., and Watson, C. S. (1989). "Pattern perception by humans," in *The Comparative Psychology of Audition*, edited by R. J. Dooling and S. Hulse (Lawrence Erlbaum, Hillsdale, NJ), pp. 67–94.
- Farabaugh, S. M., and Dooling, R. J. (1996). "Ecology and evolution of acoustic communication in parrots: Laboratory and field studies in parrots," in *Ecology and Evolution of Acoustic Communication in Birds*, edited by D. E. Kroodsma and E. H. Miller (Cornell University Press, Ithaca), pp. 97–117.
- French, N. R., and Steinberg, J. C. (1947). "Factors governing the intelligibility of speech sounds," *J. Acoust. Soc. Am.* **19**, 90–119.
- Freyman, R. L., and Nelson, D. A. (1987). "Frequency discrimination of short versus long duration tones by normal and hearing-impaired listeners," *J. Speech Hear. Res.* **30**, 28–36.
- Gray, L., and Rubel, E. W. (1987). "Development of auditory thresholds and frequency difference limens in chickens," in *Measurement of Audition and Vision in the First Year of Postnatal Life: A Methodological Overview*, edited by G. Gottlieb and N. Krasnegor (Ablex, Norwood, NJ), pp. 145–165.
- Green, D. M., and Swets, J. A. (1966). *Signal Detection Theory* (Wiley, New York).
- Hall, J. W., and Wood, E. J. (1984). "Stimulus duration and frequency discrimination for normal hearing and hearing impaired subjects," *J. Speech Hear. Res.* **27**, 252–256.
- Hartmann, W. M., Rakerd, B., and Packard, T. N. (1985). "On measuring the frequency-difference limen for short tones," *Percept. Psychophys.* **38**, 199–207.
- Henning, G. B. (1966). "Frequency discrimination of random amplitude tones," *J. Acoust. Soc. Am.* **39**, 336–339.
- Henning, G. B. (1970). "Effects of duration on frequency and amplitude discrimination," in *Frequency Analysis and Periodicity Detection in Hearing*, edited by R. Plomp and G. F. Smoorenburg (A. W. Sijthoff, Leiden, The Netherlands), pp. 350–359.
- Hienz, R. D., Sinnott, J. M., and Sachs, M. B. (1977). "Auditory sensitivity of the redwing blackbird (*Agelaius phoeniceus*) and the brown-headed cowbird (*Molothrus ater*)," *J. Comp. Physiol. Psychol.* **91**, 1365–1376.
- Johnston, J. C., and McClelland, J. C. (1980). "Experimental tests of a hierarchical model of word identification," *J. Verbal Learn. Verbal Behav.* **19**, 503–524.
- Kuhn, A., Leppelsack, H. J., and Schwartzkopff, J. (1980). "Measurement of frequency discrimination in the starling (*Sturnus vulgaris*) by conditioning of heart rate," *Naturwissenschaften* **67**, 102.
- Langemann, U., and Klump, G. M. (1992). "Frequency discrimination in

- the European starling (*Sturnus vulgaris*): A comparison of different measures," *Hearing Res.* **63**, 43–51.
- Leek, M. R., and Watson, C. S. (1984). "Learning to detect auditory pattern components," *J. Acoust. Soc. Am.* **76**, 1037–1044.
- Leek, M. R., and Watson, C. S. (1988). "Auditory perceptual learning of tonal patterns," *Percept. Psychophys.* **43**, 389–394.
- Lohr, B., and Dooling, R. J. (1998). "Detection of changes in timbre and harmonicity in complex sounds by zebra finches (*Taeniopygia guttata*) and budgerigars (*Melopsittacus undulatus*)," *J. Comp. Psych.* **112**, 36–47.
- Miller, G. A. (1981). *Language and Speech* (W. H. Freeman, San Francisco).
- Moore, B. C. J. (1973). "Frequency difference limens for short-duration tones," *J. Acoust. Soc. Am.* **54**, 610–619.
- Moore, B. C. J. (1997). *An Introduction to the Psychology of Hearing* (Academic, San Diego).
- Okanoya, K., and Dooling, R. J. (1987). "Hearing in passerine and psittacine birds: A comparative study of absolute and masked auditory thresholds," *J. Comp. Psych.* **101**, 7–15.
- Park, T. J., and Dooling, R. J. (1985). "Perception of species-specific contact calls by budgerigars (*Melopsittacus undulatus*)," *J. Comp. Psych.* **99**, 391–402.
- Park, T. J., and Dooling, R. J. (1986). "Perception of degraded vocalizations by budgerigars (*Melopsittacus undulatus*)," *Anim. Learning Behav.* **14**, 359–364.
- Penner, M. J. (1995). "Psychophysical methods," in *Methods in Comparative Psychoacoustics*, edited by G. M. Klump, R. J. Dooling, R. R. Fay, and W. C. Stebbins (Birkhauser Verlag, Basel, Switzerland), pp. 47–57.
- Prosen, C. A., Halpern, D. L., and Dallos, P. (1989). "Frequency difference limens in normal and sensorineural hearing impaired chinchillas," *J. Acoust. Soc. Am.* **85**, 1302–1313.
- Quine, D. B., and Konishi, M. (1974). "Absolute frequency discrimination in the barn owl," *J. Comp. Physiol.* **93**, 347–360.
- Reicher, G. M. (1969). "Perceptual recognition as a function of meaningfulness of stimulus material," *J. Exp. Psychol.* **81**, 275–280.
- Saunders, J. C., Rintelmann, W. F., and Bock, G. R. (1979). "Frequency selectivity in bird and man: A comparison among critical ratios, critical bands, and psychophysical tuning curves," *Hearing Res.* **1**, 303–323.
- Sinnott, J. M., Sachs, M. B., and Hienz, R. D. (1980). "Aspects of frequency discrimination in passerine birds and pigeons," *J. Comp. Physiol. Psychol.* **94**, 401–415.
- Turnbull, W. (1944). "Pitch discrimination as a function of tonal duration," *J. Exp. Psychol.* **34**, 302–316.
- Watson, C. S., Kelly, W. J., and Wroton, H. W. (1976). "Factors in the discrimination of tonal patterns. II. Selective attention and learning under various levels of stimulus uncertainty," *J. Acoust. Soc. Am.* **60**, 1176–1186.
- Watson, C. S., Wroton, H. W., Kelly, W. J., and Benbassat, C. A. (1975). "Factors in the discrimination of tonal patterns. I. Component frequency, temporal position, and silent intervals," *J. Acoust. Soc. Am.* **57**, 1175–1185.
- Wheeler, D. D. (1970). "Processes in word recognition," *Cogn. Psychol.* **1**, 59–85.

Feedback path variability modeling for robust hearing aids

B. Rafaely,^{a)} M. Roccasalva-Firenze, and E. Payne

*Institute of Sound and Vibration Research, University of Southampton,
Southampton SO17 1BJ, United Kingdom*

(Received 13 April 1999; revised 21 October 1999; accepted 28 January 2000)

Acoustic feedback is a common problem in hearing aids with vented earmolds. Hearing aids designed to work under normal conditions become unstable when the feedback path varies under changing conditions. A comprehensive study of the variability of the feedback path under various conditions and for different users is presented in this paper. A multiplicative uncertainty bound widely used in H -infinity robust control is suggested to model the variations, which is then used to formulate a robust stability condition for the hearing aid. The upper limit of the closed-loop acoustic gain of the hearing aid for maintaining robust stability is also derived. Examples of robust constant amplification hearing aids, which maintain stability in the face of the given variations in the feedback path, are presented. The robust stability condition is also suggested as a tool to design more robust digital hearing aids. © 2000 Acoustical Society of America.

[S0001-4966(00)01605-2]

PACS numbers: 43.66.Ts, 43.66.Yw [SPB]

INTRODUCTION

Acoustic feedback is a severe problem in hearing aids. The feedback occurs when the receiver of the aid produces an acoustic signal that leaks back to the microphone. This causes instability and oscillations leading to a high-frequency whistling sound produced by the hearing aid. Reducing the hearing aid gain reduces the feedback signal, but also degrades the hearing aid performance. In most cases the vent is the principle cause of the acoustic feedback in the hearing aids. Venting, however, is necessary to provide a more natural sound by reducing the occlusion effect that hearing aid users often experience when venting is insufficient, as discussed by the Hearing Aid Audiology Group, BSA (1984), and to reduce the low-frequency gain. The feedback problem is worse for in-the-ear and in-the-canal hearing aids due to the shorter distance between the microphone and the vent opening, and the increased coupling between the receiver and the microphone.

Hearing aids that perform well when fitted might go unstable when the feedback path varies due to face movements or when placing an object such as a telephone handset near the ear. Large variations in the feedback path force lower gains on the hearing aid if stability is to be maintained, thus degrading the hearing aid performance and utility. Various approaches have been reported which attempt to reduce the feedback problem. Low-pass filtering has been used to roll off the high-frequency response of the hearing aid, which is where instability usually occurs, but this can detrimentally attenuate the high-frequency gain of the hearing aid, as shown by Preves *et al.* (1986). More complex methods for feedback suppression involve various techniques of digital signal processing, where adaptive filters are used to generate an estimate of the feedback signal and subtract it from the feedback signal measured by the microphone, thus

reducing its effect. Variations of this approach with different implementations of the adaptive filter are reported in recent studies by Kates (1991), Joson *et al.* (1993), Maxwell and Zurek (1995), Siqueira *et al.* (1996), Wyrsh and Kaelin (1997), and Kates (1999).

Although many of the adaptive feedback cancellation systems are useful in reducing the feedback problem, the adaptive feedback-path model is never perfect, and so the feedback signal is never completely canceled. Instability can therefore still occur, although at higher hearing aid gains. Therefore, to maintain stability in a hearing aid system, whether an adaptive feedback-path cancellation filter is used or not, the hearing aid system must be designed to allow for some feedback signal without experiencing instability. In other words, robust stability must be maintained for some given variability or uncertainty in the feedback path, as discussed in Skogestad and Postlethwaite (1996), for example.

The aim and contribution of this paper is the comprehensive analysis of the variability of the feedback path between users and under various conditions, and the formulation of a robust stability condition from the bounds on the feedback-path variability. This is then used to derive an upper limit on the closed-loop acoustic gain of a robustly stable hearing aid. The results presented here can be employed in the design of more robust hearing aids.

The paper is organized as follows. Section I presents the theoretical formulation of the robust stability condition for the hearing aid system, and in Sec. II the experimental analysis of the feedback-path variability is described. The multiplicative uncertainty model for the variations are evaluated in Sec. III, and in Sec. IV the performance of constant-gain robust hearing aids is presented, followed by conclusions.

I. THEORY

A. Hearing aid system

A modified Starkey 11M in-the-ear hearing aid attached to a customized earmold, as used in this study, is shown in

^{a)} Author to whom correspondence should be addressed. Electronic mail: br@isvr.soton.ac.uk

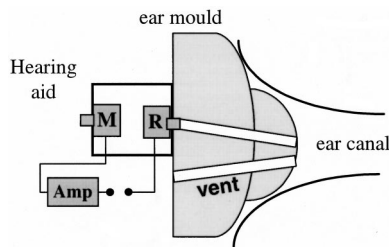


FIG. 1. The hearing aid system.

Figs. 1 and 2. A description of the modified hearing aid is detailed in Sec. II. The principal components of the system are the microphone (M), the hearing aid circuit (amplifier), the receiver (R), and the vent. The external sound is detected by the microphone, amplified by the hearing aid circuit, and transmitted into the ear canal via the receiver. The vent connects the ear canal to the outside, and transmits the sound back from the ear canal to the microphone causing the feedback problem. The hearing aid feedback mechanism is well explained by Kates (1988) and Egolf *et al.* (1985), for example, and is illustrated by the block diagram in Fig. 3, where P_{ext} is the sound pressure outside the ear, P_{ear} is the sound pressure in the ear canal, A is the transfer function of the unamplified signal path directly through the vent into the ear canal, B is the transfer function of the acoustic feedback path from the ear canal through the vent and other leaks, M is the electroacoustic transfer function of the microphone, R is the electroacoustic transfer function of the receiver, and H is the transfer function of the hearing aid circuit, in this study it represents a constant gain over frequency.

The open-loop acoustic gain is defined in the frequency domain, omitting the explicit dependence on frequency for simplicity, as

$$F = MHR, \quad (1)$$

and the open-loop gain with unity amplification, $H=1$, is defined as

$$F_0 = MR. \quad (2)$$

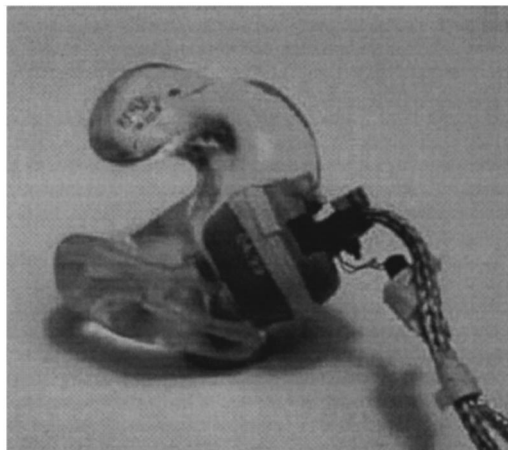


FIG. 2. A modified Starkey 11M modular in-the-ear hearing aid attached to an earmold.

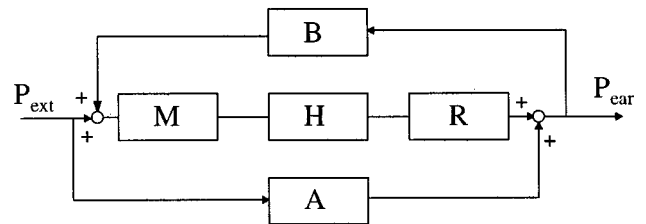


FIG. 3. A system block diagram for a vented hearing aid.

The feedback path, which is the response from the input signal of the receiver to the output signal of the microphone, is defined as

$$G = RBM, \quad (3)$$

and the complete acoustic transfer function of the system, also referred to as the closed-loop acoustic gain, is written using Fig. 3 as

$$\frac{P_{\text{ear}}}{P_{\text{ext}}} = A + \frac{MHR}{1 - MHRB} = A + \frac{F}{1 - GH} = A + \frac{F_0H}{1 - GH}. \quad (4)$$

B. Feedback path and closed-loop stability

Feedback occurs when a fraction of the receiver signal is transmitted back to the microphone, through the acoustic path B as shown in Fig. 3. The acoustic path is the most significant contributor to the feedback signal, although electrical and mechanical paths also exist. In this study, only the acoustical component is considered, although the contributions from the other paths could also be analyzed in a similar manner. Instability occurs when a frequency component of the feedback signal arrives at the microphone in phase with the external signal component, but with greater magnitude. Then oscillation will occur, driving the hearing aid to its maximum level and impeding its normal operation.

A block diagram of the feedback part in the hearing aid system is illustrated in Fig. 4, where G is the feedback path and H is the amplifier gain. The open-loop transfer function is defined as $L = -GH$, with the minus sign resulting from the positive summation of the feedback signal. The system of Fig. 4 is stable according to the Nyquist criterion if the Nyquist plot of L , which is itself stable in this case, does not encircle the critical $(-1, j0)$ point [see, for example, Skogestad and Postlethwaite (1996)]. Bisgaard and Dyrland (1991) suggested that a gain margin of approximately 5 dB from the instability point might be sufficient to stop the howling in normal use.

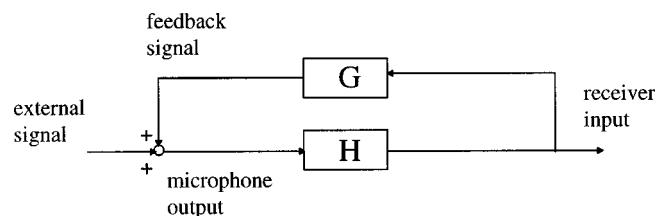


FIG. 4. A system block diagram of the hearing aid as a feedback system.

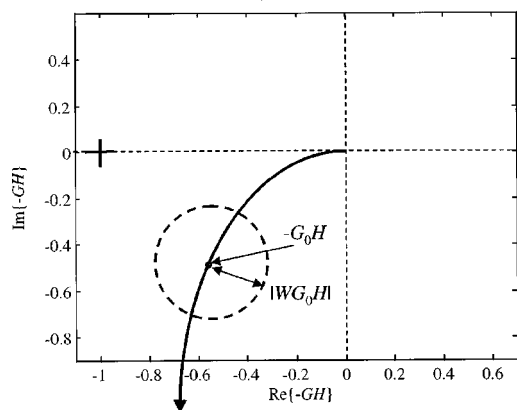


FIG. 5. Nyquist plot of the open-loop system $(-GH)$, showing the multiplicative uncertainty bound W as a circle around the nominal value.

C. Feedback path variability and robust stability

The knowledge of the variability of the feedback path can be useful in designing robustly stable hearing aids that maintain stability even when the response of the path changes. One way to describe variations or uncertainty, which is widely used in H_∞ feedback control, as described in Skogestad and Postlethwaite (1996), for example, is by a multiplicative uncertainty around a nominal model. If G_0 is the nominal response of the feedback path, i.e., the response with normal fit under normal conditions of use, and \tilde{G} is a set of measured feedback-path response functions taken under various conditions of use, then any response G from the set \tilde{G} can be written in terms of the nominal response as

$$G = G_0(1 + \Delta), \quad (5)$$

where Δ can take any complex value, and is a function of frequency. Δ can be written as

$$\Delta = \frac{G}{G_0} - 1, \quad (6)$$

and its maximum value at any frequency for the complete set \tilde{G} , denoted by the real frequency-dependent function W , is given by

$$W = \max_{G \in \tilde{G}} |\Delta| = \max_{G \in \tilde{G}} \left| \frac{G}{G_0} - 1 \right|. \quad (7)$$

Considering an illustration of the Nyquist plot of an open-loop system $-GH$ as shown in Fig. 5, the system with responses described by the multiplicative uncertainty bound W for a given frequency is represented as being inside a circle of radius $|WG_0H|$ centered at $-G_0H$. This uncertainty model fits reasonably well the variations of the feedback path, as shown in Sec. III below, especially at high frequencies, and leads to a simple stability criterion, as shown next. The hearing aid system, as shown in Fig. 4, will be robustly stable, i.e., stable for all given variations in the feedback path, as long as the circle of uncertainty does not encircle the $(-1, j0)$ point for all frequencies. This follows directly from the Nyquist stability criterion discussed above. Given

$$L_0 = -G_0H \quad (8)$$

as the open-loop response for the nominal feedback path G_0 , such that the distance between the circle center and the $(-1, j0)$ point is $|1 + L_0|$, then to maintain stability the circle radius must be smaller than that distance. The condition for robust stability is therefore written as

$$|WL_0| < |1 + L_0| \Rightarrow \frac{|WL_0|}{|1 + L_0|} < 1 \text{ for all frequencies,} \quad (9)$$

which can be written as

$$|WT_0| < 1 \text{ for all frequencies,} \quad (10)$$

or using the ∞ norm, as

$$\|WT_0\|_\infty < 1, \quad (11)$$

where T_0 is the nominal complementary sensitivity function given by $L_0/(1 + L_0)$. The reader is referred to Skogestad and Postlethwaite (1996), for example, for further details on the derivation above. The condition for robust stability given by Eq. (10) means that the hearing aid system with gain H and nominal feedback path G_0 will maintain stability for any variation of the feedback path from G_0 , as long as this variation is bounded by W , and Eq. (10) is maintained.

The maximum gain of the hearing aid response $|P_{\text{ear}}/P_{\text{ext}}|$ that satisfies the robust stability condition can now be calculated. It is first assumed that $|A| \ll |F_0H/(1 - GH)|$ in Eq. (4), which means that the contribution of the forward acoustic path, A , also referred to as the real-ear occluded response (REOR) in Mueller *et al.* (1992), is negligible compared to the contribution of the amplified path. Then, the hearing aid response can be approximated as

$$\frac{P_{\text{ear}}}{P_{\text{ext}}} \approx \frac{F_0H}{1 - GH}. \quad (12)$$

The robust stability condition requires $|WT_0| < 1$, or

$$\left| \frac{H}{1 - G_0H} \right| < \left| \frac{1}{G_0W} \right|. \quad (13)$$

Substituting this inequality in Eq. (12), the maximum hearing aid gain can be written as

$$\left| \frac{P_{\text{ear}}}{P_{\text{ext}}} \right|_{\text{max}} \approx \left| \frac{F_0}{G_0W} \right|. \quad (14)$$

Given the forward response with unit amplifier gain, F_0 , and the response of the nominal feedback path, G_0 , the maximum achievable closed-loop acoustic gain while maintaining robust stability depends on the feedback path variability, W , and is given above.

Setting $W=1$, which gives a gain margin of 6 dB, the maximum gain becomes

$$\left| \frac{P_{\text{ear}}}{P_{\text{ext}}} \right|_{\text{max}} \approx \left| \frac{F_0}{G_0} \right|. \quad (15)$$

This is equivalent to the condition presented by Hellgren *et al.* (1999) as the limit on the real-ear aided gain (REAG), or the real-ear aided response (REAR), as defined in Mueller *et al.* (1992), which was derived by the stability condition

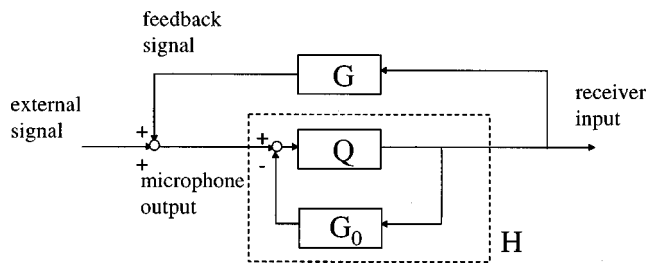


FIG. 6. A system block diagram of the hearing aid with a feedback cancellation filter.

$|GH| < 0.5$, i.e., a 6-dB gain margin. However, the condition derived here in Eq. (14) takes into account the actual variability of the feedback path, as a function of frequency, and in general is less conservative than the condition with a scalar W , and can therefore allow for higher hearing aid gain.

D. Robust stability with a feedback cancellation filter

Contemporary digital hearing aids employ a feedback cancellation filter, connected in parallel to the physical feedback path to cancel the feedback signal, as described in the papers by Kates (1999) and Maxwell and Zurek (1995), for example. This allows higher gains before feedback occurs and therefore better performance. Figure 6 shows a block diagram of a hearing aid system where the hearing aid filter H is now replaced by a hearing aid filter Q and a feedback cancellation filter G_0 . The cancellation filter is denoted by G_0 since it is assumed that it models the nominal feedback path, i.e., the response under normal conditions, as denoted above by G_0 . The overall response of filter H can therefore be written as

$$H = \frac{Q}{1 + QG_0}. \quad (16)$$

Substituting this equation in the condition for robust stability, Eq. (10), the new condition for robust stability is given by

$$|WQG_0| < 1 \quad \text{for all frequencies.} \quad (17)$$

This is a simpler condition which directly relates the response of the two filters, Q and G_0 , to robust stability. This condition is similar to the robust stability condition derived for internal model controllers (IMC), as detailed in Morari and Zafriou (1989), where in this case the feedback cancellation filter is equivalent to the internal model.

A condition on the magnitude limit of the hearing aid filter Q can be introduced by rearranging the variables in Eq. (17) as follows:

$$|Q| < \frac{1}{|WG_0|} \quad \text{for all frequencies.} \quad (18)$$

This is a useful gain limit on the hearing aid amplifier or forward filter, which depends on the gain of the feedback cancellation filter G_0 , and the multiplicative bound on the feedback-path variations from G_0 .

The condition in Eq. (17) can be used in two ways. First, for a hearing aid system with fixed (nonadaptive) filters, the filter Q can be designed for best compensation while main-

taining the robust stability condition, given G_0 as the model of the feedback path under normal conditions. A simple way to achieve the highest possible gain while maintaining robust stability is to set the magnitude of Q equal to the limit in Eq. (18). The impulse response of a finite impulse response (FIR) filter Q can then be derived simply by taking the inverse Fourier transform of the magnitude value, to form a symmetric linear phase filter, or using other filter design methods to generate minimum-phase filters, for example. Moreover, a more general constrained optimization can be solved for setting the magnitude of Q to a desired response required from the hearing compensation, while maintaining the robust stability condition. The design of such robust digital filters is the subject of current research with initial results described in Rafaely *et al.* (1999).

Second, for adaptive feedback cancellation filters, Eq. (17) can be used as a constraint on the adaptive filter, which will ensure robust stability. In this case the value of W would be reduced compared to a fixed cancellation filter implementation, since W , which represents the maximum deviation of the feedback path from the model G_0 , is smaller if the model is made adaptive to track variations in the feedback path. Kates (1999) recently presented an adaptive feedback cancellation filter with a constraint on the filter to avoid cancellation of tones in the input signal. Kates used a constraint on the deviation of the feedback cancellation filter from its initial value, which was set to a constant 10 dB. The work presented here proposes to take into account the actual bound on the variability between the model and the feedback path, as a function of frequency, using a multiplicative uncertainty model, and then use the resulting constraint in the adaptation process. The estimation of W for an adaptive filter and the use in an adaptive feedback cancellation system are the subjects of current research.

II. EXPERIMENT 1—THE FEEDBACK-PATH VARIABILITY

The feedback path of a vented hearing aid plays a dominant role in the feedback problems encountered by hearing aid users. Once fitted, the hearing aid could become unstable if the feedback path changes. Therefore, knowledge of the feedback path and its variability in real ears under realistic conditions will assist in the design of hearing aids that are more robust and suffer less from the feedback problem. The acoustical feedback path G in human ears can change for different subjects, ears, and conditions. The causes of these changes are outlined below.

- Increase in the sound leaking from the ear to the outside. The movements of the head and of the mouth when smiling, laughing, or eating, and other conditions causing changes in the hearing aid fit, can change the level of the sound that leaks to the microphone.
- Change in the sound-pressure level inside the ear canal. The acoustic characteristics of each ear determined by the ear-canal dimensions and eardrum impedance are unique, so the sound pressure developed inside the ear in response to the same receiver input will vary between ears.

- (c) Change in the level of the feedback signal transmitted from the vent and other leak paths to the microphone. A hand or a telephone handset covering the ear can increase the level of the feedback signal detected by the microphone.

An experiment was conducted on each ear of four otologically normal subjects to analyze the variability of the feedback path in various conditions that cause the changes described above. An earmold (Starkey 2206 clear hard acrylic full-shell) was produced from a silicon impression of the ear for the left and right ears of all subjects. The earmolds included a vent with a length varying from 12–16 mm, and a diameter varying from 1.6 to 2.1 mm. A second tube was formed in the mold to acoustically connect the receiver output to the ear canal, which was of length similar to the vent and a diameter of 1.4 mm. The custom earmolds were then fitted to a modular hearing aid (Starkey ITE 11M), which was modified to facilitate the measurements. First, the receiver (Knowles ED-1975) and microphone (Knowles EM-3046) were disconnected electrically from the hearing aid amplifier, but remained fixed in their location. External wires were then connected to the transducers' terminals. The microphone was then connected to an external microphone amplifier, which was built around two Analog Device OP27 operational amplifiers. The amplifier was designed to produce constant gains of 20, 40, or 60 dB. Figures 1 and 2 illustrate the hearing aid used in this experiment and the way it was attached to the mold.

The receiver was connected to the output of a signal generator that was part of a two-channel spectrum analyzer (Advantest R9211C), and the microphone amplifier was connected to the analyzer input. The signal generator produced a broadband noise in the range 300–10 000 Hz and a level of 0.2 V rms at the receiver input. This amplitude was chosen to ensure that the noise level inside the ear canal was lower than 91 dBA, to avoid any physical damage to the ear, as suggested by the safety guidelines in ISVR (1996). The frequency response of the feedback path was then measured as the response between the analyzer output signal and the microphone amplifier output signal using an averaging procedure, and recorded on the analyzer disk.

The subjects were seated in an anechoic chamber during the experiment, where an otoscopic examination of their ears was performed prior to fitting the earmold with the modular hearing aid attached. The analyzer was placed outside the chamber and the operator communicated with the subject through the chamber's glass window. The frequency response of the feedback path in each ear was then measured under eight conditions. The conditions were chosen to represent various situations of everyday use, as described below.

- (1) First normal fitting. The hearing aid attached to the earmold was fitted to the subject's ear. The subject was sitting still.
- (2) Second normal fitting. The earmold and hearing aid were removed from the ear and refitted.
- (3) Hand positioned very close to the ear, almost covering the ear completely. This represented a situation of a hand or a telephone handset placed near the ear.

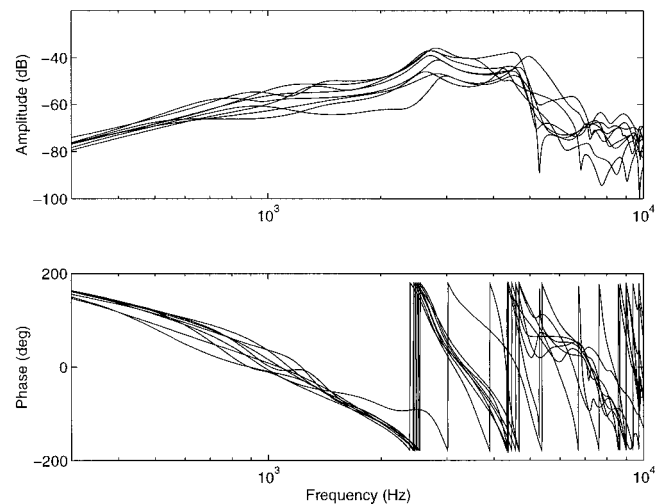


FIG. 7. Frequency-response curves of the feedback path G of all ears for the first normal fitting (condition 1).

- (4) Hand positioned 10 cm from the ear. This condition represents a hand or other object placed some distance from the ear.
- (5) Hand positioned 20 cm from the ear. This condition represents a hand or other object placed further away from the ear than in condition (4).
- (6) Open mouth. This condition occurs during speaking and eating.
- (7) Smiling face.
- (8) A wrong fitting, with the top prong of the earmold removed from its position in the helix of the outer ear, such that the earmold is no longer fitted tightly to the ear.

Figure 7 shows the various frequency-response curves of the feedback path G in the normal fit (condition 1), as measured in this experiment for the left and right ears of all subjects. From 300 Hz to about 1 kHz the difference between the various curves is about 10 dB, after which the difference increases to a value of about 50 dB near 5 kHz. Figures 8–10 show the frequency-response curves of G with the hearing aid fitted to one subject, under the different measurement conditions. This is presented as an example of a set of feedback-path response functions that could be used to compute the bound on the variability as in Eq. (7). In the measurements with a hand near the ear (Fig. 8, conditions 3–5), it is evident that only when the hand is very close to the ear, the variation from the normal response (the first fitting) is large, particularly below 1 kHz, where there is about 20 dB difference. Above 1 kHz, the effect of a hand near the ear seems to be smaller.

Figure 9 shows the measurement of the feedback path with an open mouth and smiling face (conditions 6 and 7), which was found to be very similar to the measurement with the normal fit, i.e., movement of the face and mouth was found to have little effect on the feedback path. It is acknowledged, however, that when face movements do change the fitting of the hearing aid and increase the leaks, they could have a more significant effect on the feedback path. The measurements with wrong fitting and a second normal

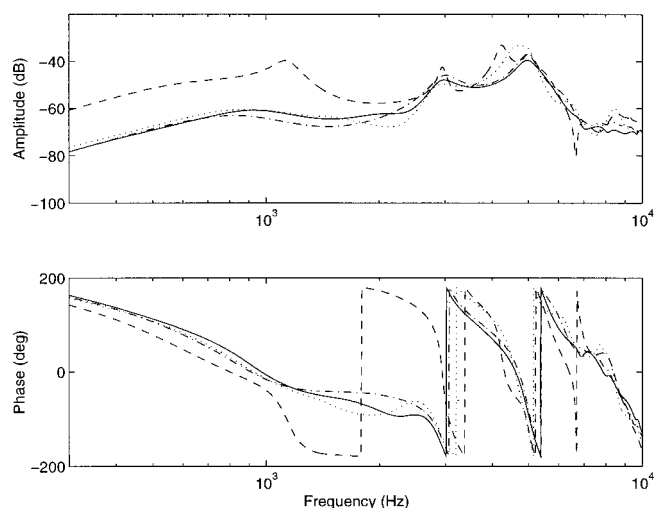


FIG. 8. Frequency-response curves of the feedback path of one ear in different conditions of measurement: first normal fitting (condition 1, solid curve), hand positioned very close to the ear (condition 3, dashed curve), hand positioned 10 cm from the ear (condition 4, dotted curve), and hand positioned 20 cm from the ear (condition 5, dash-dotted curve).

fitting (conditions 2 and 8) are shown in Fig. 10. Very small differences from the first normal fit were found below 2.8 kHz, while above 2.8 kHz the amplitude of the curve corresponding to the wrong fit varied up to 10 dB from the normal fit.

Most conditions seem to produce a variation in the feedback path which is smaller than 5–10 dB, except the condition with a hand very close to the ear, which produced variations of about 20 dB. Variability among subjects also produced changes in the magnitude of the feedback path on the order of 20 dB.

III. EVALUATION OF THE MULTIPLICATIVE UNCERTAINTY MODEL

The variation of the measured feedback path was analyzed at four discrete frequencies, i.e., 0.5, 1, 3, and 5 kHz.

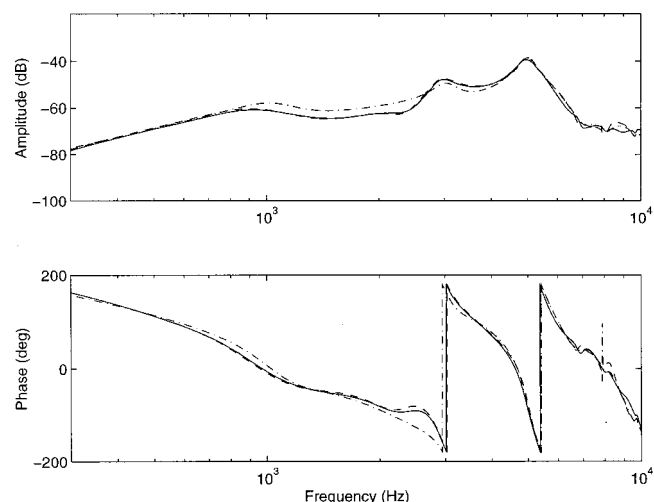


FIG. 9. Frequency-response curves of the feedback path of one ear in different conditions of measurement: first normal fitting (condition 1, solid curve), open mouth (condition 6, dashed curve), and smiling face (condition 7, dash-dotted curve).

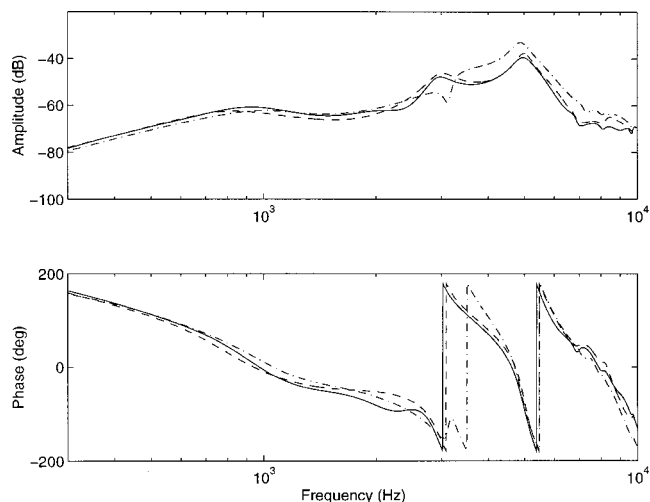


FIG. 10. Frequency-response curves of the feedback path of one ear in different conditions of measurement: first normal fitting (condition 1, solid curve), second normal fitting (condition 2, dashed curve), and wrong fitting (condition 8, dash-dotted curve).

The complex values of the frequency response functions at these frequencies were plotted using a Nyquist plot, to evaluate the “spread” of the response in the various conditions. Two sets of measurements were defined and analyzed separately. The first included the complete measurement set with all ears under all conditions. The second set included data from a single ear under all conditions, taking into account that hearing aids are customized to individual subjects, thus excluding any intersubject variability.

Figure 11 shows, the Nyquist plot of the feedback path of one ear under normal conditions at the full frequency range (solid curve), and the complex values of the feedback path of all ears under all conditions at the selected frequency only, marked by “×.” The figure also shows a circle that forms a bound around all complex values, which is used in

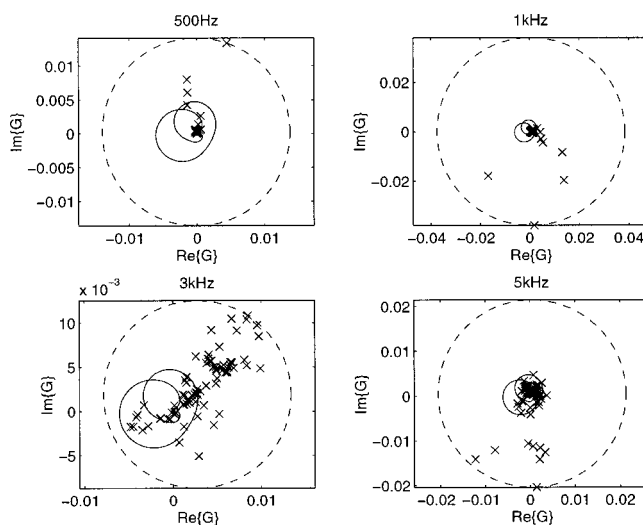


FIG. 11. Nyquist plot of the feedback-path variability, which includes all subjects under all conditions at several discrete frequencies, as shown above. The response of one ear under normal conditions is presented at the full frequency range (solid curve), while all other complex response values marked by “×” are presented at the four discrete frequencies. A circle enclosing all response values is also presented.

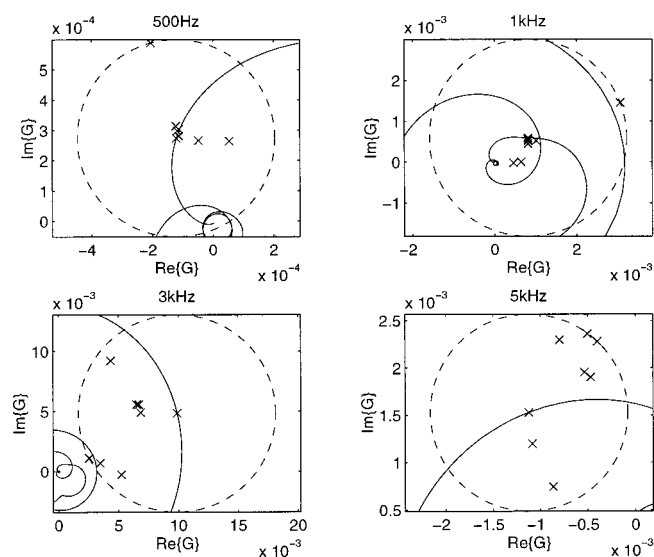


FIG. 12. Nyquist plot of the feedback-path variability, which includes one ear under all conditions at several discrete frequencies as shown above. The response for that ear under normal conditions is presented at the full frequency range (solid curve), while all other complex response values marked by “x” are presented at the four discrete frequencies. A circle enclosing all response values is also presented.

the robust stability condition as discussed above. At lower frequencies, 500 and 1000 Hz, most values are clustered together, while few values spread far from the rest. These extreme values are due to the more extreme condition of placing a hand very close to the ear, which generates large low-frequency variations. At higher frequencies, 3 and 5 kHz, the variations are spread more evenly, which suggests that these variations occur more randomly between subjects. Note that the panels are scaled to better illustrate the various complex values at the selected frequencies, and that the solid curve representing the feedback path of one ear under normal conditions is actually the same in all four panels.

Figure 12 shows the feedback-path variations at the same discrete frequencies but for an individual ear, where only the measurements with that ear under all conditions were included. It can be seen that here, as well, the values at low frequencies are clustered closer together except for the value corresponding to the condition with a hand placed very close to the ear, while at high frequencies they are spread more evenly in the complex plane.

From the analysis of these measurement sets, it is evident that the spread of the values in the complex plane form more clustered shapes that fill only a small part of the circular bound at low frequencies, while at higher frequencies the spread is more even and the various values fill the circular bound more evenly. This analysis shows that the multiplicative uncertainty model is a reasonable model in this case, especially at the higher frequencies.

Figure 13 shows the multiplicative uncertainty bound W calculated using Eq. (7) for the measurement set that included all data (upper curve), and for the measurement sets that included data from individual ears (lower curves). With all data included, the multiplicative uncertainty bound showed high values at lower frequencies due to the condition of placing a hand very close to the ear, and at high frequen-

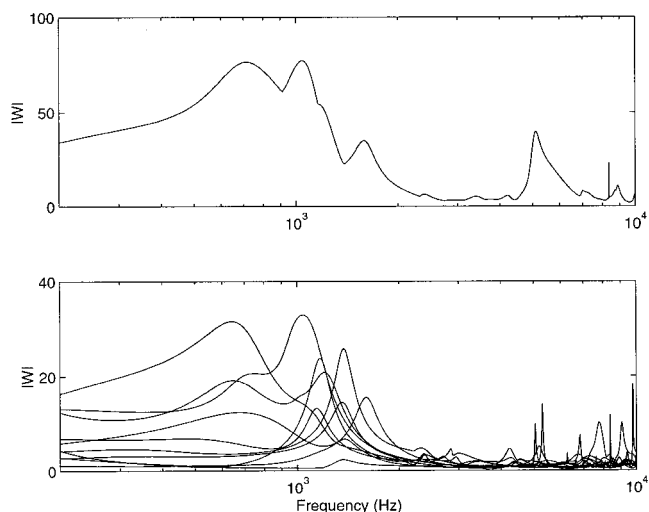


FIG. 13. The multiplicative uncertainty bound W for the measurement set with all subjects under all conditions (upper curve), and the multiplicative uncertainty bound W for individual ears under all conditions (lower curve).

cies due to the intersubject variability and the other conditions. The response with a normal hearing aid fitting for one specific subject was used as the nominal response G_0 in this case. With data from individual ears (lower curves), the multiplicative uncertainty bound changes significantly between subjects, due to the intersubject variability, and is also characterized by high values at both the low-frequency and high-frequency bands. The response with a normal hearing aid fitting for each individual ear was used as the nominal response G_0 in this case.

IV. EXPERIMENT 2—ROBUSTLY STABLE HEARING AID

A second experiment was conducted to examine the performance of constant-gain hearing aids that are designed to be robustly stable for the variations in the feedback path as measured above. The performance was assessed from the closed-loop acoustic gain of the hearing aid, denoted by the ratio $P_{\text{ear}}/P_{\text{ext}}$ in Fig. 3, which was measured in this experiment. A probe microphone measurement was performed according to the procedure described by Mueller *et al.* (1992).

A probe tube of inner diameter 0.7 mm, outer diameter of 1.2 mm, and a length of 85 mm was connected to the outlet of a Knowles EM3046 microphone and inserted into the ear canal to within 5 mm of the eardrum. The modified modular hearing aid and custom earmold used in the previous experiment were then fitted to the ear of one subject, and another Knowles EM3046 microphone was placed outside the ear near the hearing aid. Both microphones were connected to the microphone amplifiers used in the previous experiment. This way, the acoustic response between the pressure outside the ear and the pressure inside the ear could be estimated from the response measurement between the two microphones. The response was measured by the spectrum analyzer (Advantest R9211 C), which generated broadband random noise at a level of 85 dBA via a loudspeaker placed 50 cm from the subject. To account for the differences between the response of the microphones and the ef-

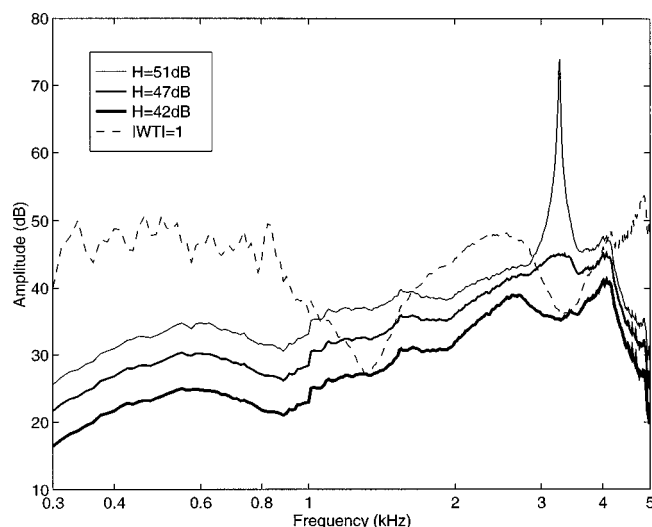


FIG. 14. Hearing aid response curves (closed-loop acoustic gain) calculated using Eq. (4) from measured data, for three values of the amplifier gain as noted above (solid curves), and the upper limit of the hearing aid gain calculated from Eq. (14) (dashed curve).

fect of the probe tube, a calibration was performed, as described in Mueller *et al.* (1992), by placing the tube opening close to the outlet of the outer microphone, both placed about 50 cm from the loudspeaker, and measuring the frequency response between the two microphones. This response was then used to correct the measured response functions when the tube microphone was placed in the ear. In addition, an otoscopic examination was conducted before the insertion of the hearing aid to the subject's ear in order to ensure correct placement of the probe tube in the ear canal.

The experiment was conducted with one subject, where various responses were measured as follows:

- (1) The response between the two microphones with the hearing aid fitted to the ear but switched off was first measured, denoted by the real-ear occluded response (REOR). This response is equal to the forward acoustic path denoted by A in Fig. 3.
- (2) The response between the two microphones with the hearing aid fitted and switched on was then measured, denoted by the real-ear aided response (REAR), and is equivalent to the response $P_{\text{ear}}/P_{\text{ext}}$ presented in Eq. (4). An amplification of 40 dB was used for the hearing aid amplifier in this case.
- (3) The response of the feedback path between the input of the receiver and the output of the hearing aid microphone amplifier was also measured, as in the first experiment, for the eight different conditions previously outlined.

The feedback-path variability bound, W , was calculated from Eq. (7) for the various feedback-path measurements, and the robust stability condition as in Eq. (10) was then computed using the feedback path under normal conditions (condition 1), as the nominal response G_0 , and the gain of the amplifier, H .

The response F_0 as in Eq. (4) was calculated from the measured response functions of A (the REOR), G (under

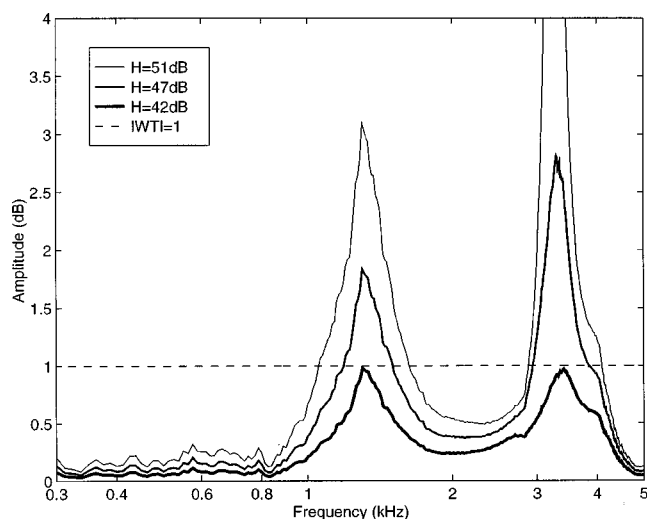


FIG. 15. The robust stability condition $|WT_0|$ corresponding to the four hearing aid response curves shown in Fig. 14.

condition 1), the REAR denoted by $P_{\text{ear}}/P_{\text{ext}}$ in the equation, and the hearing aid amplifier gain H , which was set to 40 dB. Given F_0 , the REAR was then calculated for other values of H using Eq. (4).

Figure 14 presents the calculated REAR for values of $H=42$, 47, and 51 dB (solid lines). With $H=51$ dB, the hearing aid is close to instability, and howling occurs at 3.2 kHz. Reducing the gain to $H=47$ dB stops the howling, while reducing the gain to $H=42$ dB reduces the hearing aid response even further at the howling frequency. The peak of the response at 600 Hz is due to the Helmholtz resonator between the receiver tube and the ear-canal volume, and the peaks at the higher frequencies are due to the response of the receiver. Figure 15 shows the quantity $|WT_0|$ used in the robust stability condition as presented in Eq. (10) for the three hearing aid gains as described above. The large values for $H=51$ dB at 3.2 kHz suggest that the hearing aid is already close to instability, as was indeed shown above. With a gain of $H=47$ dB, the hearing aid is stable but does not satisfy the robust stability condition of $|WT_0| < 1$, and could become unstable if the feedback path was to vary. Reducing the gain to $H=42$ dB ensures robust stability, since $|WT_0| < 1$ for all frequencies, but in this case the gain of the hearing aid response is also reduced, thus reducing performance.

The dashed curve in Fig. 14 represents the maximum gain of the hearing aid response $|P_{\text{ear}}/P_{\text{ext}}|$ that satisfies the robust stability condition, as derived in Eq. (14). The corresponding robust stability is represented by the dashed curve in Fig. 15, and is equal to unity at all frequencies, as expected from a robust hearing aid with a maximal gain. It is clear that at a wide frequency range, a higher gain can be achieved without compromising robust stability, although at two frequencies, 1.5 kHz and 3.4 kHz, the gain is already maximal, which is also illustrated by the robust stability term $|WT_0|$ being equal to unity at these frequencies. This shows that better performance could potentially be achieved by designing hearing aid filters that satisfy robust stability more

tightly. The detailed design methods of such filters, based on computing the response of digital FIR filters that satisfy Eq. (18), are the subject of current research work, as described in Rafaely *et al.* (1999).

V. CONCLUSIONS

The variability of the feedback path in hearing aids was studied in this paper. An experiment was conducted to measure the variability of the feedback path for different subjects under various conditions. It was shown that the variations could be modeled reasonably well using a multiplicative uncertainty bound, widely used in robust control. A robust stability condition for the hearing aid system, which is a function of the uncertainty bound, was then formulated. Constant amplification (over frequency) hearing aids were then designed for maximum gain while maintaining robust stability under various conditions, showing clearly the tradeoff between performance and robustness. The robust stability condition was also used to formulate an upper limit on the closed-loop acoustic gain of the hearing aid for maintaining robust stability. The design of robust hearing aids can be extended to digital hearing aids, where best compensation and robust stability are achieved simultaneously. This is currently being studied.

ACKNOWLEDGMENT

The authors would like to thank the University of Southampton for the support under Grant No. A97/26.

- Bisgaard, N., and Dyrland, O. (1991). "Acoustic feedback I Traditional feedback suppression methods," *Hear. Instrum.* **42**(9), 24–26.
- 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.

- Hearing Aid Audiology Group, BSA (1984). "The Earmould, current practice and technology," Hearing Aid Audiology Group, British Society of Audiology.
- Hellgren, J., Lunner, T., and Arlinger, S. (1999). "System identification of feedback in hearing aids," *J. Acoust. Soc. Am.* **105**, 3481–3496.
- ISVR (1996). "Guide to Experimentation involving Human Subjects," Technical Memorandum No. 808, Institute of Sound and Vibration Research, University of Southampton, UK, 1996.
- Jose, H. A. L., Asano, F., Suzuki, Y., and Sone, T. (1993). "Adaptive feedback cancellation with frequency compression for hearing aids," *J. Acoust. Soc. Am.* **94**, 3248–3254.
- 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.
- Kates, J. M. (1991). "Feedback cancellation in hearing aids: Results from a computer simulation," *IEEE Trans. Signal Process.* **39**(3), 553–562.
- Kates, J. M. (1999). "Constrained adaptation for feedback cancellation in hearing aids," *J. Acoust. Soc. Am.* **106**, 1010–1019.
- Maxwell, J. A., and Zurek, P. M. (1995). "Reducing acoustic feedback in hearing aids," *IEEE Trans. Speech Audio Process.* **3**(4), 304–313.
- Morari, M., and Zafriou, E. (1989). *Robust Process Control* (Prentice-Hall, Englewood Cliffs, NJ).
- Mueller, H. G., Hawkins, D. B., and Northern, J. L. (1992). *Probe Microphone Measurements—Hearing Aid Selection and Assessment* (Singular, San Diego, CA).
- Preves, D. A., Sigelman, J. A., and LeMay, P. R. (1986). "A feedback stabilizing circuit for hearing aids," *Hear. Instrum.* **37**(4), 35–36.
- Rafaely, B., Roccasalva-Firenze, M., and Payne, E. (1999). "Control of feedback in hearing aids," Technical Memorandum No. 836, Institute of Sound and Vibration Research, University of Southampton, UK.
- Siqueira, M. G., Speece, R., Petsalis, V., Alwan, A., Soli, S., and Gao, G. (1996). "Subband adaptive filtering applied to acoustic feedback reduction in hearing aids," *Proceeding of the IEEE Asilomar Conference on Signal System, and Computer*, Pacific Grove, November, 1996, pp. 778–782.
- Skogestad, S., and Postlethwaite, I. (1996). *Multivariable Feedback Control* (Wiley, New York).
- Wyrsh, S., and Kaelin, A. (1997). "A DSP implementation of a digital hearing aid with recruitment of loudness compensation and acoustic echo cancellation," *Workshop on application of signal processing to audio and acoustic*, New Paltz, NY, October 1997, pp. 19–22 (unpublished).

Experience with a yes–no single-interval maximum-likelihood procedure

Marjorie R. Leek^{a)}

Army Audiology and Speech Center, Walter Reed Army Medical Center, Washington, D.C. 20307-5001

Judy R. Dubno,^{b)} Ning-ji He, and Jayne B. Ahlstrom

*Department of Otolaryngology-Head and Neck Surgery, Medical University of South Carolina,
P.O. Box 250150, Charleston, South Carolina 29425*

(Received 7 July 1997; revised 1 November 1999; accepted 4 February 2000)

The report in 1993 by Green [J. Acoust. Soc. Am. **93**, 2096–2105 (1993)] describing the application of a new psychophysical method requiring few trials and little time to measure auditory thresholds has generated considerable interest among experimentalists. The procedure uses a single-interval stimulus presentation, requests a yes–no decision by subjects, and implements a maximum-likelihood calculation to determine the next trial stimulus level within an adaptive track, as well as the final threshold estimate. Data are presented here describing separate experiences with this procedure in two laboratories in both detection and discrimination tasks. Issues addressed include comparisons with more traditional psychophysical methods, variability in threshold estimates, experimental time required, and possible minor modifications to improve the basic procedure. Results using this procedure are comparable in terms of variability of estimates to those emerging from more lengthy procedures. However, because it may be difficult for some listeners to maintain a consistent criterion and because attentional lapses may be costly, experimenters must be willing to monitor performance closely and repeat some tracks in cases where excessively high variability is noted. Further, this procedure may not be suitable for tasks for which the form of the psychometric function is not well-established. Modifications allowing a variable slope parameter in the maximum-likelihood evaluations of psychometric functions may be of benefit. © 2000 Acoustical Society of America. [S0001-4966(00)03305-1]

PACS numbers: 43.66.Yw, 43.66.Cb, 43.66.Dc, 43.66.Fe [JWH]

INTRODUCTION

Efforts to improve psychophysical procedures are often directed toward reducing the variability of the estimates and, hence, increasing the procedure's efficiency. Adaptive methods are known for their measurement efficiency compared with other methods, estimating points on a psychometric function by consulting performance on previous trials to select stimulus levels for subsequent trials. A new adaptive procedure incorporating a yes–no single interval task, introduced by Green (1993), promises to produce reliable thresholds in a limited amount of experimental time. Rather than an up–down adaptive algorithm, popularized by Levitt (1971), trial-by-trial stimulus selection and final threshold estimates are based on a maximum-likelihood evaluation of an array of candidate psychometric functions, selecting the one best representing past performance within the current adaptive track. In Green's yes–no procedure, listeners were asked a rather simple question on each trial presentation, i.e., did they or did they not hear the signal, reducing possible subject confusions regarding the probabilistic nature of data collection in forced-choice paradigms. The subject's task was simply to indicate "yes" (i.e., "I heard the tone") or "no" ("I did not hear the tone"). Subjects were free to set their own criterion for saying yes or no. There was no feed-

back (other than an indication of response registration) because the judgment was a subjective one, incapable of external verification (with the possible exception of responses to mute or "catch" trials). Convergence to a most-likely threshold value was generally very quick, reducing subject fatigue and increasing experimenter satisfaction because more data could be collected.

In a series of studies, Green and his colleagues (Green, 1993, 1995; Gu and Green, 1994) used a combination of computer simulations and human data collection to become acquainted with the characteristics and limitations of the procedure. It became apparent that the specification of a false-alarm rate, defined as the rate of responding yes to very low-intensity stimuli, was a significant problem, in that the estimation of a false-alarm rate was biased low. This problem was addressed fairly successfully by Gu and Green (1994) by including catch trials within the track, during which no stimulus was presented. Simulations indicated the importance of several experimental parameters controlling the adaptive track. The spacing of candidate psychometric functions along the stimulus axis (essentially a step-size variable) was not critical to the determination of threshold, nor was the initial starting value of the adaptive track. The assumed slope of the psychometric function was important in the precision of a threshold measurement, but Green suggested that even if the assumed psychometric function slope was quite different from the "true" slope, the threshold value would not be excessively compromised. The length of

^{a)}Electronic mail: marjorie.leek@na.amedd.army.mil

^{b)}Electronic mail: dubnojr@musc.edu

the track was the strongest determinant of precision of estimate. Through computer simulations, Green found that, after some minimum number of trials, the expected relationship between precision of measurement and number of trials held, i.e., the standard deviation of the threshold estimates, was inversely proportional to the square root of the number of trials. Using laboratory staff members as listeners, Green (1993) determined that threshold estimates with about a 3-dB standard deviation could be collected in as few as 12 trials in a track. Finally, comparison was made between results from the yes–no maximum-likelihood procedure and those from a more traditional two-alternative forced-choice (2AFC) up–down procedure, with findings indicating significant agreement between the two. However, variability of the estimates was larger in the yes–no procedure than in the 2AFC procedure.

In this communication, experiences with the yes–no single-interval maximum-likelihood procedure will be described, involving several psychoacoustic experimental tasks, subjects with normal hearing and those with hearing loss, and subjects of varying ages. Data collection was carried out in two independent laboratories in several experiments. Tasks involved measures of detection threshold for a tone in quiet and a tone in noise with threshold searches involving either a fixed signal or a fixed noise. Also included are discrimination tasks, measuring both frequency and intensity difference limens. Each of these tasks presents a different set of experimental problems and procedural decisions, such as the choice of psychometric function, and decisions concerning false-alarm rates and catch trials. Modifications to Green's original formulation of the basic procedure have been implemented to address specific difficulties encountered by each laboratory for particular tasks. Most of the subjects who have taken part in these experiments were inexperienced in psychometric experimentation and the subjects with hearing loss may have an additional disadvantage due to the auditory distortions caused by their damaged systems. Reported here is success with the procedure, some suggestions and cautions in its implementation and use with human subjects, and some recommendations for further development.

I. GENERAL METHODOLOGY

The general procedure, as described by Green (1993) and Gu and Green (1994), is designed to find the psychometric function among a set of possible candidate functions that most probably underlies the behavior reflected in a data set. It is an adaptive procedure in that the stimulus level on each trial is selected from the current "best" psychometric function as determined from the previous data collected. The procedure requires several assumptions. One assumption is the shape of the psychometric function [in Green's work this has been either a logistic function or a Gaussian-type function called a "stimulus power function" (see Gu and Green, 1994)]. A set of candidate functions of the assumed shape extends above and below the estimated threshold region. Each function is described with a location along the stimulus axis, a slope or spread factor (the estimated change in performance with a given change in stimulus level), and an

assumed false-alarm rate. Throughout the threshold search to follow, each of the candidate functions is evaluated after each trial, and its probability upgraded based on performance on all previous trials, to determine which most probably underlies the subject's behavior. In practice, the number of candidate functions narrows rapidly after a few trials to converge on the best possible function. Once a given precision of estimate or a given number of trials has been reached, the threshold search ends and a threshold value is extracted from the best psychometric function.

II. APPLICATION OF THE PROCEDURE

The primary advantage of using the single-interval maximum-likelihood procedure is its promise of significant time savings in data collection. Potential costs of the procedure, however, may be a reduction in measurement precision and increase in experimental bias relative to more traditional psychometric procedures. A further concern is that the more subjective yes–no procedure may be problematic for inexperienced or clinical subjects. In the following sections, questions concerning use of this procedure with clinical or aged populations, the degree of time savings, and the potential penalties associated with this adaptive procedure will be addressed. Section II A compares threshold measurements and within-subject variability in a typical 2AFC task with measures taken in the yes–no maximum-likelihood procedure. An analysis of time savings using this adaptive procedure is presented in Sec. II B, along with a comparison of performance on the procedure by hearing-impaired and normal-hearing experimental subjects, and some suggestions for modifications meant to reduce test time required and variability of results. Finally, Sec. III B compares the use of the procedure to measure auditory discrimination with a more traditional method of constant stimuli. In such experiments, the characteristics of the psychometric function underlying performance may not be well-established. In this section, the ability of elderly subjects to produce reliable measures is also evaluated.

A. How does the procedure compare with a forced-choice procedure?

1. Method

The work reported by Green and his colleagues (Green, 1993; Gu and Green, 1994) focused on detection of tones. This was also the task used in the Dubno laboratory in a study of simultaneous and forward masking by normal-hearing (NH) and hearing-impaired (HI) subjects (Dubno and Ahlstrom, 1999). Absolute and masked-detection thresholds were measured using both the 2AFC procedure and the maximum-likelihood yes–no procedure with NH subjects. Both procedures were also used to measure absolute thresholds in HI subjects.

The 2AFC procedure was an up–down adaptive paradigm with feedback converging on 70.7%-correct performance (Levitt, 1971). The signal level was initially set at a level at least 30 dB above the estimated threshold and then adjusted in 6-dB steps. After six track reversals (i.e., changes in direction of signal level), the step size was reduced to 2 dB, and, after two more direction changes, threshold was

determined by the subsequent ten track reversals. Threshold was computed as the mean levels of the track peaks (reduced by one half step) minus the mean of the track valleys (increased by one half step). Prior to the start of each adaptive track, subjects heard repeated samples of the signal (and noise, when appropriate) presented at signal levels that were clearly audible.

The yes–no maximum-likelihood procedure was essentially the same as described by Green (1993). The procedure started with a set of candidate psychometric functions, each described by a logistic function with variations in its middle point (m , in dB) and/or false-alarm rate (α).

$$P(\text{yes}) = \alpha + \frac{(1 - \alpha)}{1 + e^{-k(x - m)}} \quad (1)$$

Following Green (1993), the slope parameter, k , was assumed to be 0.5 (k is dimensionless). The number of candidate functions was determined by the range of the signal levels (usually 40 to 50 dB) in 0.5 dB steps multiplied by nine α levels ranging from 0 to 0.4. The precise range and starting level varied with frequency and subject. During each trial only one stimulus was presented. Subjects responded by clicking one of the two mouse buttons corresponding to the responses “yes, I heard the tone” and “no, I did not hear the tone.” There was no feedback. Each signal presented was used to calculate the probability, $P(\text{yes})$, for each of the candidate functions when the subject’s response was yes. When the subject’s response was no, the probability was $1 - P(\text{yes})$. The probability for each candidate function was upgraded by cumulatively multiplying it with the probability obtained in the previous trials. Then, a searching procedure was used to find the psychometric function with the maximum probability. This psychometric function was taken as the most likely function describing all previous performance. Its “sweetpoint,” defined as the point on the function leading to the minimum variability of threshold estimates, was selected as the stimulus level for the next trial [see Green (1993) for a discussion of sweetpoints]. For this study, the sweetpoint was calculated after each trial using Eq. (5) of Green (1993), instead of the simplified 60% point suggested by Green. When $\alpha = 0$, the sweetpoint is equal to 50%. Each adaptive track was terminated with a fixed number of trials set at 24, including 4 catch trials during which no signal was presented. The catch trials occurred randomly within a track and followed the same calculation procedure as other trials. Prior to the start of each adaptive track, subjects heard repeated samples of the signal (and noise, if appropriate) presented at or above the starting level of the subsequent track.

For the NH subjects, simultaneous-and forward-masked thresholds were measured using 2AFC and maximum-likelihood methods. All thresholds were obtained using one psychophysical method, followed by threshold obtained using the other method. Order of detection measure and psychophysical method were counterbalanced across subjects.

Both HI and NH groups of subjects participated in measures of absolute thresholds. Two sets of thresholds were obtained with each psychophysical method in order to evaluate within-subject variability. The time to obtain each abso-

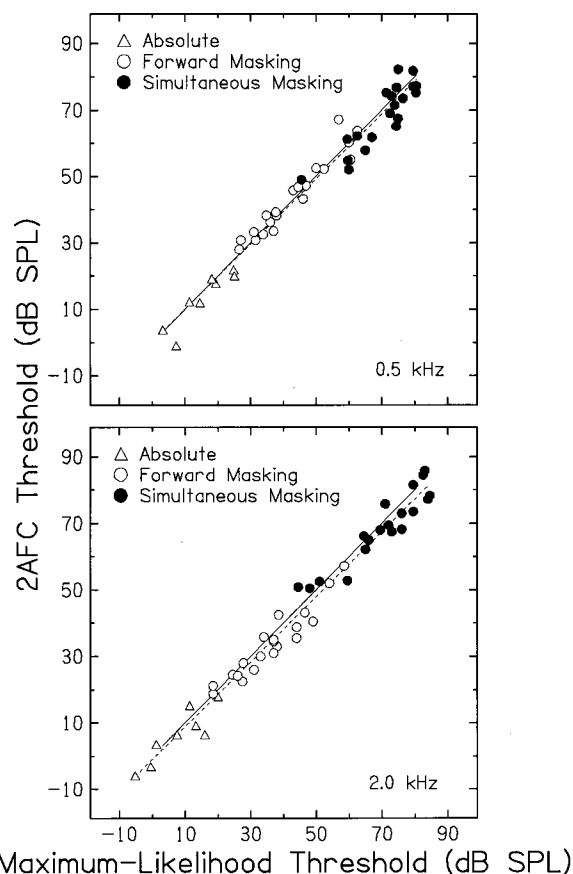


FIG. 1. Scatter plots of 2AFC and yes–no maximum-likelihood thresholds at 0.5 and 2.0 kHz for normal-hearing subjects. The solid line in each panel has a slope of 1 and the dashed line is the best-fit linear regression function fit to all thresholds.

lute threshold was recorded, as was the total number of trials for the 2AFC adaptive procedure.

Subjects across both groups ranged in age from 18 to 69 years. Signals were pure tones at several frequencies, with duration of either 350 or 20 ms. Maskers were broadband noises, some with spectral notches. (The specific characteristics of the signals and maskers are not relevant to the focus of this article.)

2. Results

Data from these detection measures may be used to compare thresholds and between-and within-subject variability for the two psychophysical procedures.

Figure 1 shows scatter plots of 2AFC and maximum-likelihood thresholds at 0.5 and 2.0 kHz for NH subjects, including absolute, forward-masked, and simultaneous-masked thresholds. The solid line in each panel has a slope of 1 and the dashed line is the linear regression fit to all thresholds using a least-squares procedure. Including all thresholds, Pearson correlation coefficients (r) between thresholds for the two procedures were 0.98 (0.5 kHz, $p < 0.01$) and 0.99 (2.0 kHz, $p < 0.01$). When the three detection measures were examined separately, correlation coefficients remained high, ranging from 0.96 and 0.95 (for 0.5 and 2.0 kHz, respectively, $p < 0.01$) for forward masking to 0.92 and 0.89 ($p < 0.01$) for absolute thresholds. Note that

when detection thresholds are examined separately, correlations are affected by range and sample size limitations. For both 0.5 and 2.0 kHz (across all detection tasks), the slopes of the regression lines (0.98 and 0.98) were not significantly different from 1 ($p > 0.05$) and the y intercepts (0.13 and -1.05) were not significantly different from zero ($p > 0.05$). These results suggest that the overall agreement between the two psychometric procedures was excellent.

Across all conditions, yes–no thresholds were slightly higher than 2AFC thresholds. This difference was expected because 2AFC measurements converged on the 70.7% point on a psychometric function that ranged from 50% to 100%, whereas yes–no measures converged on the 50% point on a psychometric function that ranged from 0% to 100% (for $\alpha = 0$). Using the rescaling formula suggested in Eq. (1) of McKee *et al.* (1985), a 2AFC measurement converging on 70.7% on a psychometric function ranging from 50% to 100% is equivalent to 41.4% on a psychometric function ranging from 0% to 100%. Here, maximum-likelihood thresholds converged on the 50% point on the yes–no psychometric function. Assuming a slope factor of 0.5 and $\alpha = 0$, the level difference corresponding to a change from 50% to 41.4% is 0.7 dB. Therefore, at a minimum, 0.7 dB of the difference between yes–no maximum-likelihood and 2AFC thresholds can be accounted for by the difference in criterion performance between the two procedures. Of course, as the slope of the psychometric function decreases (i.e., becomes less steep), the level difference corresponding to a change from 50% to 41.4% increases. In addition, as the false-alarm rate increases from 0, the sweetpoint will increase from 50%, which also increases the level difference from 41.4%.

Thresholds for the three detection tasks were examined separately to determine if the agreement between the two psychophysical procedures was consistent for forward-masking, simultaneous-masking, and absolute thresholds. Across masker conditions, forward-masked thresholds measured using the maximum-likelihood procedure were 0.79 dB higher than those using the 2AFC procedure. This value is nearly equal to the predicted threshold difference of 0.7 dB owing to criterion performance differences for the two procedures. An analysis of variance (ANOVA) with repeated measures on psychophysical procedure and signal frequency supported a conclusion of no difference in forward-masked threshold due to procedure ($p = 0.49$). A significant interaction of procedure by frequency ($p = 0.04$) was attributed to a larger threshold difference due to procedure for 2.0 kHz than for 0.5 kHz.

Simultaneous-masked thresholds measured using the maximum-likelihood procedure were, on average, 1.82 dB higher than those using the 2AFC procedure. The overall difference was about 1.1 dB greater than predicted for these two procedures. This difference was statistically significant ($p = 0.009$) even when a 0.7-dB correction was applied ($p = 0.03$), whereas the procedure by frequency interaction was nonsignificant ($p = 0.82$). The larger influence of procedure observed for simultaneous than for forward masking may reflect greater uncertainty, inattentiveness, or task difficulties for simultaneous than forward masking, or it may indicate

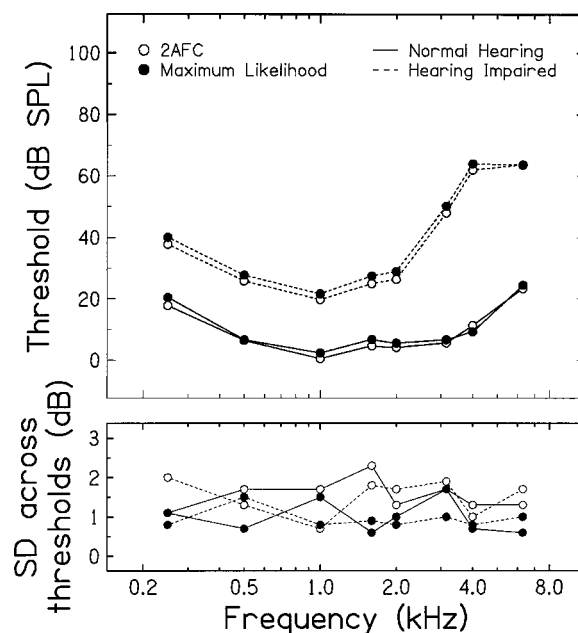


FIG. 2. Mean absolute thresholds as a function of frequency for the two subject groups measured using 2AFC and yes–no maximum-likelihood methods (top panel). Mean between-subject standard deviations for normal-hearing and hearing-impaired subjects for two methods (bottom panel).

that the slope of the psychometric function for simultaneous masking is somewhat shallower than the assumed slope of 0.5.

Comparisons of within-subject variability for the two procedures may be assessed by examining results of absolute detection in which two threshold measures for each procedure were obtained for five NH and five HI subjects. The top panel of Fig. 2 shows mean absolute thresholds as a function of frequency for the two subject groups measured using 2AFC and maximum-likelihood methods. The values shown are means of the two threshold measures. Across frequency and subject group, maximum-likelihood thresholds were higher than 2AFC thresholds by an average of 1.40 dB. This difference was statistically significant ($p = 0.008$) but became nonsignificant when the 0.7-dB correction was applied ($p = 0.12$). Interactions were nonsignificant ($p > 0.23$). Thus, thresholds measured with the 2AFC procedure and the maximum-likelihood procedure were consistent for two groups of subjects, a wide range of frequencies, and across two threshold measurements.

Means of the standard deviations across two threshold measurements are shown in the lower panel of Fig. 2 for the two subject groups and two procedures. Across signal frequency, the mean within-subject standard deviations for NH listeners were 1.56 and 0.97 dB for the 2AFC and maximum-likelihood procedures, respectively, and 1.51 and 0.94 dB for HI listeners. An ANOVA with hearing status as a grouping factor and repeated measures on frequency and procedure revealed a significant main effect of procedure ($p = 0.004$); no other significant main effects or interactions were found ($p > 0.46$). Thus, for both NH and HI subjects, less within-subject variability was observed for the maximum-likelihood procedure than for 2AFC. The within-subject variability observed for the maximum-likelihood procedure is remarkable

given that 2AFC thresholds were obtained with an average of 53 trials (with two listening intervals per trial), whereas the maximum-likelihood thresholds were obtained with only 24 trials (one interval per trial). This corresponded to an average of 2:13 and 2:43 (min:s) per 2AFC threshold for NH and HI subjects, respectively, and 0:45 and 0:56 (min:s) per maximum-likelihood threshold. These results are further evidence of the efficiency of the maximum-likelihood procedure, given that 2AFC-equivalent thresholds were measured in less time with less variability.

These data indicate that in measures of detection, whether in quiet or in the presence of a masking noise, the maximum-likelihood procedure shows comparable thresholds and reliability to measures made using the more traditional forced-choice procedure. Further, the more subjective yes–no procedure was not adversely affected when used with a clinical population compared to measures obtained from a normal-control subject group.

B. How fast is the procedure and what levels of variability are expected?

1. Method

Tone detection in noise was also the task used by Leek and Summers (1996) in a study of notched-noise auditory filters at a variety of signal levels, but in this experiment the level of the signal was held constant while noise level needed to mask the tone was allowed to vary. However, several modifications to the original yes–no procedure were implemented in order to ensure a specified level of variability within each adaptive track, to incorporate a stopping rule by tying the length of the adaptive track to a measure of variability, to include a mechanism of familiarization periodically throughout the experimental sessions, and to allow an estimation of repeatability of threshold estimates.

For each threshold, the yes–no maximum-likelihood procedure measured the level of noise required to mask a 1.0- or 2.0-kHz tone at levels ranging from 40 to 100 dB SPL. The noise was either broadband or bandstopped, with a spectral notch in the frequency region of the signal.

Sets of three adaptive tracks were collected for each measurement and the mean threshold and standard deviation were calculated from them. If the standard deviation across the three measures was larger than 3 dB, that set was discarded and a further set of three tracks was collected. At the beginning of each set of three tracks, the noise level was set very low, so that the tone was clearly audible and the response was “yes, I heard the tone.” On each subsequent trial during this familiarization phase, the noise level was increased by 5 dB until the first no response occurred. The next trial initiated the maximum-likelihood procedure, beginning with a noise level that was randomly selected to be from 5 to 15 dB below the noise level of the first no response. The next two tracks in the three-track set were begun without additional familiarization but with an initial level again determined randomly to be 5–15 dB below the last familiarization trial. This procedure served to ensure that the first trial in a track was clearly a yes trial and to remind the listener of the sound of the tone to be detected. Both the reminder of the signal and the automatic range adjustment for each listener

and condition proved to be especially valuable for HI listeners. The reminder was of considerable value in cases where there was significant distortion caused by the hearing loss, where the listener tended to “lose” the signal. The range adjustment provided assurance that the levels of the candidate psychometric functions were appropriate for different individuals with widely varying dynamic ranges of hearing.

The length of each track was determined by a running calculation of the 95%-confidence interval across levels visited after the first four trials in the track. The track was required to be at least 12 trials, but was terminated thereafter when the confidence interval of previous levels dropped below 1 dB. If that level of track consistency had not been reached by 50 trials, the search was stopped and a threshold was estimated from the last function selected. This occurred on 2% of tracks for NH listeners and on 3% and 5% of tracks for HI listeners for signal frequencies of 1.0 and 2.0 kHz, respectively. This suggests that the distortion associated with the greater hearing loss at 2.0 kHz may have made a consistent criterion difficult to maintain by those listeners.

The final modification made to the original tracking procedure was selection of the shape of the psychometric function. Gu and Green (1994) suggested that the “stimulus power function” provided a better representation of psychometric performance on the tone-in-noise detection task. However, because the track adapted on the level of the noise, rather than the level of the tone, the direction of the function was reversed. The underlying psychometric function assumed in this procedure was

$$P(\text{yes}) = 1 - \Phi[k[m + c - 10 \times \log_{10}(10^{m/10} + 10^{s/10})]], \quad (2)$$

where Φ is the cumulative Gaussian function, m is the mid-point of the function, c represents the listener’s criterion, and s is the increase in power (in dB) when the signal is added to the noise. This function follows Eq. (3) in Gu and Green (1994). The slope parameter, k , was assumed to be -0.30 , and c was chosen to produce false-alarm rates of 0.0, 0.1, 0.2, or 0.3.

At the initiation of each track, a set of candidate functions was determined to span a 40-dB range of thresholds and the four possible false-alarm rates. The range (and the subsequent number of candidate functions) was reduced, if necessary, to avoid overall noise levels greater than 105 dB SPL. On each trial, there was a 20% probability of a mute signal (catch trial) to improve the estimation of false-alarm rate. After each trial, the currently “best” psychometric function was consulted and the next trial was presented with a noise level that would produce 70% yes responses on that function. At the end of the track, this value (for the current best function) was selected as the threshold estimate.

2. Results

The nature of this design and selection of experimental conditions produces information regarding stability of tracks, variability within subjects across three threshold estimates, and the amount of time necessary to collect enough data to estimate the means and variability of threshold. These findings are available for subjects with normal hearing and those

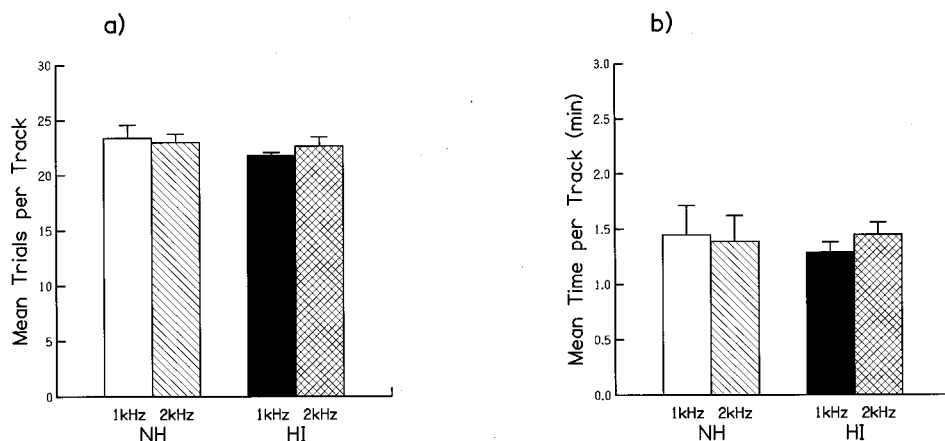


FIG. 3. Mean number of trials per track (a), and time required per track (b) for two groups of subjects at two signal frequencies in a yes-no maximum-likelihood procedure measuring tone detection in noise. Error bars represent ± 1 standard deviation of the means.

with moderate sensorineural hearing loss. Results will be expressed in terms of signal frequency (1.0 and 2.0 kHz) and subjects' hearing status. Frequency is probably not a determining factor for this procedure, but this group of HI listeners had much greater losses at 2.0 kHz than at 1.0 kHz. Most of the measures of time or variability reflect better (more consistent, faster) performance at 1.0 kHz for the HI subjects than at 2.0 kHz, suggesting more signal distortion, less adequate noise rejection in favor of signal, and in general greater uncertainty and difficulty of listening. Because no comparable difference as a function of frequency was observed for NH subjects, this loss of consistency is likely not due to frequency or subjects *per se*, but rather to the degree of hearing loss.

a. Track length and time required for threshold estimation. Recall that the track length was not fixed in this experiment, but was allowed to vary in order to reach a criterion level of stability. Data were obtained with several noise configurations and levels as maskers, and each listener produced at least three threshold estimates for each masker. There were, then, 50 to 90 thresholds measured for each listener. Figure 3(a) shows the mean number of trials per track for the four NH and the five HI subjects for each signal frequency. Error bars represent the standard errors of the means across subjects. The mean number of trials across frequency for the NH listeners was 23.2, whereas for the HI listeners it was 22.3. An unbalanced ANOVA with hearing status as a grouping factor and repeated measures on frequency supports a conclusion of no difference in length of tracks to reach criterion stability due to group membership or signal frequency ($p > 0.10$ for all main effects and interactions).

Trials were self-paced and so the amount of time taken to reach a threshold is a function of both the number of trials in the track and the pace at which subjects chose to work. The time involved in presentation of a trial is, of course, much reduced from forced-choice methods because only one presentation is required for each response rather than two or more. Figure 3(b) shows the time required per track to meet the criterion stability, again separated by group and signal frequency. The mean time for each track, combining across frequency, was 1.4 min for both groups of subjects. Measures of mean threshold and variability in threshold estimate

across three tracks (as was taken in this study) could be obtained in just over 4 min on average, regardless of hearing status of the subjects. For comparison, three threshold estimates using a 2AFC procedure in a 50-trial adaptive track might require perhaps 15 to 20 min. An unbalanced ANOVA indicated no significant differences in time for thresholds due to the main effects of group membership or signal frequency ($p > 0.20$), but the interaction of group by frequency was significant ($p = 0.02$).

Although differences in experimental time were not different for NH and HI subjects, the group-by-frequency interaction suggests that the pace of the trials may have differed for the two frequencies. Figure 4 shows a scatter plot of the mean time per track for each subject as a function of their audiometric threshold at the signal frequency. As indicated by the ANOVA and in the previous figure, there was considerable overlap in time required per track for members of the

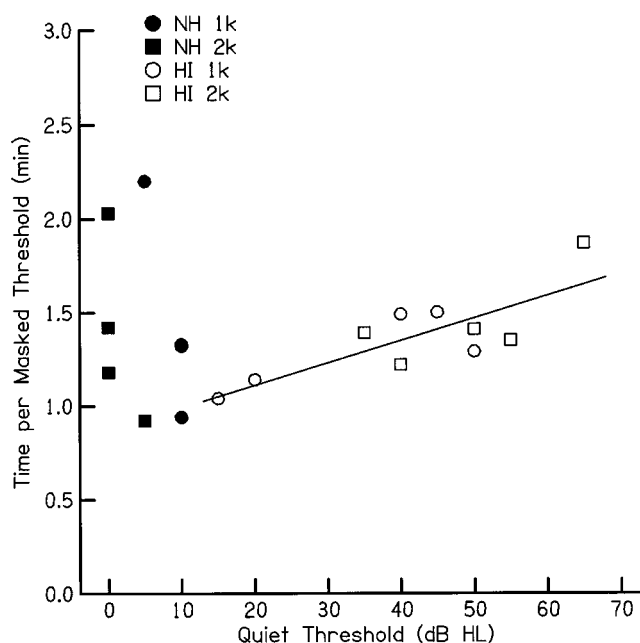


FIG. 4. Relationship between time required for threshold estimation in the yes-no procedure and absolute threshold. The solid line is the best-fit regression line representing the effect of threshold on experimental time for hearing-impaired listeners.

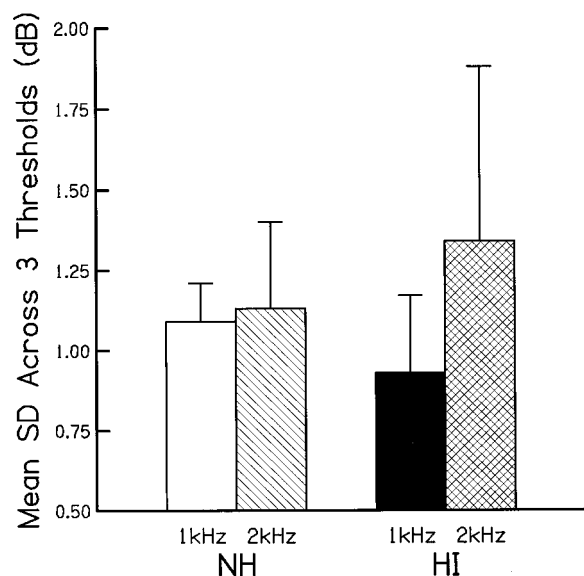


FIG. 5. Mean standard deviations across three threshold estimates of tone detection in noise using the yes–no maximum-likelihood procedure for two groups of subjects and two signal frequencies. Error bars are standard deviations across subjects.

two groups. However, the time per track was significantly correlated with the amount of hearing loss within the HI listeners ($r=0.80$, $p=0.005$). The greater losses at 2.0 kHz compared to 1.0 kHz no doubt underlie the significant interaction of group by frequency, and suggest that the pace of response was in part determined by the difficulty inherent in a more or less distorted signal.

b. Within-subject variability across repeated threshold estimates. For each threshold estimated in this study, the means and standard deviations over three tracks were calculated. If the standard deviation of these three measures exceeded 3 dB, an additional set of three thresholds was obtained, and the first set was excluded. This occurred on 15% of the threshold measures (sets of three) for the NH listeners and 12% of the measures for HI listeners. The means of the standard deviations are shown in Fig. 5 for the two groups and two signal frequencies. Across signal frequency, the mean standard deviation for NH listeners was 1.11 dB and for HI listeners 1.14 dB. These group standard deviations were not significantly different ($p>0.8$), nor was the interaction between group and frequency ($p=0.10$). However, there was a significant main effect of signal frequency ($p<0.05$), conditioned mainly by an increased variability for 2.0 kHz by all the HI listeners over that shown for 1.0 kHz. Two of the NH listeners also had slightly greater variability at 2.0 kHz, but most of the significant frequency effect was probably related to the larger hearing losses at 2.0 kHz. Nevertheless, none of the listeners demonstrated a mean standard deviation across all their measures greater than 2 dB. The additional time taken to achieve these low levels of variability by rerunning sets of threshold tracks when indicated did not appreciably decrease the time savings due to this experimental procedure. Inflating the time per threshold reported earlier by 15% suggests that the time to obtain three threshold measures with a standard deviation of less than 3 dB would be 4.8 min—still considerably less time than would be

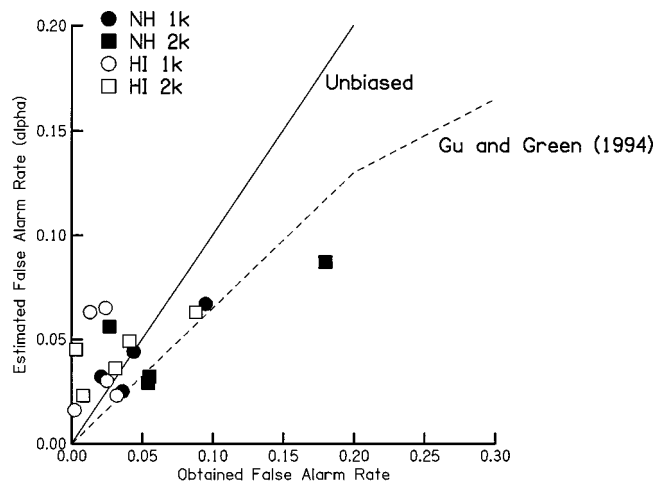


FIG. 6. Comparison of estimated false-alarm rate α from the maximum-likelihood procedure and the obtained false-alarm rate (yes responses to catch trials/number of catch trials). Data points are averages across 27 threshold estimates for each subject and signal frequency. The solid line shows the expectation from an unbiased procedure, while the dashed line is taken from Gu and Green (1994).

required for a more traditional forced-choice procedure.

c. What about the problem of estimating false alarms?

One of the most intransigent problems identified by Green (Green, 1993; Gu and Green, 1994) is obtaining an acceptable measure of the false-alarm rate. Gu and Green (1994) suggested including catch trials on a percentage of the presentations, with the stimulus level set very low. In the Leek and Summers (1996) study, catch trials were provided by muting the signal presentation so that only noise was present with 20% probability on each trial. The ratio of yes responses to a catch trial over the number of catch trials was used to define the false-alarm rate.

Gu and Green (1994) used simulations to determine how accurately the procedure could estimate false-alarm rates when the actual rate varied from 0 to 0.30, and compared actual and estimated rates for several catch-trial probabilities. The data reported by Gu and Green for a 20% probability of a catch trial, as used in the procedure described here, are reproduced in Fig. 6 as the dashed line. The solid line shows the prediction if the actual false alarms occur at the same rate as the estimated rate. The data points in Fig. 6 show the mean estimated false-alarm rates for the NH and HI listeners from the best-fitting psychometric functions on the ordinate and the obtained false-alarm rates, calculated from yes responses to catch trials, on the abscissa. Means of the individual subjects for the two signal frequencies are shown.

Figure 6 indicates that the estimate of false-alarm rate, given a 20% probability of a catch trial, tends to be more nearly an unbiased estimate for this group of subjects than Gu and Green found for their simulated subjects, in that the values tend to cluster more around the solid line than the dashed one. However, there is a tendency for the HI subjects (open symbols) to produce slightly higher false-alarm estimates in the maximum-likelihood procedure than may be observed in their actual responses to catch trials. This bias is relatively small, however, when compared to the severe un-

derestimation of false-alarm rates reported by Green (1993) for the original version of the procedure.

C. How does the procedure perform in measures of discrimination?

Most of the work reported to date using the single-interval maximum-likelihood procedure has involved detection tasks. It is not obvious how the special considerations involved in measuring differential thresholds might interact with the procedure, and whether the benefits observed in detection tasks are as great for this more difficult task.

One notable difference between detection and discrimination tasks is that the shapes and slopes of the underlying psychometric functions change with varying stimulus differences. He *et al.* (1998) used this procedure to measure frequency discrimination (Δf) and intensity discrimination (ΔL) for tones. Because the appropriate psychometric functions were not apparent from earlier studies, a method of constant stimuli (MCS) was used to describe the functions before implementation of the maximum-likelihood procedure. For this experiment, then, data are available that show how the maximum-likelihood procedure compares with a more traditional procedure in a discrimination task.

1. Method

Psychometric functions for both frequency and intensity discrimination were measured using a method of constant stimuli and a same/different task with a single interval on each trial. Characteristics of the psychometric functions determined in the MCS were then used to implement the maximum-likelihood procedure. For both measurement procedures, there was one listening interval during which two tones were presented sequentially. The first tone was always the standard stimulus and the second was either the same as the standard or higher (in frequency or level) than the standard. On each trial, subjects indicated with a button-press whether they heard two “different” tones or two “same” tones. This is one form of a same/different task called a “reminder” design by Macmillan and Creelman (1991), and was found to be easier for naive or aged subjects than a more traditional discrimination paradigm.

Intensity and frequency discrimination were measured in separate blocks of trials. In the MCS procedure, comparison stimuli differed from the standard stimuli on a given trial by 0 Hz or 0 dB, up to as much as 1.25% times the standard frequency or 10 dB (for Δf or ΔL measurements, respectively). A psychometric function was constructed for each of four standard frequencies and two intensity levels based on 50 trials each of 11 comparison levels or frequencies, presented in random order, for a total of 550 trials per psychometric function.

For each subject and standard frequency, the obtained psychometric functions were fitted by the logistic function shown in Eq. (1) with m and x in Hz or dB, in order to estimate slopes associated with each discrimination and standard frequency. Mean slope parameters across subjects and frequencies were $k=3.0$ for frequency discrimination, and

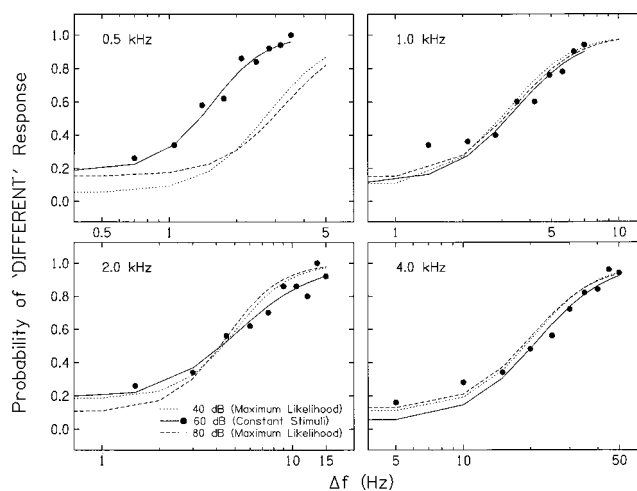


FIG. 7. Psychometric functions for frequency discrimination at four standard frequencies and three standard levels obtained from one subject. The filled circles are responses obtained with a constant-stimulus method using a 60-dB standard level and the solid lines are logistic functions fit to those data. The functions for 40- and 80-dB standard levels (dotted and dashed lines, respectively) were estimated from adaptive maximum-likelihood results. These data were previously shown in He *et al.* (1998), Fig. 3.

$k=1.4$ for intensity discrimination. These values were subsequently used in the implementation of the maximum-likelihood procedure.

In the maximum-likelihood procedure, frequency and intensity discrimination were measured at four standard frequencies at standard levels of 40 and 80 dB SPL. The procedure was essentially the same as described above for detection thresholds, with $P(\text{yes})$ replaced by $P(\text{different})$. The minimal variable magnitude was 0 (i.e., no difference) and the maximum varied with task and subject. The set of candidate psychometric functions was determined by the range and variable magnitudes incorporating four α values: 0, 0.1, 0.2, and 0.3. The number of trials was fixed at 50, including three catch trials during which the minimal variable magnitude was used. Thresholds corresponding to $d'=1$ were taken from the best psychometric function after each maximum-likelihood block was completed.

Prior to each session's data collection with the maximum-likelihood procedure, subjects heard several familiarization trials over the complete range of stimulus differences. These practice trials also served to determine appropriate variable ranges for the maximum-likelihood procedure. Differential thresholds using the maximum-likelihood procedure were measured four times for each standard frequency. The interval between two consecutive threshold estimates ranged from 15 min to several weeks. Four young subjects (≤ 33 years) and four aged subjects (≥ 65 years), all with normal hearing, participated in each experiment, for a total of 13 subjects (i.e., not all subjects participated in both experiments).

2. Results

a. How does the procedure compare with a constant-stimuli procedure? Data from this experiment may be used to make several comparisons between the adaptive maximum-likelihood procedure and the more time-consuming MCS

procedure. Comparisons may be made in terms of the resultant psychometric functions generated by both procedures and in terms of the differential estimates of thresholds. Figure 7 shows psychometric functions for frequency discrimination at four standard frequencies at three standard levels obtained from one subject. The panels include the best psychometric function averaged across four replications of the maximum-likelihood procedure for 40- and 80-dB standard levels, using estimated values of m and α and the fixed value of k ($k = 3$ in this case). The functions fitted to the experimental data obtained with the constant-stimuli procedure at a standard level of 60 dB (solid lines and filled symbols) were computed from the estimated values of m , k , and α ; Δf in Hz was transformed into natural logarithmic units [$\ln(\Delta f)$] for curve fitting. The logistic function provides a generally good fit to the Δf data.

Frequency discrimination was not expected to vary substantially between 40 and 80 dB and the maximum-likelihood psychometric functions reflect only small differences at the two levels. In Fig. 7, for this subject, comparisons between fitted functions from the constant-stimuli procedure and functions generated in the maximum-likelihood procedure were generally very similar, except at 0.5 kHz. At that standard frequency, there is a sizable difference in the threshold estimated from the constant-stimuli and maximum-likelihood procedures as indicated by the horizontal distance between the 60-dB function and 40- and 80-dB functions. A review of Δf thresholds for all subjects and frequencies, however, revealed no systematic difference in either mean thresholds or variability due to psychophysical method, although somewhat higher maximum-likelihood thresholds were observed at 0.5 kHz without an increase in variability. Differences between thresholds for the two methods averaged about 0.02%, which was well within the range of intersubject variability. Comparable results were observed for intensity-discrimination functions. This generally good agreement between functions generated by the two procedures is remarkable, given that each psychometric function from the constant-stimuli procedure included 550 trials, whereas each function generated by the maximum-likelihood procedure included only 50 trials. Further, although a fixed, 50-trial stopping rule was incorporated, *post hoc* examination of the adaptive tracks revealed that the procedure converged on the best possible function well before 50 trials.

b. Within-subject variability across repeated threshold estimates. Each differential threshold was measured four times (tracks) over several sessions. Within-subject variability is of interest for these measures due to the difficult nature of the discrimination task and the inclusion of young and aged subjects who were inexperienced with both the task and the adaptive nature of the procedure. For frequency and intensity discrimination, differential thresholds obtained for each of the four tracks were submitted to separate ANOVAs with subject age as a grouping factor and repeated measures on standard level, standard frequency, and track. For frequency discrimination, the main effect of track was not significant ($p = 0.06$) and there were no significant interactions with track ($p > 0.12$). For intensity discrimination, a significant track main effect was found ($p = 0.01$) and was attrib-

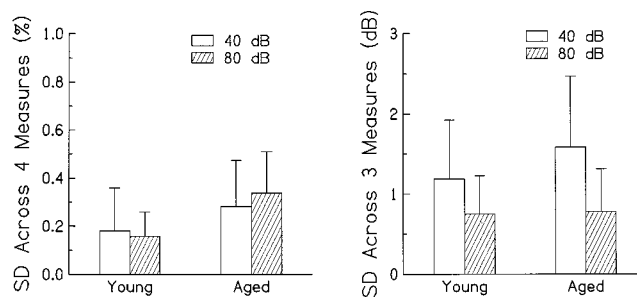


FIG. 8. Mean standard deviations across four measures of frequency discrimination (left panel) and three measures of intensity discrimination (right panel) for two standard levels obtained from young and aged subjects with normal hearing. Results shown are pooled over four standard frequencies.

uted to unusually large ΔL s measured in the first track for one standard level for aged subjects only. Because this outcome was not likely to be related to the maximum-likelihood procedure, subsequent analyses of within-subject variability for intensity discrimination included only the last three tracks.

For frequency and intensity discrimination, standard deviations across four or three tracks were submitted to separate ANOVAs with subject age as a grouping factor and repeated measures on standard level and standard frequency. The means of the standard deviations are shown in Fig. 8 for the two subject groups and two standard levels. Within-subject variability did not vary with standard frequency and there were no significant interactions with frequency; as such, results shown are pooled over standard frequency. For frequency discrimination (left panel), across standard level, the mean standard deviation was 0.16% for young subjects and 0.31% for aged subjects. Although within-subject variability appears to be slightly larger for aged than young subjects, these group differences were not significantly different ($p = 0.07$), nor were any other main effects or interactions ($p > 0.2$). For intensity discrimination (right panel), across standard level, the mean standard deviation was 0.97 dB for young subjects and 1.18 dB for aged subjects. These group differences were not significant ($p = 0.4$), nor was the interaction between group and standard level ($p = 0.24$). However, there was a significant main effect of standard level ($p = 0.005$) attributed to increased within-subject variability for the 40-dB standard for both groups. The increased variance for the lower-level standard may reflect greater uncertainty, or inability to maintain a consistent criterion for this condition. Nevertheless, the discrimination results reveal very small standard deviations across measures, given the difficulty of the task, the inexperience of the subjects, and the length of time between some measures.

III. GENERAL DISCUSSION

For all of the experimental tasks and subject groups described in these experiments, the single-interval yes-no maximum-likelihood procedure gave results that were nearly equivalent to results obtained in more time-consuming, but more bias-free, procedures such as 2AFC. In tone-detection tasks, the correlation of data from the two procedures was nearly perfect, ranging from 0.95 to 0.99 on tasks of simul-

taneous and forward masking. In the discrimination measures, a “reminder” same/different task was used, similar to the yes–no task in that a single response is given. Correspondence between discrimination measures from a method of constant stimuli and from an adaptive maximum-likelihood procedure was generally good, but inconsistencies were observed between the two procedures in the generated psychometric functions for a few conditions, notably at a standard frequency of 0.5 kHz. Such discrepancies may reflect the difference between adaptive and nonadaptive procedures, rather than being specific to the maximum-likelihood adaptive procedure *per se*.

In the experiments described here, both a logistic psychometric function and a Gaussian-type function were used. Although Gu and Green (1994) developed the stimulus power function [Eq. (2)] because they saw evidence that the logistic did not provide a sufficiently good fit to the data, the differences between the two were found to be small. The logistic function [Eq. (1)] is certainly more computationally accessible and there is little evidence from the data sets evaluated here to prefer one over the other. Therefore, in future work, it may be of little additional value to use the computationally less tractable Gaussian function.

Of more importance in controlling variability are issues of slope of the candidate functions and the selection of a sweetpoint. Green (1993) suggested that even if the slope of the assumed function were considerably different from the true underlying function in a listener, the procedure would still converge on a reasonable and valid threshold estimate. There are hints in the data reported here, however, that a poor selection of slope might increase the variability, requiring either longer tracks to converge or more replications of the measurement. This was seen in the slightly poorer agreement of the yes–no and 2AFC tasks for simultaneous masking than for forward masking, indicating that perhaps the slope of the psychometric function might differ for these two experimental tasks. The psychometric parameter slope has not been adequately addressed, either in previous reports of simulation or human data. It may be valuable to include several estimates of slope in the array of candidate psychometric functions, so that slope is allowed to vary.

The selection of sweetpoint differed between the two laboratories reporting data in this article. Green (1993) described the sweetpoint as the point on the currently best psychometric function that minimizes the measurement variability in the estimate of threshold. The sweetpoint may be calculated after each trial in the yes–no procedure, and the next stimulus level presented is based on that calculation. By always selecting presentation levels based on the sweetpoint of the currently best candidate psychometric function, the actual percent-correct level of the presentation is allowed to vary from trial to trial. A variable sweetpoint was used in the detection data measured in Sec. II A. In contrast, a fixed level of proportion of yes responses, based loosely on the approximate value of the sweetpoint, may be used, as was suggested by Green (1993), and implemented in the tone-in-noise detection data shown in Sec. II B. A rough comparison of within-subject variability from data taken with a variable sweetpoint (Sec. II A) and with a fixed presentation level of

70% correct may be made by consulting Figs. 2 and 5. Because these variability measures are standard deviations across either two or three threshold estimates, the standard errors may be compared. In Fig. 2, the mean standard error across frequencies and listeners is about 0.68 dB, whereas the mean standard error for the variability shown in Fig. 5 is about 0.64 dB. As Green suggested, there is little, if any, penalty associated with using an estimate of sweetpoint in this procedure, rather than tying stimulus presentation level to each sweetpoint as determined by the current best candidate function. However, a systematic approach to the selection of stimulus level based on the sweetpoint might take on greater importance in cases where the psychometric function shape and slope are not yet established rather than in a relatively simple detection task. As there is little computational effort involved in calculating presentation level based on a sweetpoint determined by the current psychometric function, in less well-understood experimental tasks, it may be of benefit to allow the stimulus presentation level to be linked to the “true” sweetpoint after each trial.

The yes–no maximum-likelihood procedure imposes some burdens on the experimenter to monitor results carefully. Gu and Green (1994) and Green (1995) have evaluated the effects of lapses in subject attention during this procedure, showing that, depending on when during the adaptive track such a lapse occurs, there is greater or lesser impact. If a lapse occurs early in the track (within the first few trials), it is difficult for the track to “recover” within the specified number of trials remaining (Gu and Green used 15 trials in these simulations), and convergence on an appropriate psychometric function may be delayed significantly, resulting in inaccurate estimates of thresholds. Lapses occurring later in the adaptive track had less impact on the final threshold values. Because a lapse may occur due to a momentary loss of attention or even a simple erroneous response, it is necessary to monitor results carefully and be willing to rerun tracks in cases of excessively high variability. Imposition of a stopping rule for the track tied to track variability, rather than using a fixed number of trials, can assist both in keeping the tracks as short as is consistent with an acceptable amount of track variability, as well as alerting the experimenter to lapses in attention that adversely affect the final threshold estimate. Interestingly, the number of trials in tracks terminated by a stopping rule was, on average, about the same number of trials selected for the fixed-track-length data shown in Sec. II A. If a fixed track length is desired, 24 trials appears to be well supported by three separate estimates of variability from three laboratories: one reported by Green (1993) and Gu and Green (1994), one from the tracks from Dubno’s lab reported in Sec. II A here, and the mean track length resulting from a variability-linked stopping rule implemented in Leek’s lab reported in Sec. II B. However, a stopping rule may add to efficiency of measurement by allowing consistently and systematically converging adaptive tracks to stop early, while permitting more variable tracks to extend long enough to reach an acceptable convergence.

Performance by different groups of subjects provides additional understanding of this procedure. Hearing-impaired subjects tended to have slightly different patterns of results

in terms of variability and time for thresholds that appeared to be related to the degree of hearing loss. Some indications that this task may have been slightly more difficult for listeners with hearing loss than those with normal hearing may be observed in the relationship between time per track and threshold sensitivity (Fig. 4), and greater within-subject variability for tones in a region of greater hearing loss (Fig. 5). These two findings may reflect more difficulty maintaining a criterion for saying yes when the hearing loss presumably affects the clarity of the tone. However, these results are rather subtle, in that the mean track lengths across groups of normal-hearing and hearing-impaired listeners were not significantly different, nor was the mean time required for each track. Procedures used with these subjects included an initial familiarization sequence at the beginning of each track or three-track set, which may have served to remind the listener of the signal on a regular basis. There was not a significant difference on frequency discrimination measures using the yes–no maximum-likelihood procedure between young subjects and more elderly listeners, suggesting that age is not a factor in the ability to maintain a threshold criterion.

Agreement between thresholds obtained with the yes–no maximum-likelihood procedure and results from a forced-choice procedure was excellent. Time to obtain those thresholds, however, was greatly reduced for the former method relative to the latter. The results reported here, along with studies carried out through simulation and human data collection in Green's laboratory and in a clinical setting (Formby *et al.*, 1996), provide separate verifications of this procedure's success in providing fast, consistent, and valid thresholds in tone-detection tasks. Use of this method in other psychoacoustic tasks, such as gap-detection thresholds (Florentine *et al.*, 2000), and the discrimination tasks reported here, is also supported. A caveat must be noted however, in that this procedure may be problematic in tasks that have underlying psychometric functions that are highly variable across conditions or subjects, or that are not well understood, and should be carefully and watchfully implemented.

ACKNOWLEDGMENTS

This work was supported by research Grant Numbers R01 DC00626, R01 DC00184, and P50 DC00422 from the National Institute on Deafness and Other Communication Disorders, National Institutes of Health. All subjects participating in this research provided written informed consent prior to beginning the study. We are grateful to Dr. Craig Formby and Dr. Joseph L. Hall for valuable comments on an earlier version of this paper. The opinions or assertions contained herein are the private views of the authors and are not to be construed as official or as reflecting the views of the Department of the Army or the Department of Defense.

- Dubno, J. R., and Ahlstrom, J. B. (2000). "Forward- and simultaneous-masked thresholds in bandlimited maskers in subjects with normal hearing and cochlear hearing loss," *J. Acoust. Soc. Am.* (submitted).
- Florentine, M., Buus, S., and Geng, W. (2000). "Toward a clinical procedure for narrowband gap detection I: A psychophysical procedure," *Audiology* (in press).
- Formby, C., Sherlock, L. P., and Green, D. M. (1996). "Evaluation of a maximum-likelihood procedure for measuring pure-tone thresholds under computer control," *J. Am. Acad. Audiol.* **7**, 125–129.
- Green, D. M. (1993). "A maximum-likelihood method for estimating thresholds in a yes–no task," *J. Acoust. Soc. Am.* **93**, 2096–2105.
- Green, D. M. (1995). "Maximum-likelihood procedures and the inattentive observer," *J. Acoust. Soc. Am.* **97**, 3749–3760.
- Gu, X., and Green, D. M. (1994). "Further studies of a maximum-likelihood yes–no procedure," *J. Acoust. Soc. Am.* **96**, 93–101.
- He, N.-j., Dubno, J. R., and Mills, J. H. (1998). "Frequency and intensity discrimination measured in a maximum-likelihood procedure from young and aged normal-hearing subjects," *J. Acoust. Soc. Am.* **103**, 553–565.
- Leek, M. R., and Summers, V. (1996). "Effects of signal level on auditory filter shapes in hearing-impaired listeners," Abstracts of the 19th Midwinter Research Meeting, Association for Research in Otolaryngology, p. 139.
- Levitt, H. (1971). "Transformed up–down methods in psychoacoustics," *J. Acoust. Soc. Am.* **49**, 467–477.
- Macmillan, N. A., and Creelman, C. D. (1991). *Detection Theory: A User's Guide* (Cambridge University Press, Cambridge).
- McKee, S. P., Klein, S. A., and Teller, D. Y. (1985). "Statistical properties of forced-choice psychometric functions: Implications for probit analysis," *Percept. Psychophys.* **37**, 286–298.

Phonological conditioning of peak alignment in rising pitch accents in Dutch

D. R. Ladd, Ineke Mennen,^{a)} and Astrid Schepman^{b)}

University of Edinburgh, Department of Theoretical and Applied Linguistics, Adam Ferguson Building,
George Square, Edinburgh EH8 9LL, United Kingdom

(Received 26 July 1999; accepted for publication 7 January 2000)

This paper deals with the factors that influence the alignment of *F0* movements with phonetic segments. It reports two experiments on the alignment of rising prenuclear pitch accents in Dutch. In experiment 1, it is shown that the final peak of the rise is aligned at the end of the vowel if the accented syllable contains a long vowel, but during the following consonant if the accented syllable contains a short vowel. The beginning of the rise is consistently aligned at the beginning of the accented syllable. Experiment 2 attempts to distinguish between two explanations for this finding: (1) a durational account, in which the *F0* rise takes a certain amount of time and overruns into the following consonant if the vowel is short; and (2) a structural account, in which the peak of the rise is seen as a tonal target aligned with the end of the syllable (which is structurally earlier for long vowels than for short vowels). The data partially support both accounts. There is an alignment difference despite a lack of durational difference, which supports the structure-based account. However, the effect is reduced compared to experiment 1, showing that time pressure may work against the ideal alignment. © 2000 Acoustical Society of America. [S0001-4966(00)03004-6]

PACS numbers: 43.70.Fq, 43.70.Bk [AL]

INTRODUCTION

A. Segmental anchoring of tonal targets

A number of recent papers have demonstrated the existence of striking regularities in the way fundamental frequency (*F0*) movements are aligned with the segmental string (Silverman and Pierrehumbert, 1990; Caspers and van Heuven, 1993; Caspers, 1994; Prieto *et al.*, 1995; Arvaniti *et al.*, 1998; Xu, 1998). Caspers and van Heuven (1993) studied the effects of various kinds of “time pressure” on the realization of linguistically significant *F0* movements (“pitch accents”) in Dutch. Their motivation was to determine which aspect of pitch accents would be most stable under time pressure—the shape of the accent, the size of the pitch excursion, or the alignment of the accent with the segmental string. They created time pressure in three different ways: by asking speakers to vary their speaking rate, by asking speakers to put the same pitch movements on CVC words containing short vowels as well as long vowels, and by asking speakers to produce two major pitch movements in quick succession. While their findings were rather complex, one clear finding was that the onset of a rising pitch accent (i.e., the *F0* minimum at the beginning of the rise) is consistently aligned at the beginning of the accented syllable. The *F0* peak at the offset of the pitch accent, by contrast, is aligned in quite variable ways.

In a similar vein, Prieto *et al.* (1995) found that the *F0* minimum at the beginning of a rising pitch accent in Mexi-

can Spanish is consistently aligned with the onset of the accented syllable. Like Caspers and van Heuven, they found that the *F0* maximum at the end of the rise is affected by a considerable number of factors, whose effects they model with a multiple regression formula. A similar approach was taken by Silverman and Pierrehumbert (1990). More recently, comparable results have been found by Xu (1998) for the *F0* movements that manifest lexical tone in Chinese, suggesting that the principles of *F0* alignment are broadly similar in all languages regardless of the linguistic function of pitch. However, Xu’s findings were interestingly different from those on European languages. Specifically, he found that the *end* of the *F0* rise for lexical tone 2 (mid-high rising) in Mandarin is consistently aligned at the end of the syllable to which it applies. The *beginning* of the rise, on the other hand, seems less consistently aligned, being affected by speaking rate, by the segmental content of the syllable and by tonal crowding.

All of these findings are compatible with a model in which the accent is some sort of unit gesture, anchored at only one particular point in the segmental string, and lasting a specified amount of time or moving through a specified *F0* interval or trajectory. This approach has been adopted in the past for use in speech synthesis (e.g., Fujisaki, 1983; ‘t Hart *et al.*, 1990). However, a new set of findings in this general line of alignment research casts doubt on this view. Arvaniti *et al.* (1998) showed that in Modern Greek, when the location of nearby word boundaries and other accents is carefully controlled, *both the beginning and the end* of rising accents are anchored to specific points in the segmental string. Arvaniti *et al.*’s finding was corroborated in a study by Ladd *et al.* (1999), in which it was shown that the beginning and end of rising accents in English are anchored to specific points in the segmental string regardless of changes in seg-

^{a)}Present address: Department of Speech, University of Newcastle, King George VI Building, Queen Victoria Road, Newcastle-upon-Tyne NE1 7RU, United Kingdom.

^{b)}Present address: Division of Psychology, School of Social & Health Sciences, University of Abertay Dundee, Marketgait House, 158 Marketgait, Dundee DD1 1NJ, United Kingdom.

mental duration brought about by modifications of speech rate. Arvaniti *et al.* argue that their finding provides evidence for the theory assumed in much modern phonological work on intonation and tone (e.g., Bruce, 1977; Pierrehumbert, 1980; Ladd, 1996), namely that *F0* movements should be analyzed structurally as *sequences of tonal targets* occurring at specified points in the utterance. This view need not contradict the “unit gesture” theory entirely: an analogy might be to the affricates /tʃ/ and /dʒ/ in English, which are units at one level of analysis but consist of two distinct subunits which determine the articulatory trajectories at another level of analysis. It does suggest, however, that the notion of segmentally anchored local tonal targets will be of value in modeling *F0*.

The present paper provides further evidence for the independent specification of the beginning and end of rising pitch accents, and hence of the tonal target view. Specifically, it demonstrates that the alignment of the peak of rising pitch accents in Dutch is actually conditioned by the phonological vowel length of the accented syllable. It also provides evidence for the effect of “time pressure” of the sort studied by Caspers and van Heuven (see first paragraph of this section), but argues that in order to describe the effects of time pressure accurately we must make explicit reference to both the beginning and the end of the *F0* movements.

B. Phonological vowel length and accent peak alignment

The variable manipulated in these experiments is phonological vowel length. In Dutch, as in English and German, there are two series of vowels, generally known either as long and short or as tense and lax, respectively. The reason for the multiple terminology is that the phonetic basis of the distinction is a matter of some disagreement, and may vary from language to language or dialect to dialect; crucially, as we shall see, it is not always phonetic duration. The structural or phonological reality of the distinction, however, is not in doubt. In all three languages there is a basic restriction on the form of monosyllabic words, which is that they *cannot end in a short vowel*. To take an illustration from English, there are long–short pairs ending in a consonant, such as *beat-bit* or *bait-bet*; but in monosyllabic words without a following consonant only the long vowels can occur. Thus *bee* and *bay* are real words, but /bɪ/ and /bɛ/ are, in principle, impossible. An identical restriction on short vowels in monosyllabic words applies in Dutch and German, though of course the exact vowel inventories are not the same. Other reflexes of the vowel length distinction are: (1) syllable structure: speakers readily syllabify *beaten* as *bea-ten* but with *bitten* hesitate among *bi-tten*, *bitt-en*, and *bit-ten* [cf. Treiman and Danis (1988) for English; Schiller *et al.* (1997) for Dutch]; and (2) possible morpheme shapes: for example, morpheme-internal consonant sequences are permissible following a short vowel, but not after a long vowel; thus monomorphemic *guilder* and bimorphemic *builder* are both possible, but *fielder* can be only bimorphemic.

Our initial observations of the alignment of accent peaks in both Dutch (Mennen, 1999) and English have suggested that it is affected by phonological vowel length. Specifically,

it appears that accent peaks are aligned earlier in syllables with phonologically long vowels than with phonologically short ones—during the vowel in the case of long vowels, and during the following consonant in the case of short ones. Such a difference can also be inferred from experimental findings for Dutch by Caspers and van Heuven (1993). Although they do not report peak alignment as such (but see Caspers, 1994, Appendix D.2: Tables Ib and IIb), they do report that rise time was about 18 ms longer in syllables with long /a:/ than in syllables with short /a/. Recall that they found that the onset of the rise aligned with the onset of the syllable in both conditions. The difference in vowel duration for the pair /a:/ vs /a/ is typically on the order of 70 ms (Nooteboom and Slis, 1972). This means that, *relative to the vowel offset*, the end of the rise must have been considerably *earlier* with long vowels than with short vowels. Our first goal in this paper (reported in experiment 1) is to put this finding on a sound empirical footing.

One natural explanation for such a finding, assuming it is confirmed, would be in terms of a minimum duration for the rise in *F0*. Suppose that there is a fixed anchor point for the *F0* rise at the beginning of the accented syllable, and that the rise takes a certain length of time to complete such that it peaks late in a long vowel. Given this rise duration, then by the time the *F0* rise is completed, the speaker has already finished the short vowel and is into the following consonant. We will refer to this line of reasoning in what follows as the *durational explanation*. However, there is another possible explanation, which we will term the *structural explanation*. Suppose that the end of the rise, like the beginning, is anchored to a specific place in structure, as found by Arvaniti *et al.* (1988) and by Ladd *et al.* (1999). Then the difference in alignment might be explained by assuming that the peak of the rise is aiming for a fixed anchor point at the “end of the syllable” in both cases. (Recall that Xu found the end of the syllable to be the fixed alignment point for tone 2 rises in Mandarin.) Since, as we just saw, the syllable structures associated with long and short vowels are probably different—the syllable includes some or all of the consonant in the case of the short vowel, but not in the case of the long vowel—we would expect the peak to be aligned near the end of a long vowel but during the consonant following a short vowel. The second goal of this paper (in experiment 2) is to distinguish experimentally between these two possible explanations.

I. EXPERIMENT 1

A. Method

The general method involved measuring the alignment of *F0* events with segmental events in prenuclear rising accents.¹ These were elicited from speakers who read prepared sentences of the following sort (see also Fig. 1):

- (1) Ik kan de melige grappen van Seth Gaaikema niet meer aanhoren. (I can no longer stand Seth Gaaikema’s corny jokes.)
- (2) Wij konden de rennende atleten met geen mogelijkheid bijhouden. (There was no way we could keep up with the running athletes.)

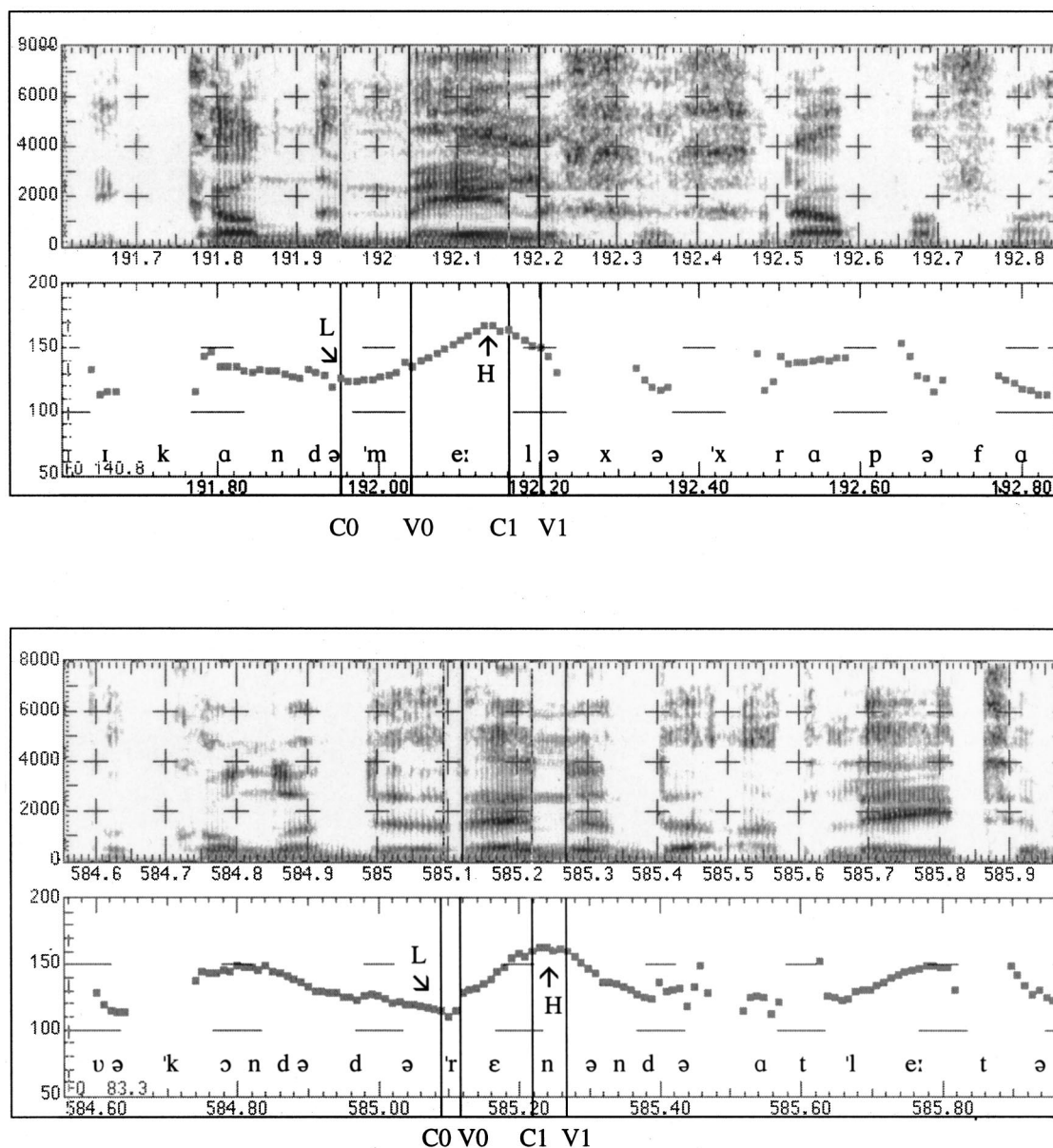


FIG. 1. Measurement points for both experiments, illustrated with example sentences from experiment 1. The top panel shows a sentence including a test word with a phonologically long vowel. The bottom panel has a test word with a phonologically short vowel. Note the earlier peak alignment, relative to C1 (the end of the vowel), in the test word with the phonologically long vowel.

In sentence (1), *melige* /'me:ləxə/ is the test word, and *me-* is the test syllable, in this case containing a phonologically long vowel; in sentence (2), *rennende* /'rɛnɛndə/ is the test word, and *re-* is the test syllable, containing a phonologically short vowel. (We do not prejudge the results of the experiment by referring to *ren-* as the test syllable at this stage. As noted in the Introduction, however, there is strong reason to assume that, in the case of short vowels, the test syllable would include at least some of the following consonant.) For readers not familiar with Dutch, it may be helpful to mention that long vowels are generally followed by a single consonant letter in the orthography, while short vowels are followed by two.

In both types of test syllable, we identified the location of the F_0 minimum (L) at the beginning of the accentual F_0 rise and the location of the F_0 maximum or peak (H) at the

end of the rise. We expected that, irrespective of vowel length, L would align with the onset of the test syllable, as found by Caspers and van Heuven (1993) for Dutch and as observed in several studies of other languages as well (e.g., Mexican Spanish—Prieto *et al.*, 1995; Greek—Arvaniti *et al.*, 1998; and English—Ladd *et al.*, 1999). Our central hypothesis, as discussed in the Introduction, was that H would be aligned earlier, relative to the vowel offset, when the vowel is phonologically long than when it is phonologically short.

1. Speech materials

There were 20 sentences with one of the short vowels /ɪ/, /ɛ/, /a/, /ɔ/, and /ʏ/ and 20 sentences with one of the long vowels /i:/, /e:/, /a:/, /o:/, and /y:/.

sentences are given in the previous subsection. Each test word was preceded by between two and five unstressed syllables and the test word always had lexical stress on the antepenultimate syllable, to avoid stress clash or tonal crowding (cf. Prieto *et al.*, 1995; Arvaniti *et al.*, 1998). The test syllable's vowel nucleus was flanked by sonorants or, in a few cases, voiced obstruents. In all but two cases, the test word was an adjective followed by a noun; in the two remaining cases it was a noun followed by a semantically closely linked prepositional phrase [e.g., *beloningen van de politie* (police rewards), item 31]. In all cases the expected stress pattern was a prenuclear accent on the test word, followed by either another prenuclear accent or a nuclear ('pointed hat') accent on the following noun. In five sentences the test vowel was followed by a consonant cluster. The other 35 items had a singleton consonant. The 40 test sentences were randomly interspersed among 80 filler items to prevent speakers from detecting the regularities of the test set (for instance, the regular syntactic structure of subject-verb-adjective-noun beginnings used in most test sentences). A full list of test items can be found in the Appendix.

2. Speakers

The sentences were read by five adult native speakers of Dutch. Three speakers were female and two were male. Speakers were not paid for their participation.

3. Recording and analysis procedures

Speakers read the sentences twice from a set of cards, each of which had one test sentence typed on it. The order of the sentences was random. Speakers were asked to read the sentences as naturally as possible, and were asked to repeat any misread sentences. Materials were recorded on Digital Audio Tape (DAT) in studios at the Universities of Amsterdam or Groningen, or in a quiet room at the speaker's home.

Recorded materials were digitized at a sampling rate of 16 kHz, with appropriate low-pass prefiltering. The second author selected the first acceptable token for further measurement. A token was considered unacceptable if it contained a disfluency, when the recording was noisy, when the test word was immediately followed by a phrase boundary, or when otherwise unexpected intonation patterns were produced (e.g., question intonation or deaccenting of the test word). In 84% of all cases, repetition 1 was chosen. Repetition 2 was used in 14.5% of cases, and in 1.5% of cases data were missing.

The selected sentences were analyzed using a Sun SPARC workstation running ESPS Waves+. F_0 tracks were obtained using 49-ms \cos^4 windows moving in 10-ms steps. Durational measurements were made by marking selected points in a simultaneous display of the waveform, wide-band spectrogram, and F_0 track. The marking was done by hand by the second author.

The two tonal targets were marked as follows. The location of the highest F_0 around the end of the accented syllable of the test word was marked as H , and L was defined as the F_0 minimum in the accented syllable of the test word.

As consonants can cause perturbations on the F_0 track (for instance, nasal and laterals can cause slight dips and rises in the F_0 contour at their onsets and offsets, respectively), we compensated for any confounding effects of this phenomenon by ignoring F_0 points that were clearly perturbations, and taking the next highest or lowest point as our measurement point. However, no attempt was made to compensate for microprosodic effects on F_0 when there was doubt whether the F_0 was a true (intonational) minimum or maximum or a false (microprosodic) F_0 point. In those cases, the absolute F_0 maxima and minima in the vicinity of the test syllables were chosen. If two points of equal F_0 value existed, we consistently chose the one that occurred first. On the whole, there were clear valleys and peaks, which enabled us to measure according to the objective criteria described above.

The four segmental points were: the onset of the accented test syllable (C_0); the onset of the accented vowel (V_0); the end of the accented vowel (C_1); the onset of the following vowel (V_1). On the basis of these segmental landmarks, the following dependent variables were derived:

- (1) Alignment of L (L minus C_0 in ms; negative values indicate alignment of L before C_0);
- (2) Alignment of H (H minus C_1 in ms; negative values indicate alignment of H before C_1);
- (3) Duration of the accented vowel (C_1 minus V_0 in ms).

An example of the measurement points can be seen in Fig. 1.

Boundaries between vowels and nasals or laterals were marked at the point where sudden changes in both amplitude and formant structure occurred. The onset of /r/ was marked by a combination of cues, e.g., the occurrence of noise, a change in formant structure, or an increase in local energy at the boundary between the vowel and liquid. If the change in formant structure was gradual, the segment boundaries were drawn at the midpoint of the transition from vowel to liquid. For obstruents, the start of closure was marked as the onset, and the start of high amplitude periodicity was marked as the onset of the next vowel.

B. Results

All data were analyzed in overall two-way (2×5) mixed design ANOVAs, with items as the random factor, Phonological Vowel Length as a between-items factor, and Speaker as a within-items factor. In addition, all measures were analyzed by individual speaker, using one-way ANOVAs, with items as the random factor and Phonological Vowel Length as a single between-items fixed factor. Individual speakers' means and analyzes for all measures are reported in Table I. Graphic presentation of the results is postponed until the report of experiment 2, in order to facilitate the comparison between the two sets of data (see Fig. 2).

1. Alignment of L

The F_0 minimum is aligned close to the onset of the syllable in both conditions. Mean alignment is 2.6 ms after the consonant onset in long vowel syllables and 1.2 ms before the consonant onset in short vowel syllables. The two-way ANOVA showed no significant effect of Phonological

TABLE I. Means (standard errors), and individual analyses for experiment 1. Relevant measures are indicated in the table. The subjects' initials are prefixed with F (Female) or M (Male).

Speaker	Long	Short	<i>df</i>	<i>F</i>	<i>P</i>
C0 to <i>L</i> (alignment of the <i>L</i> relative to C0 in ms)					
FCF	1.2 (5)	-6.9 (5)	1,38	0.11	0.743
FLH	-3.0 (5)	10.2 (5)	1,37	2.71	0.108
FMS	6.5 (8)	10.5 (8)	1,36	1.71	0.199
MDG	11.1 (7)	-3.7 (6)	1,38	2.19	0.147
MEH	-2.8 (5)	-15.9 (5)	1,38	1.53	0.223
Overall	2.6 (3)	-1.2 (2)			
C1 to <i>H</i> (alignment of <i>H</i> relative to C1 in ms). Negative values mean that <i>H</i> is aligned before the end of the vowel and positive values mean that it is aligned after the end of the vowel, i.e., in the following consonant.					
FCF	-11.3 (7)	29.6 (7)	1,38	18.1	<0.0001
FLH	-21.2 (6)	25.8 (5)	1,37	42.0	<0.0001
FMS	-9.3 (7)	22.1 (7)	1,36	8.2	0.007
MDG	-16.8 (4)	18.8 (4)	1,38	41.2	<0.0001
MEF	-1.3 (7)	27.5 (7)	1,38	9.5	0.004
Overall	-12.0 (3)	24.8 (2)			
V0 to C1 (duration of the vowel in ms)					
FCF	126.6 (6)	70.5 (3)	1,38	73.1	<0.0001
FLH	142.8 (8)	73.8 (4)	1,37	59.6	<0.0001
FMS	143.1 (7)	85.1 (2)	1,36	63.3	<0.0001
MDG	126.7 (7)	83.0 (3)	1,38	33.5	<0.0001
MEF	126.3 (8)	72.5 (2)	1,38	36.9	<0.0001
Overall	133.1 (3)	77.0 (1)			
<i>H</i> (Hz) minus <i>L</i> (Hz) (<i>F0</i> change from <i>L</i> to <i>H</i>)					
FCF	81.7 (3)	84.1 (4)	1,38	2.0	0.657
FLH	98.6 (5)	105.5 (6)	1,37	0.7	0.402
FMS	90.5 (4)	94.3 (4)	1,36	0.4	0.512
MDG	52.2 (3)	47.6 (2)	1,38	1.4	0.244
MEF	39.7 (2)	34.3 (2)	1,38	4.7	0.037
Overall	72.5 (3)	73.1 (3)			

Vowel Length ($F < 1$). However, there was a significant effect of speaker, $F(4,140) = 3.61$, $p = 0.008$, and a significant interaction, $F(4,140) = 2.71$, $p = 0.033$.

The overall outcome clearly replicates the findings of Caspers and van Heuven that the $F0$ minimum is consistently aligned close to the onset consonant of the accented syllable, regardless of vowel length. The analysis also reveals some speaker idiosyncrasies, leading to both the main effect of speaker (different overall alignment) and the interaction (differences in magnitude and direction of alignment differences in the two vowel length conditions). However, as the differences in alignment for the two types of vowel were extremely small, we believe it is appropriate to conclude that phonological vowel length has no systematic influence on the alignment of the beginning of the $F0$ rise.

2. Alignment of *H*

The $F0$ maximum was aligned 12 ms before the offset of the vowel when the vowel was long and 24.8 ms into the next consonant when the vowel was short. The overall analysis revealed a significant main effect of Phonological Vowel Length, $F(1,35) = 42.139$, $p < 0.0001$, but no effect of speaker and no interaction (F 's < 1.5). All individual analyses were also significant for the factor Phonological Vowel Length. This replicates Caspers' (1994) results [see also

Caspers and van Heuven (1993) from which similar results can be inferred, as described in the Introduction]. The effect can be clearly seen in the two example sentences in Fig. 1.

3. Vowel duration, rise duration, and onset consonant duration

The mean duration of the accented vowel was 133.1 ms for long vowels, and 77 ms for the short vowels. The overall analysis revealed a significant main effect of Phonological Vowel Length, $F(1,35) = 56.2$, $p < 0.0001$, which reflected the individual analyses. There was also a significant effect of speaker, $F(4,140) = 7.41$, $p < 0.0001$, and an interaction, $F(4,140) = 3.38$, $p = 0.011$. Once again, these two effects can be attributed to speaker idiosyncrasies and are not of concern. The main finding was that vowel duration differed by 56.1 ms in the two conditions, and that this difference was statistically significant.

We report a further measure to facilitate comparison with the findings of Caspers and van Heuven (1993). The distance from the syllable onset to the peak (*H* minus *C0*) was 193 ms in short vowels and 206 ms in long vowels. Although this difference was not significant, $F(1,35) = 1.67$, $p = 0.205$, the pattern is comparable to that reported by Caspers and van Heuven (1993), who found an 18-ms difference in the same direction. The fact that our effect was slightly smaller and nonsignificant can be attributed to the fact that we used a larger variety of vowels, which adds variability to the data. A final relevant measure is the duration of the onset consonant of the test syllable. This was, on average, 91 ms in syllables with short vowels, and 86 ms in syllables with long vowels. The difference between these was not significant [$F(1,35) < 1$], showing that the duration of the onset consonant(s) did not form a confounding factor in our data.

4. $F0$ change

As we have just seen, rises accompanying phonologically long vowels tend to be longer than those accompanying short vowels. As a result, it might be expected that rises accompanying long vowels would show larger $F0$ excursions than those accompanying short vowels. To check whether this was the case, we calculated the difference in Hz between the $F0$ at *L* and that at *H*. The $F0$ difference was 72.5 Hz in long vowels and 73.1 Hz in short vowels.

In the overall ANOVA there was no effect of Phonological Vowel Length on the $F0$ change measure ($F < 1$). There was a significant effect of Speaker, $F(4,140) = 101.8$, $p < 0.0001$, due to the wider $F0$ range in female than in male speakers. The interaction was not significant ($F < 1$). In the individual analyses only one speaker showed a significant difference (5.4 Hz greater difference in long vowels than in short vowels), but other speakers showed nonsignificant effects in both directions.

Overall, then, there was no significant link between phonological vowel length and amount of $F0$ change. There is no reason to assume that the $F0$ movements associated with long and short vowels differ consistently in the size of the interval they span.

C. Discussion

The results consistently replicate Caspers and van Heuven's findings and confirm our initial observations. The *L* of the Dutch prenuclear rising accent, like similar accents in other European languages, aligns close to the onset of the syllable with both long and short vowels. The *H* of this accent aligns in the general vicinity of the end of the accented vowel, which is considerably earlier than the alignment of the peak of the corresponding accent in Greek (Arvaniti *et al.*, 1998; cf. Mennen, 1998, 1999), but similar to findings for Mexican Spanish (Prieto *et al.*, 1995) and English (Ladd *et al.*, 1999).

In addition, however, as shown by Caspers' (1994) findings, there is a strong effect of phonological vowel length: the *H* aligns shortly before the offset of the long vowel, and some distance after the offset of the short vowel. As discussed in the Introduction, one can imagine two quite different explanations for this finding. The durational explanation is that the rise takes a certain amount of time, and the short vowel is too short for the rise to be fully realized on the vowel itself. The structural explanation, by contrast, assumes that tonal targets are anchored to specific places in structure, and explains the difference in alignment in terms of the structural differences between long and short vowels in Dutch: The accent peak may be anchored to the end of the syllable, which is equivalent to the end of the vowel if the vowel is long, but which is during the following consonant if the vowel is short.

The goal of experiment 2 is to decide between these two explanations.

II. EXPERIMENT 2

As it happens, Dutch provides us with an ideal test-bed for distinguishing between the durational and structural effects of phonological vowel length on the alignment of accent peaks. It has long been noted in phonetic descriptions of the language that the phonologically long high vowels /i:/, /y:/, and /u:/, are phonetically rather short (Nootboom and Slis, 1972). The logic of experiment 2 exploits this fact by applying rising prenuclear pitch accents to segmentally controlled materials containing "long" and "short" high vowels in the accented test syllables. If the long and short vowels are the same duration (as found by Nootboom and Slis) and if the durational explanation of the alignment difference in experiment 1 is correct, then there should be no difference in alignment in these cases: For both the short and long vowels the *F0* peak should be aligned in the following consonant. If, on the other hand, the explanation in terms of syllable structure is correct, then there should be a difference in alignment even though the vowels have the same phonetic duration: The *F0* peak should still be aligned at the offset of the phonologically long vowel and well after the offset of the phonologically short vowel.

A. Method

The general approach was the same as for experiment 1, except that the vowels of the test syllables were restricted to

/i:/ and /ɪ/. However, for practical reasons there was a further difference between the two experiments, namely that in experiment 2 we measured only the alignment of the accentual peak (*H*) and did not measure the *F0* minimum (*L*) at the beginning of the accentual rise. In order to measure both *L* and *H* accurately, we would have needed test words in which the accented vowel is flanked by sonorants (e.g., /l/, /n/; we avoided /r/ as this makes a preceding /i:/ phonetically long). It turned out to be impossible to find enough test words containing /i:/ or /ɪ/ in the test syllable, if we imposed the sonorant-only condition on both flanking consonants and met all the other criteria of experiment 1. We therefore had to choose between (1) accurate measurement of both *L* and *H*, but with an insufficient sample size, or (2) accurate measurement of only *H*. Because the alignment of the *L* with the onset of the accented syllable seems well established now (not only for Dutch but also, as noted previously, for several other European languages), we decided to measure only the alignment of *H*.

On the expectation that there would be no difference in duration between the "long" and "short" accented vowels /i:/ and /ɪ/, we measured the alignment of *H* relative to the end of the accented vowel. If the durational explanation of the results of experiment 1 is correct, the difference in phonological vowel length should yield no difference in the alignment of the *H*. If, on the other hand, the structural explanation is correct, we should observe the same difference in the alignment of *H* observed in experiment 1, viz., *H* should be aligned at the end of /i:/ and during the consonant following /ɪ/.

1. Speech materials

There were 40 test sentences, all of a structure similar to the sentences used in experiment 1. All test words were adjectives. Twenty test syllables contained the vowel /i:/ and twenty the vowel /ɪ/, followed by a nasal or lateral. Great care was taken to avoid confounding vowel length with other potentially relevant factors, such as presence of a single consonant or a consonant cluster at the onset of the test syllable and the identity of the consonants at the offset of the test syllable. The list of test adjectives was matched as closely as possible for these factors [e.g., *grieneude* (/ˈxri:nəndə/; sobbing) and *grinnikend* (/xɪnəkənt/; sniggering), items 2 and 22, respectively]. None of the test syllables ended in consonant clusters. A full list of items is shown in the Appendix.

The items were presented on A4 sheets, with ten sentences per sheet. The test items were pseudo-randomly interspersed among 40 fillers. The first five and the last three sentences on the list were fillers. At the end of the list, there were three additional dummy sentences, announcing the end of the experiment.

2. Speakers

The speakers were seven undergraduate students at the University of Nijmegen, all native speakers of Dutch. Four were female and three were male. The speakers were paid for their participation.

TABLE II. Means and individual analyses for experiment 2. Standard errors in parentheses.

Speaker	Long	Short	<i>df</i>	<i>F</i>	<i>P</i>
V0 to C1 (duration of the vowel in ms)					
FMF	69.7 (2)	66.2 (5)	1,38	0.461	0.501
FND	66.3 (3)	59.6 (3)	1,37	2.932	0.095
FSH	67.0 (3)	67.1 (3)	1,38	0.002	0.961
FSK	63.1 (2)	63.5 (3)	1,38	0.014	0.908
MAC	65.7 (3)	70.7 (4)	1,38	1.018	0.319
MJP	56.0 (3)	56.4 (2)	1,38	0.01	0.923
MXC	63.8 (3)	55.9 (2)	1,38	4.085	0.050
Overall	64.5 (1)	62.8 (1)			
C1 to <i>H</i> (alignment of <i>H</i> relative to C1 in ms)					
FMF	11.2 (6)	21.6 (7)	1,37	1.270	0.267
FND	28.0 (4)	43.0 (5)	1,36	5.388	0.026
FSH	26.0 (8)	29.7 (7)	1,37	0.117	0.734
FSK	16.5 (7)	38.1 (7)	1,38	4.684	0.037
MAC	-2.3 (7)	8.1 (6)	1,32	1.091	0.304
MJP	36.9 (7)	33.8 (7)	1,38	0.104	0.749
MXC	27.9 (8)	46.1 (4)	1,38	4.546	0.040
Overall	20.6 (3)	31.5 (3)			
<i>F0</i> level at <i>H</i> (Hz)					
FMF	321 (8)	302 (6)	1,37	3.8	0.060
FND	334 (5)	329 (6)	1,36	0.4	0.517
FSH	286 (5)	286 (5)	1,37	0.0	0.978
FSK	297 (4)	295 (3)	1,38	0.2	0.690
MAC	155 (2)	150 (3)	1,32	1.7	0.203
MJP	142 (2)	138 (2)	1,38	2.3	0.134
MXC	183 (3)	181 (4)	1,38	0.3	0.591
Overall	246 (7)	240 (7)			

3. Recording and measurement procedures

The recording procedure was similar to that for experiment 1, with some minor differences, introduced largely for practical reasons. The main such difference was that speakers read the list of sentences only once. This was done mainly to save time, and proved not to lead to great loss of data, so it is not a crucial difference between the two procedures. Subjects were all recorded in the phonetics laboratory of the University of Nijmegen. The measurement procedure was identical to that of experiment 1, except for the fact that the third author made the measurements in consultation with the second author.

B. Results

As for experiment 1, all measures were analyzed using two-way (2×7) mixed design ANOVAs, with items as the random factor, Phonological Vowel Length as a between-items factor and Speaker as a within-items factor. In addition we carried out individual analyses, with Phonological Vowel Length as the single between-items factor. Individual means and analysis results are reported in Table II. Data were missing in 3.6% of cases, mainly due to speech errors or unexpected pitch contours (e.g., deaccenting of the target word).

1. Duration of the vowel and onset consonant(s)

The mean duration of /i:/ was 64.5 ms, and that of /ɪ/ 62.8 ms. Overall analyses showed that this difference was not significant, $F(1,37) < 1$. Although there was a significant effect of Speaker, $F(6,222) = 4.82$, $p < 0.0001$, reflecting

overall differences in vowel duration (possibly linked to differences in speech rate), crucially, the interaction between Phonological Vowel Length and Speaker was not significant, $F(6,222) = 1.22$, $p = 0.299$. The individual analyses revealed that only one speaker showed a significant effect of Phonological Vowel Length on vowel duration.

These analyses, on the whole, meet the central expectation of this experiment, namely that there should be no difference in phonetic vowel duration despite the difference in phonological vowel ‘length’ or tenseness. The findings also concur with those of Nooteboom and Slis (1972). This lack of difference of duration is, as noted in the introduction to experiment 2, crucial for the logic of the experiment.

As in experiment 1, the duration of the onset consonants did not form a confounding factor. The duration was 113 ms before phonologically long vowels and 108 ms before phonologically short vowels. This difference was not significant [$F(1,37) < 1$]. Note that, overall, the onset consonants are longer in this experiment than they were in experiment 1. This is a result of the inclusion of more words with consonant clusters in the onset of the test syllable in the latter experiment, which was a consequence of the constraints placed on the target words (see Sec. II A). We return to the duration of the onset consonant(s) in the Discussion.

2. Alignment of the *F0* maximum

For both long and short vowels the *F0* peak was aligned after the offset of the vowel, as predicted by the durational explanation of the findings of experiment 1. However, as predicted by the structural explanation, the peak was aligned earlier following /i:/ (20.6 ms after the vowel offset) than following /ɪ/ (31.5 ms after the vowel offset). This difference was significant, $F(1,28) = 9.64$, $p = 0.004$. There was also an effect of Speaker, $F(6,168) = 6.62$, $p < 0.0001$, but no interaction $F(6,168) < 1$. In the individual analyses, all but one speaker showed effects in the same direction, although it reached significance for only three speakers.

The alignment and vowel duration data for both experiments are presented graphically in Fig. 2. These results provide some support for both the durational and the structural explanations. In favor of the structural account is the fact that, even though the vowels are not of significantly different duration, the alignment of the peak is nevertheless significantly affected by the phonological length of the vowel. However, in support of the durational account is the fact that with both long and short vowels the alignment of the peak is during the consonant. Moreover, the difference in alignment between ‘long’ /i:/ and ‘short’ /ɪ/ is neither as large nor as statistically well-supported as it was in experiment 1. We return to this issue in the Discussion.

3. *F0* level

As before, we inspected *F0* level. Unlike in experiment 1, we could not calculate *F0* change from *L*, as *L* was not measured. Instead, we simply used the *F0* (in Hz) at *H* as a measure.

The maximum *F0* level was slightly but nonsignificantly higher in syllables with /i:/ (246 Hz) than in syllables

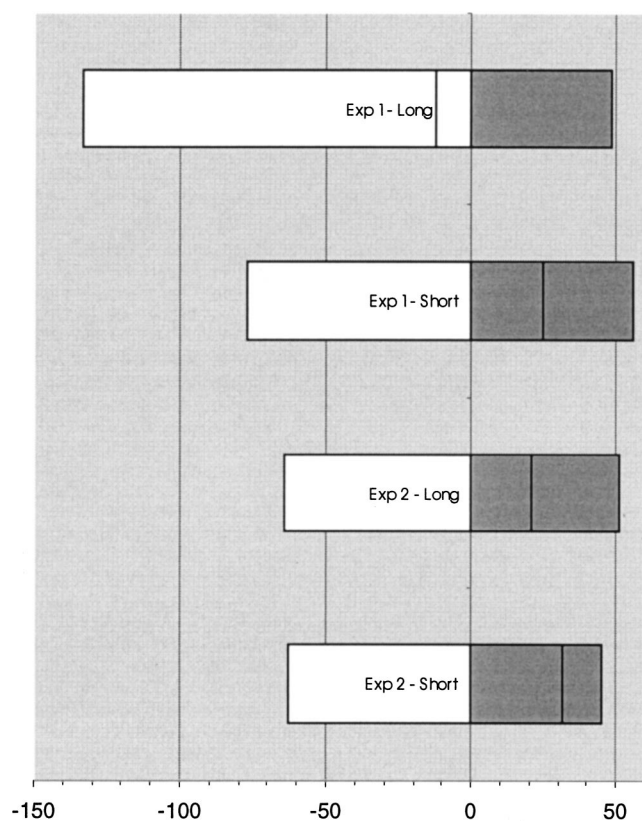


FIG. 2. Segmental durations and alignment for phonologically long and short vowels in experiments 1 and 2. C1 (the end of the vowel) is set at zero. White bars indicate the duration of the vowel. Shaded bars indicate the duration of the consonant. Vertical lines indicate the location of the H (F_0 maximum).

with /t/ (240 Hz), $F < 1$. The effect of Speaker was highly significant, $F(6,168) = 581.3$, $p < 0.0001$, but there was no significant interaction, $F(6,168) = 1.32$, $p > 0.25$. However, the individual analyses reported in Table II suggest that the effect of vowel length on F_0 level is negligible: In the individual analyses only one speaker showed a significant effect of Phonological Vowel Length. Once again, then, we can conclude that F_0 is not affected by differences in Phonological Vowel Length or by any associated differences in alignment.

C. Discussion

The results of this experiment suggest that we must give some credence to the durational explanation for the differences in alignment of F_0 maxima with phonologically long and phonologically short vowels. The F_0 peak is not aligned at the end of the vowel /i:/, which is what is predicted by a strict interpretation of the structural account. Instead, it is aligned somewhat later, during the following consonant. This could be explained in essentially durational terms, if we assume that /i:/, like /t/ and like the other short vowels in experiment 1, is somehow “too short” for the rise to be completed during the vowel.

At the same time, however, experiment 2 provides evidence of the effect predicted by the structural explanation as well. The vowels /i:/ and /t/ differ neither substantially nor significantly in duration, yet the alignment of the peak of an

accompanying F_0 rise is significantly different depending on which vowel is involved. The idea of a constant rise duration, which is the basis of the durational explanation, cannot account for this difference. One possibility, of course, is that the difference, though statistically significant, is so small as to be meaningless. However, we believe that it is possible to take the experimental results at face value, and to reconcile the two explanations in an interesting way.

Assume, with the structural explanation, that the “alignment point” for the peak of the F_0 rise is at the end of the long vowel, but in the middle of the following consonant after a short vowel. Assume further, with the durational explanation, that the F_0 rise takes a certain amount of time to achieve comfortably. Specifically, suppose that the time between the beginning of the accented syllable and the end of the vowel /i:/ is too short to complete the F_0 rise, but the time between the beginning of the accented syllable and the middle of the final consonant after a short vowel is just adequate. If all these premises are valid, then in the case of /i:/ the rise is subject to what Caspers and van Heuven called “time pressure,” but in the other cases we have considered it is not. There is no time pressure with mid or low long vowels, because the vowels themselves are longer than /i:/, and there is no time pressure with the short vowels, because the rise does not need to be completed until the middle of the following consonant.²

This more complex account—which we may call the “revised structural explanation”—makes a prediction that we can begin to test on our data. Given that the beginning of the F_0 rise is firmly fixed to the beginning of the syllable [cf. Caspers and van Heuven (1994) and our own results in experiment 1], then the actual duration of the syllable should affect the alignment of the peak in the case of /i:/, but not in the case of /t/. The reasoning is as follows. According to the revised structural explanation, there is no “time pressure” in the case of short vowels, so the time interval between the onset of the syllable and the middle of the following consonant should not affect the alignment of the F_0 peak. In the case of /i:/, however, we hypothesize that there is insufficient time to realize the F_0 rise completely, and that therefore the F_0 peak is aligned after the end of the vowel. Yet the longer the stressed syllable, the less severe this time pressure should be: There should be less deviation from the canonical peak alignment in the case of syllables with long onsets (consonant clusters, or consonants with intrinsically long duration such as /s/), and greater deviation in the case of syllables with short onsets (single consonants of intrinsically short duration, such as /n/ or tapped /r/).

While our materials were not set up to test this specific prediction, we can evaluate the hypothesis in a preliminary way. If the hypothesis is true, there should be a negative correlation between the length of the syllable and the alignment of the F_0 peak in the case of /i:/ (the longer the syllable, the earlier the F_0 peak should be aligned, relative to C1). There should be no correlation in the case of /t/, because time pressure is hypothesized not to be relevant.

We calculated the correlations in the following way. For each item, we averaged H minus C1 and C1 minus C0 over the seven speakers. We then ran separate Pearson correla-

tions for the /i:/ and /ɪ/ items. The results are consistent with the hypothesis: For /i:/, $r = -0.5795$, $N = 20$, $p = 0.007$, and for /ɪ/, $r = -0.2989$, $N = 20$, $p = 0.201$ (n.s.). Thus for /i:/, there is a clear correlation: The longer the stressed syllable, the earlier the alignment. For /ɪ/, there is some tendency in the same direction, but the correlation is smaller and more likely to have occurred by chance. This difference between the two cases is predicted by the revised structural explanation, but not by the durational explanation alone.

This result in turn is relevant to the central question posed by Caspers and van Heuven (1994), namely which features of a given F_0 movement are preserved when there is insufficient time to realize the movement fully. In the case of prenuclear rises and Dutch /i:/, the answer seems clear: The alignment of the peak is delayed so that the target F_0 level can be achieved. If this is the correct conclusion, one further prediction follows for our data: Unlike what we found with alignment, there should be *no* correlation between syllable duration and F_0 peak for either /i:/ or /ɪ/. That is, if the F_0 peak in fact has priority, and the alignment is adjusted to allow for time pressure, then syllable duration should not account for any variability in the F_0 level of the peak.

We analyzed the correlations between syllable duration and F_0 peak in the same way as those involving alignment. The results are as follows: For /i:/, $r = 0.1472$, $N = 20$, $p = 0.53$, and for /ɪ/, $r = 0.2702$, $N = 20$, $p = 0.24$. This lack of correlation is consistent with the conclusion just proposed. While it is, of course, unwise to rely on negative results, this specific result is very much in line with other work. In the present study, we have failed to find any effect of our experimental manipulations on the F_0 excursions (experiment 1) or the F_0 peaks (experiment 2) associated with the F_0 movements studied. The same lack of effect was noted by Prieto *et al.* in their study of Spanish (1995: p. 435) and by Arvaniti *et al.*'s work on Greek (1998). Note that, as pointed out by Arvaniti *et al.* (1998: p. 24), the apparent invariability of F_0 levels under these experimental manipulations is consistent with the phonological interpretation of the F_0 movement as a transition between intonational targets that have specified alignments and F_0 levels.

III. GENERAL DISCUSSION

The evidence from our two experiments points clearly to the relevance of phonological structure in determining the alignment of F_0 movements in speech. Specifically, we have suggested that the Dutch F_0 rises under study are anchored to two places in the segmental string: The beginning of the rise is aligned with the beginning of the onset consonant of the accented syllable, and the end of the rise with the end of the syllable. In our materials, the "end of the syllable" is at the end of the vowel if the vowel is phonologically long, and in the following consonant if the vowel is short.

This finding superficially vindicates Xu's (1998) insistence on the importance of the syllable in determining the alignment of F_0 movements. However, by its very conformity with Xu's views this finding may actually create a problem for them. Recall that Arvaniti *et al.* found that the end of the prenuclear accentual rise in Greek is at the beginning of the following unstressed vowel (i.e., well outside the bounds

of the syllable with which the F_0 movement is associated), while Xu himself found that the beginning of the rise in Chinese rising tone need not occur until well into the vowel of the affected syllable. By contrast, in our Dutch data the alignment points appear to be straightforwardly at the beginning and end of the syllable. Similarly, recall that results from the European languages (including our own results from experiment 2) show that time pressure affects the alignment of the F_0 peak, whereas Xu's findings for Chinese show that the alignment of the F_0 valley is more variable. If, as Xu maintains, the basic alignment principle for F_0 movements is that they align with the syllable, then it is necessary to explain why the details are so different from one language to another.

We have also shown that, when the actual duration of the "long" vowel is relatively short (as regularly happens in the case of Dutch high vowels), the end of the rise is aligned later than the end of the vowel, though still earlier than the end of the rise accompanying a short vowel. This is apparently because there is not sufficient time to realize the rise fully during the interval between the structurally specified beginning and ending points. As just noted in the discussion section of experiment 2, this interpretation is strengthened by the existence of a significant correlation between syllable duration and alignment for the "long" vowel, but not for the short vowel. This finding adds to Caspers and van Heuven's work on the effects of "time pressure" on F_0 movements, but again it provides a new level of detail that may make it difficult to maintain some of Caspers and van Heuven's conclusions. Specifically, we have provided evidence that the alignment distinction between long and short vowels in Dutch is not really about "time pressure" at all—which is what Caspers and van Heuven surmised—but about differently specified alignment points based on the different syllable structures for long and short vowels. True effects of time pressure may be restricted to the case of the phonetically short "long" vowels like /i:/.

In summary, our results shed light on a number of issues in recent work on F_0 alignment. We have provided support for Xu's view that syllable structure is relevant, and more generally for the finding (first reported by Silverman and Pierrehumbert, 1990) that a number of structural variables are relevant for modeling alignment. We have also, building on Caspers and van Heuven's work, provided further details of the effects that "time pressure" exerts on the precise realization of F_0 rises, while at the same time showing that phonological vowel length does not constitute time pressure in and of itself. We believe that we can best integrate our new data into these recent lines of research by adopting a phonological model in which F_0 movements are specified in terms of two distinct targets that can be independently aligned with the segmental string. In approximate terms, it is correct to say (with Xu) that F_0 movements align with syllables, and (with Caspers and van Heuven) that time pressure may affect the way F_0 movements are realized. However, when we look more closely at the differences of detail that have been documented between one syllable structure and another or between one language and another, we find that

both the beginning and the end of the movement can and must be precisely specified.

ACKNOWLEDGMENTS

Ineke Mennen's work on experiment 1 formed part of her Ph.D. research, which was funded in part by the University of Edinburgh and in part by the U.K. Economic and Social Research Council (ESRC) through Grant No. R000-23-5614 to Edinburgh University. Work on experiment 2 was funded by the ESRC through Grant No. R000-23-7447 to Edinburgh University, as part of the project "Alignment of Fundamental Frequency Targets in English and Dutch." We thank staff at the Universities of Amsterdam and Groningen for the use of their Phonetics Laboratories for experiment 1. We thank Carlos Gussenhoven for much practical help with experiment 2, and all our speakers for their participation in our experiments. Finally, we thank Anders Löfqvist, Yi Xu, and Johanneke Caspers for their comments on an earlier draft of this paper.

APPENDIX: MATERIALS FOR EXPERIMENT 1

Target words are preceded by asterisks, and transcribed phonemically at the end of each sentence. Although the phonologically long high vowels are usually indicated by a "half-length" mark in Dutch, we have indicated all phonologically long vowels with the standard IPA length mark.

1. Materials for Experiment 1

Sentences 1–20 have phonologically short vowels and sentences 21–40 have phonologically long vowels.

- (1) Hij wilde de *rillende kinderen tracteren op warme chocolademelk. /'rɪləndə/ He wanted to buy the shivering children some hot chocolate.
- (2) Met haar *beminnelijk gedrag kon ze iedereen om haar vinger winden. /bə'mɪnələk/ With her amiable behavior she could wrap everyone around her little finger.
- (3) Je moet haar *bedillerig gepraat maar langs je heen laten gaan. /bə'dɪlərəx/ You should just ignore her meddling talk.
- (4) Hij kon van de *Limburgse vlaaien maar geen genoeg krijgen. /'lɪmbʏrksə/ He could not get enough of the Limburg fruit tarts.
- (5) Wij konden de *rennende atleten met geen mogelijkheid bijhouden. /'rɛnəndə/ There was no way we could keep up with the running athletes.
- (6) Hij kon de *remmende auto nog net ontwijken. /'rɛməndə/ He only just managed to avoid the braking car.
- (7) Ze wilden hun *belemmerend schoeisel zo gauw mogelijk uittrekken. /bə'ləməərənt/ They wanted to remove their constricting footwear as soon as possible.
- (8) Zij had de *bedremmelde meisjes beter naar huis kunnen sturen. /bə'drɛməldə/ It would have been better if she had sent the embarrassed girls home.
- (9) Ze waren door hun *mollige gezichten ongeschikt voor modellenwerk. /'mɒləxə/ Because of their plump faces, they were unsuitable for modeling.

- (10) Ik kon mijn *morrende collega niet meer luchten of zien. /'mɒrəndə/ I could no longer stand the sight of my grumbling colleague.
- (11) Hij kon met zijn *rollende ogen alle kinderen angst aanjagen. /'rɒləndə/ He could frighten all the children with his rolling eyes.
- (12) We kunnen de *mondige studenten van tegenwoordig niet meer de baas. /'mɒndəxə/ We can't keep the upper hand with today's assertive students.
- (13) Ze hebben de *genummerde kwitanties helemaal door elkaar gegooid. /xə'nɪmərdə/ They have messed up the order of the numbered invoices.
- (14) Je kan de *lullige verhalen van Willem beter niet serieus nemen. /'lɪləxə/ You shouldn't take Willem's cruddy stories seriously.
- (15) Hij was door zijn *lummelige houding niet bepaald aantrekkelijk. /'lɪmələxə/ His oafish attitude did not make him particularly attractive.
- (16) Hij zag de *brullende gorilla's opgewonden heen en weer rennen. /'brɪləndə/ He saw the roaring gorillas run up and down full of excitement.
- (17) Ze wordt door haar *mannelijk karakter vaak voor een man versleten. /'mənələk/ Because of her manly behavior, she is often mistaken for a man.
- (18) Hij kan zijn *rammelend betoog maar beter herschrijven. /'ramələnt/ He'd better rewrite his rambling story.
- (19) Je kon aan zijn *lallende uitspraak gelijk horen dat hij dronken was. /'lələndə/ You could tell he was drunk from his slurred pronunciation.
- (20) Ik was de *drammende telefoontjes van dat vervelende mens spuugzat. /'drəməndə/ I was sick of the pestering phone calls of that annoying woman.
- (21) Je moet in *aluminium pannen geen azijn gebruiken. /ɑly:'mi:nɪjəm/ You should not put vinegar in aluminium pans.
- (22) De jury kon de *lyrische stijl van zijn roman niet bepaald waarderen. /'li:rɪsə/ The jury did not particularly appreciate his novel's lyrical style.
- (23) Hij had de *alineas van zijn werkstuk veel te lang gemaakt. /a:'li:ni:ja:s/ He had made the paragraphs of his essay far too long.
- (24) Je moet je *dierlijke instinct niet altijd onderdrukken. /'di:rləkə/ You should not always suppress your animal instincts.
- (25) Ze moesten de *lenige assistente uit haar benarde positie bevrijden. /'le:nəxə/ They had to free the supply assistant from her awkward position.
- (26) Ik kan de *melige grappen van Seth Gaaikema niet meer aanhoren. /'me:ləxə/ I can no longer stand Seth Gaaikema's corny jokes.
- (27) Hij heeft voor de *dinerende gasten een optreden verzorgd. /di:'ne:rəndə/ He has arranged a performance for the dining guests.
- (28) Door zijn *belerende gepreek joeg hij iedereen de kerk uit. /bə'le:rəndə/ He managed to empty the church with his pedantic sermons.
- (29) Je moet bij *meerdere geschiktheid je persoonlijke

voorkeur laten varen. /'me:rdəɾə/ In case of greater suitability, you should abandon your personal preference.

- (30) Je moet in *romige sauzen een scheutje cognac doen; dat is lekker. /'ro:məxə/ You should put some brandy in creamy sauces; that is nice.
- (31) Hij moest de *beloningen van de politie gaan inventariseren. /bə'lo:niŋə/ He had to go and catalogue the police rewards.
- (32) Het kan in een *noordelijk klimaat 's winters flink koud zijn. /no:rdələk/ It can get very cold in winter in a northern climate.
- (33) Hij wil met de *naburige kerkdorpen een belangengroep oprichten. /na:'by:rəxə/ He wants to start an interest group with the neighbouring villages.
- (34) Hij kon de *ongedurige paarden niet meer in bedwang houden. /ɔnxə'dy:rəxə/ He could no longer keep the wild horses under control.
- (35) Ze willen de *Lunense heide weer openstellen voor het publiek. /'ly:nənsə/ They want to re-open the Lunen Heath to the public.
- (36) Je moet de *duurdere produkten vooraan in de vitrinekast zetten. /'dy:rdəɾə/ You should put the more expensive items at the front of the display.
- (37) We zoeken een *dynamische dertiger voor deze veelzijdige functie. /di:'na:misə/ We are seeking a dynamic person in their thirties for this multi-faceted job.
- (38) Hij kon de *malende gedachten niet uit het hoofd zetten. /'ma:ləndə/ He could not get the persistent thoughts out of his mind.
- (39) Hij is in zijn *manische periode; dan is hij altijd zo druk. /'ma:nisə/ He is in his manic phase; he is always hyperactive then.
- (40) Hij moest de *dralende studenten tot drie keer toe roepen. /'dra:ləndə/ He had to call the lingering students three times.

2. Materials for Experiment 2

Sentences 1–20 have phonologically long vowels, and sentences 21–40 have phonologically short vowels.

- (1) Hij kon de *priemende vragen niet beantwoorden. /'pri:məndə/ He could not answer the probing questions.
- (2) Ze had haar *grieneende achterneefje niet kunnen troosten. /'xri:nəndə/ She had not been able to console her sobbing nephew.
- (3) Ze kon de *bezielende toespraak van haar baas maar niet vergeten. /bə'zi:ləndə/ She could not forget the inspiring speech by her boss.
- (4) Hij had de *zielige weeskinderen een kop thee gegeven. /'zi:ləxə/ He had given the pitiful orphans a cup of tea.
- (5) Er stond een *minimum bedrag voor het openen van een rekening. /'mi:ni:mym/ There was a minimum amount for opening an account.
- (6) Ze had haar *ielige assistente gevraagd of ze kon helpen. /'i:ləxə/ She had asked her tiny assistant whether she could help.
- (7) Hij had de *kiemende zaadjes in de grond gestopt. /'ki:məndə/ He had put the germinating seeds into the soil.
- (8) Ze had de *striemende regen nauwelijks opgemerkt. /'stri:məndə/ She had hardly noticed the pouring rain.
- (9) Hij had de *kienende bejaarden een grote prijs gegeven. /'ki:nəndə/ He had given the pensioners who were playing bingo a large prize.
- (10) Hij had zijn *labielere patienten extra pillen gegeven. /la'bi:ləɾə/ He had given his more labile patients extra pills.
- (11) Ze had de *knielende nonnen in het klooster erg bewonderd. /'kni:ləndə/ She had admired the kneeling nuns in the nunnery.
- (12) Hij had de *wriemelende wormen in de tuin vol walgend bekeken. /'vri:mələndə/ She had looked at the wriggling worms in the garden in disgust.
- (13) Hij kon de *cynische opmerkingen van Jan niet meer verdragen. /'si:nisə/ He could no longer take Jan's cynical remarks.
- (14) Ze kon haar *klinische praktijk dit voorjaar afronden. /'kli:nisə/ She was able to complete her clinical practice this spring.
- (15) Ze vond de *intiemere sfeer van de Rozenboom geschikter. /m'ti:məɾə/ She preferred the more intimate atmosphere of the Rose Tree.
- (16) Ze had de *friemelende sollicitant niet serieus genomen. /'fri:mələndə/ She had not taken the fidgety applicant seriously.
- (17) Hij vond de *kienere eerstejaars het makkelijkst in de les. /'ki:nəɾə/ He found the keener first year easiest in class.
- (18) Hij had een *miniemere poging gedaan dan zijn collega. /mi:'ni:məɾə/ He had made a more minimal attempt than his colleague.
- (19) Ze had een *stabielere partner voor haar bedrijf gevonden. /sta:'bi:ləɾə/ She had found a more stable partner for her company.
- (20) Ze had haar *verdienende dochters geen zakgeld meer gegeven. /fər'di:nəndə/ She had stopped giving her working daughters pocket money.
- (21) Hij kon de *glimmende stenen niet goed verbergen. /'xli:məndə/ He could not easily hide the shiny stones.
- (22) Ze had een *grinnikend kamermeisje om dekens gevraagd. /'xɾi:nəkənt/ She had asked a sniggering chambermaid for blankets.
- (23) Ze kon het *bedillerig gedrag van Carla niet accepteren. /bə'diləɾəx/ She could not accept Carla's interfering behavior.
- (24) Er stonden *gillende tieners op een beroemde zanger te wachten. /'xiləndə/ Some screaming teenagers were waiting for a famous singer.
- (25) Er stond een *hinnikend veulen in de wei van boer Jansen. /'hinəkənt/ There was a whinnying foal in farmer Jansen's field.
- (26) Ze had zijn *innige begroeting maar niet kunnen vergeten. /'i:nəxə/ She had not been able to forget his warm welcome.

- (27) Hij had zijn *pinnige moeder een voorstel gedaan. /'pɪnəxə/ He had made his crabby mother a proposal.
- (28) Hij had de *trimmende mannen hartelijk aangemoedigd. /'trɪməndə/ He had cheered on the jogging men.
- (29) Hij had de *binnenste omheining een heel stuk hoger gemaakt. /'bɪnəstə/ He had made the inner fence a whole lot higher.
- (30) Hij vond haar *verschillende mening niet acceptabel. /fər'sxɪləndə/ He did not find her differing opinion acceptable.
- (31) Ze kon de *grillige vormen in zijn kunstwerk niet waarderen. /'xɪlɪəxə/ She could not appreciate the uneven shapes in his piece of art.
- (32) Ze vond de *spinnende poes van haar vriend ontspannend gezelschap. /'spɪnəndə/ She found her friend's purring cat relaxing company.
- (33) Hij had de *zinnige vragen van Peter met zorg beantwoord. /'zɪnəxə/ He had carefully answered Peter's sensible questions.
- (34) Ze had zijn *grimmige schilderij in de kelder gezet. /'xɪmɪəxə/ She had put his sombre painting in the basement.
- (35) Ze vond de *betimmerde muur in haar buitenhuis niet zo mooi. /bə'tɪmərðə/ She did not like the wood-covered wall in her house in the country.
- (36) Hij had de *schimmelende marmelade in de vuilnisbak gegooid. /'sxɪmələndə/ He had thrown the moldy marmalade in the rubbish bin.
- (37) Hij had de *rillende peuters warme chocolademelk gegeven. /'rɪləndə/ He had given the shivering toddlers some hot chocolate.
- (38) Hij had zijn *beginnende keelpijn totaal verwaarloosd. /bə'xɪnəndə/ He had neglected his beginning sore throat.
- (39) Ze had een *vrijwillige bijdrage aan het project geleverd. /frɛi'vɪləxə/ She had contributed voluntarily to the project.
- (40) Ze had haar *beminnende fans achter hun rug uitgelachen. /bə'mɪnəndə/ She had laughed at her adoring fans behind their backs.

¹In the system of 't Hart *et al.* (1990), the accents we measured are type 1 rises followed by either type B or type D falls. The 't Hart *et al.* system does not distinguish "prenuclear" from "nuclear" accents in exactly the same way as the British and American traditions, but in the present context "prenuclear" could be glossed as type 1 accents followed by type B or D falls, while "nuclear" rising-falling accents would be type 1 followed by type A, i.e., the "pointed hat" accents of 't Hart *et al.*

²It is difficult to be sure of the absolute minimum time required to complete an accentual pitch rise of the sort we are considering here. Physiological limits on F0 rise rates in nontrained singers reported by Sundberg (1979) might suggest a minimum of 100 ms for a rise of 6.5 semitones, which is

the approximate average pitch rise in our experimental data. This is considerably shorter than the test syllable durations in experiment 2, which averaged about 175 ms. However, as pointed out by a referee, Sundberg reported the F0 change rate for the middle part of the rise, ignoring the acceleration and deceleration, so that the real minimum rise time may be considerably longer than 100 ms. Furthermore, it is not clear whether it is valid to generalize from studies of singing voice to normal speaking voice, and neither is it clear whether any other factors are involved [e.g., the proposal by House (1990) that speakers prefer to produce F0 movements in periods of relative spectral stability, such as during the vowel]. These considerations are all possible directions for further investigation, but do not materially affect the discussion here: Our point is simply that *if* physical time pressure is relevant to peak alignment, then it is plausible that it might affect only rises with /i:/ and not those with /ɪ/.

- Arvaniti, A., Ladd, D. R., and Mennen, I. (1998). "Stability of tonal alignment: the case of Greek prenuclear accents," *J. Phonetics* 26, 3–25.
- Bruce, G. (1977). *Swedish Word Accents in Sentence Perspective* (Gleerup, Lund).
- Caspers, J. C. (1994). "Pitch movements under time pressure," The Hague: Holland Academic Graphics. Ph.D. Thesis, University of Leiden.
- Caspers, J., and van Heuven, V. J. (1993). "Effects of time pressure on the phonetic realization of the Dutch accent-lending rise and fall," *Phonetica* 50, 161–171.
- Fujisaki, H. (1983). "Dynamic characteristics of voice fundamental frequency in speech and singing," in *The Production of Speech*, edited by Peter F. MacNeilage (Springer-Verlag, New York), pp. 39–55.
- House, D. (1990). *Tonal Perception in Speech* (Lund University Press, Lund).
- Ladd, D. R. (1996). *Intonational Phonology* (Cambridge University Press, Cambridge).
- Ladd, D. R., Faulkner, D., Faulkner, H., and Schepman, A. (1999). "Constant segmental anchoring of F0 movements under changes in speech rate," *J. Acoust. Soc. Am.* 106, 1543–1554.
- Mennen, I. (1998). "Second language acquisition of intonation: the case of peak alignment," in *Chicago Linguistic Society 34, Volume II: The Panels*, edited by M. C. Gruber, D. Higgins, K. Olson, and T. Wysocki.
- Mennen, I. (1999). "Second language acquisition of intonation: The case of Dutch near-native speakers of Greek," Ph.D. Thesis, University of Edinburgh.
- Nooteboom, S. G., and Slis, I. H. (1972). "The phonetic feature of vowel length in Dutch," *Lang. Speech* 15, 301–316.
- Pierrehumbert, J. B. (1980). "The Phonology and Phonetics of English Intonation," Ph.D. Dissertation, MIT.
- Prieto, P., van Santen, J., and Hirschberg, J. (1995). "Tonal alignment patterns in Spanish," *J. Phonetics* 23, 429–451.
- Schiller, N. O., Meyer, A. S., and Levelt, W. J. M. (1997). "The syllable structure of spoken words: Evidence from the syllabification of intervocalic consonants," *Lang. Speech* 40, 103–140.
- Silverman, K., and Pierrehumbert, J. (1990). "The timing of prenuclear high accents in English," in *Papers in Laboratory Phonology I: Between the Grammar and Physics of Speech*, edited by J. Kingston and M. Beckman (Cambridge University Press, Cambridge), pp. 71–106.
- Sundberg, J. (1979). "Maximum speed of pitch change in singers and untrained subjects," *J. Phonetics* 7, 71–97.
- 't Hart, J., Collier, R., and Cohen, A. (1990). *A Perceptual Study of Intonation: An Experimental-Phonetic Approach to Speech Melody* (Cambridge University Press, Cambridge).
- Treiman, R., and Danis, C. (1988). "Syllabification of intervocalic consonants," *Journal of Memory and Language* 27, 87–104.
- Xu, Yi (1998). "Consistency of tone-syllable alignment across different syllable structures and speaking rates," *Phonetica* 55, 179–203.

Neurophysiologic correlates of cross-language phonetic perception

Anu Sharma^{a)} and Michael F. Dorman

Department of Speech and Hearing Science, Arizona State University, Tempe, Arizona 85287-0102

(Received 6 July 1999; revised 24 January 2000; accepted 28 January 2000)

This study examined neurophysiologic correlates of the perception of native and nonnative phonetic categories. Behavioral and electrophysiologic responses were obtained from Hindi and English listeners in response to a stimulus continuum of naturally produced, bilabial CV stimuli that differed in VOT from -90 to 0 ms. These speech sounds constitute phonemically relevant categories in Hindi but not in English. As expected, the native Hindi listeners identified the stimuli as belonging to two distinct phonetic categories (/ba/ and /pa/) and were easily able to discriminate a stimulus pair across these categories. On the other hand, English listeners discriminated the same stimulus pair at a chance level. In the electrophysiologic experiment N1 and MMN cortical evoked potentials (considered neurophysiologic indices of stimulus processing) were measured. The changes in N1 latency which reflected the duration of pre-voicing across the stimulus continuum were not significantly different for Hindi and English listeners. On the other hand, in response to the /ba-/pa/ stimulus contrast, a robust MMN was seen only in Hindi listeners and not in English listeners. These results suggest that neurophysiologic levels of stimulus processing reflected by the MMN and N1 are differentially altered by linguistic experience. © 2000 Acoustical Society of America. [S0001-4966(00)02805-8]

PACS numbers: 43.71.An, 43.71.Hw, 43.71.Pc, 43.64.Sj [CWT]

INTRODUCTION

A robust finding in the literature on infant speech perception is that infants 6–8 months old can discriminate speech sounds regardless of whether or not they are present in the infant's language environment. However, older infants and adults are less accurate in their perception of speech sounds that are not relevant to their native language (see review by Werker, 1994). These findings demonstrate that listeners' speech perception abilities are altered by experience with a particular language and that a lack of experience with a particular phonetic contrast has the effect of reducing sensitivity to that contrast. Certain nonnative contrasts appear to be easier to discriminate than others, perhaps due to exposure to their phonetic quality through some shared articulatory or phonetic features with sounds from one's own language (Best, 1994; Tees and Werker, 1984; Polka, 1991), and/or acoustic salience of the contrast (Burnham, 1986). Moreover, although for some contrasts re-learning is relatively easy using auditory training methods, for other contrasts it is often very difficult (see reviews by Pisoni *et al.*, 1994; Logan and Pruitt, 1995).

While there is overwhelming evidence for a "perceptual deficit" in discrimination of some nonnative speech contrasts for adult second language learners [e.g., Werker, 1994 (review)], the mechanisms of this deficit remain unclear. The prevailing view is that this deficit occurs due to re-alignment of cognitive categories resulting from "higher-level" attentional biases rather than a neural-sensory loss (see reviews by Pisoni *et al.*, 1994 and Werker, 1994). In fact, however,

little is known about whether underlying neural-sensory representations of speech are actually altered by an individual's linguistic experiences.

A noninvasive method of directly examining neurophysiologic correlates of perceptual processes in humans is recording auditory evoked potentials (AEPs). The N1 and Mismatch Negativity (MMN) auditory evoked potentials, in particular, have proved useful in investigating neural correlates of speech processing. The N1 is an evoked response whose latency and morphology reflects the time of onset of acoustic events *within* speech, such as burst onset relative to voicing onset in a stop-consonant vowel syllable (Sharma and Dorman, 1999) and the onset of friction relative to vowel onset in a fricative-vowel syllable (Kaukoranta *et al.*, 1987; Ostroff *et al.*, 1999). The mismatch negativity (MMN) evoked response, on the other hand, is elicited by changes *between* acoustic signals (Sharma *et al.*, 1993; Sharma and Dorman, 1999). Primary generator sites for the N1 and MMN include the auditory cortical and thalamic areas (Näätänen and Picton, 1987; Csepe, 1995). Both the N1 and MMN have traditionally been considered neurophysiologic indices of pre-attentive processing since they are recorded without active participation from subjects, whose attention is actively directed away from the eliciting stimuli by reading or watching a movie. Furthermore, the MMN can be elicited when the eliciting stimuli are not behaviorally discriminated, e.g., when stimuli are from within the same place of articulation category (Sams *et al.*, 1990; Sharma *et al.*, 1993; Maiste *et al.*, 1995). For these reasons, the MMN has been considered a measure of "sensory" or noncognitive processing (Näätänen, 1992).

However, recently two studies which have examined

^{a)}Electronic mail: anu.sharma@asu.edu

cross-language speech perception have raised doubts as to whether the MMN reflects a sensory level of processing. In one study, Näätänen *et al.* (1997) examined perception of Finnish vowels in Estonian and Finnish listeners, and in the other study Winkler *et al.* (1999) examined perception of Hungarian and Finnish vowel contrasts in speakers of those two languages. The overall conclusion of these studies was that the MMN responses elicited by contrasts that were not phonetically relevant in the listeners' native language were significantly smaller than the MMN responses evoked by vowels that constituted different phonetic categories in the native languages of those same listeners. One interpretation of these results of diminished MMN responses for nonnative vowel contrasts is that the sensory level of processing traditionally attributed to the MMN is modified by language experience. However, another interpretation of these results is that the MMN does not reflect a entirely sensory level of processing, but rather a higher level susceptible to linguistic influences.

In this report we describe the results of an experiment in which the N1 and MMN evoked potentials were obtained from Hindi and English speakers in response to speech sounds varying in voice-onset-time which constituted different phonetic categories in Hindi and but not in English. If language experience alters the sensory level of processing, then both the N1 and the MMN responses to stimuli along the continuum would be different for Hindi speakers compared to English speakers. On the other hand, if only the MMN (and not the N1) response to stimuli along the continuum were different for Hindi speakers compared to English speakers, then such a finding would strengthen the possibility that the MMN reflects a level of processing which is beyond the sensory level.

I. BEHAVIORAL EXPERIMENT

A. Method

1. Subjects

Ten native Hindi speakers were recruited from the Arizona State University (ASU) campus. These were students who had spent at least the first 20 years of their life in India and had come to ASU for graduate studies. Subjects reported that they were fluent in several Indian languages (including Hindi) as well as English. Additionally, ten monolingual speakers of American-English were recruited from the ASU campus. These subjects reported no previous exposure to the Hindi language. All subjects were paid (\$10/hour) for their participation. The subjects reported that they had no history of hearing, speech, or language problems.

2. Stimuli

The first author recorded the syllables /ba/ and /pa/ into a digital signal processing system. The system had 16 bit resolution and ran at 11.2 kHz. The vowels were produced with a slight /r/ coloring. The consonants were produced with pre-voicing as is appropriate in Hindi. The syllables were acceptable to native-English speakers as /ba/ and were acceptable to native-Hindi speakers as "baar" (which means "again" in Hindi) and "paar" (which means "side" in

Hindi). To create the experimental stimuli, the pre-voicing segment of the original utterances were edited at zero crossings to create stimuli with pre-voicing durations of 0 to 90 ms (see Fig. 1).

3. Procedures

In the identification portion of the experiment, Hindi subjects were asked to listen to the experimental stimuli through headphones and to classify each stimulus either as a /pa/ (or "paar") or /ba/ (or "baar"). Subjects were asked to indicate their responses by clicking with a computer mouse on panels marked PA and BA appearing on a computer screen. Each subject was given an initial practice session where he or she heard each stimulus along the continuum once. After that, ten repetitions of each of the nine stimuli were presented to the subject in a random order.

Following the identification task, subjects participated in a discrimination task. Based on results from a pilot experiment the stimulus pair chosen for the discrimination experiment had -10 and -50 ms VOT. An AX discrimination task was employed. On each trial, subjects heard two stimuli with an interstimulus interval (ISI) of 500 ms. Subjects were asked to determine whether the stimuli in the pair were "same" or "different." Subjects indicated their responses by clicking on panels labeled "same" or "different" on the computer screen. The presentation of stimulus pairs was randomized within the test and across subjects. To familiarize the subjects with the task, an initial practice session of 20 trials was presented (10 same and 10 different trials). The experimental session consisted of a total of 100 trials (50 same and 50 different) for each stimulus pair.

B. Results

The group mean identification scores for each stimulus token are shown in Fig. 2. As can be seen, Hindi listeners consistently (i.e., $>75\%$ of the time) identified stimuli with VOT's of 0 and -20 ms as /pa/ and -50 and -90 ms as /ba/. On the other hand, English listeners reliably identified *all* the stimuli as /ba/.

The mean discrimination scores for the -10 and -50 ms VOT stimulus pair are shown in Fig. 3. Hindi subjects were able to discriminate the stimuli as different with a high degree of accuracy, while English listeners discrimination of the stimulus pair was close to chance. A paired *t*-test revealed that Hindi subjects' discrimination of the stimuli was significantly more accurate than their English counterparts ($t = 10.8$; $df = 18$; $p < 0.0001$). Overall, for the Hindi subjects, the identification and discrimination results are consistent with previous reports of categorical perception for similar stimulus continua. On the other hand, for English listeners, the identification and discrimination results are consistent with previous reports that in the absence of training changes in pre-voicing are not phonemically perceived by monolingual speakers of American English [e.g., Werker, 1994 (review)].

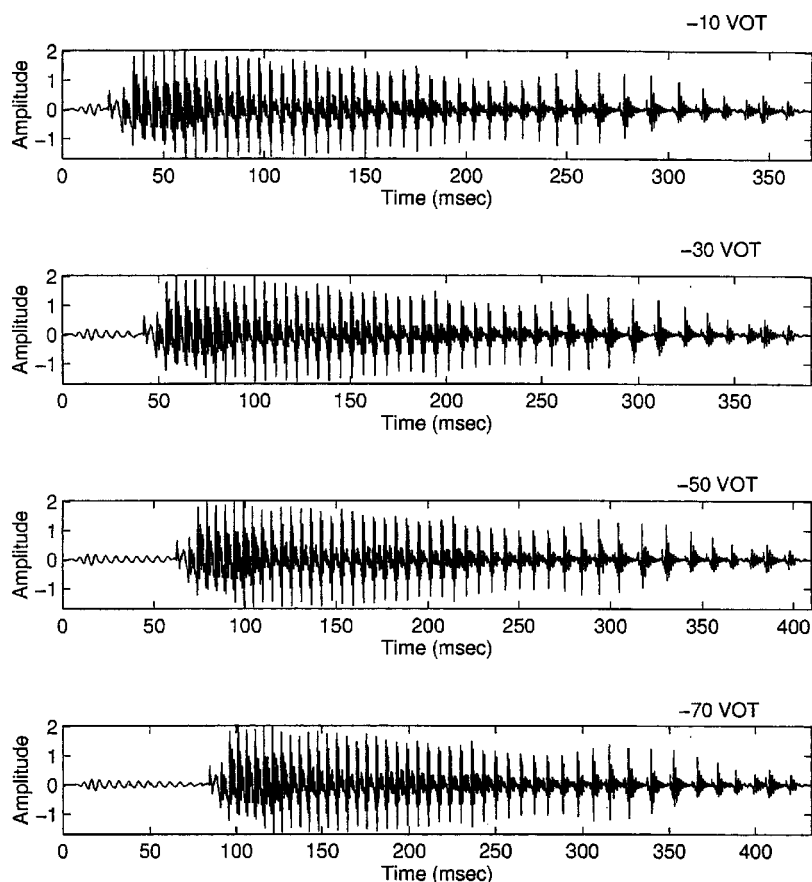


FIG. 1. Stimulus waveforms showing the duration of pre-voicing for four stimuli from the /ba/–/pa/ continuum. The voice-onset time (VOT) is indicated at the upper right-hand corner of the waveforms.

II. ELECTROPHYSIOLOGIC EXPERIMENT

A. Method

1. Subjects

For the electrophysiologic recording sessions the subjects were seated comfortably in a sound booth. To control for arousal and to minimize subjects' attention to the test stimuli, subjects watched a videotaped movie of their choice. Video tape audio levels were kept below 40 dB SPL. Subjects were asked to ignore the test stimuli that were presented through an insert earphones at 75 dB SPL in the right ear.

2. Stimuli

a. MMN. For the MMN recordings, the stimuli were the same as those used in the behavioral discrimination experiment (i.e., the -10 and -50 ms VOT pair). The MMN was elicited using an oddball paradigm in which repetitive presentations of a “standard” stimulus were occasionally replaced with a “deviant” or “target” stimulus. The deviant stimulus had a probability of occurrence of 15%. The -10 ms VOT stimulus was the standard and -50 ms VOT stimulus was the deviant. The stimuli were presented at an offset-to-onset ISI of 510 ms.

b. N1. The stimulus continuum was identical to the one used in the behavioral identification experiment. Repeated presentations of each stimulus separated by an onset-to-offset ISI of 800 ms were used to elicit the N1. The order of presentation of stimuli was counterbalanced across subjects.

3. Recording procedures

MMN and N1 evoked potentials were recorded using a NeuroScan, Inc. data acquisition system. Silver-chloride electrodes were placed on the midline (Fz, Cz, and Pz). The reference electrode was placed on the right mastoid and the ground electrode was placed on the forehead. Eye movements were monitored with a bipolar electrode montage (supraorbital to lateral canthus). Averaging was suspended when the eye channel registered blinks. The recording window included a 100-ms prestimulus period and 500 ms of post-stimulus time. AEPs were bandpass filtered on-line from 0.1 to 100 Hz (slope 12 dB/octave) and the recording gain was set at 1000. Responses that were greater than 100 μ V were judged as noisy and rejected offline.

a. MMN. In the oddball paradigm, 2000 sweeps of the response to the standard and 300 sweeps of the response to the deviant stimulus were collected. By definition, the MMN is a response to stimulus change and occurs only when the deviant stimulus is presented in the context of a sequence of standard stimuli. A control condition was employed to ensure that the response reflects a change as opposed to reflecting simply the stimulus difference between the standard stimulus and the deviant stimulus. In this condition, the response which occurred to the deviant stimulus in the oddball paradigm was compared to the response evoked by the same stimulus when it was presented alone. If an MMN is present, then a relative negativity will be apparent only in the response elicited in the context of the oddball paradigm and not when the deviant stimulus was presented alone (Kraus *et al.*, 1995a).

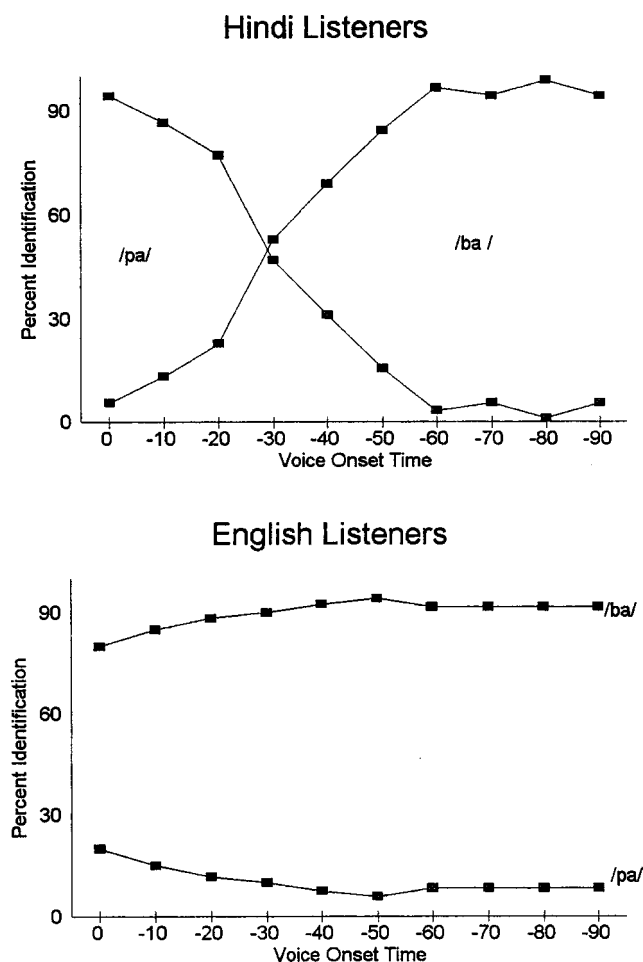


FIG. 2. Mean identification functions for the Hindi listeners (upper panel) and English listeners (lower panel).

b. N1. For the N1 recording, 300 sweeps elicited in response to each stimulus from the continuum were collected.

4. Data analysis

a. MMN. For individual subjects, sweeps were averaged to compute an individual average waveform elicited by the deviant stimulus when it occurred in the context of the standard stimulus (i.e., in the oddball paradigm) and by the de-

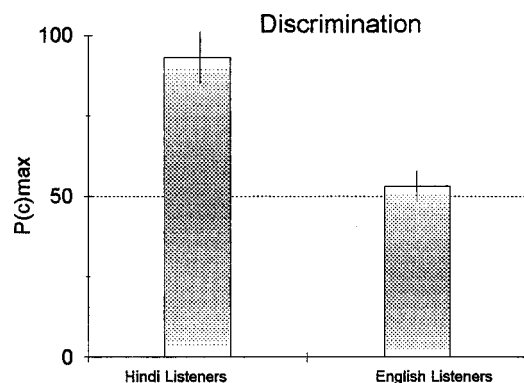


FIG. 3. Mean discrimination scores for the -10 and -50 ms VOT voicing contrast. Hindi subjects were significantly better at discriminating the stimulus pair as compared to English subjects ($p < 0.0001$). Error bars indicate ± 1 standard error.

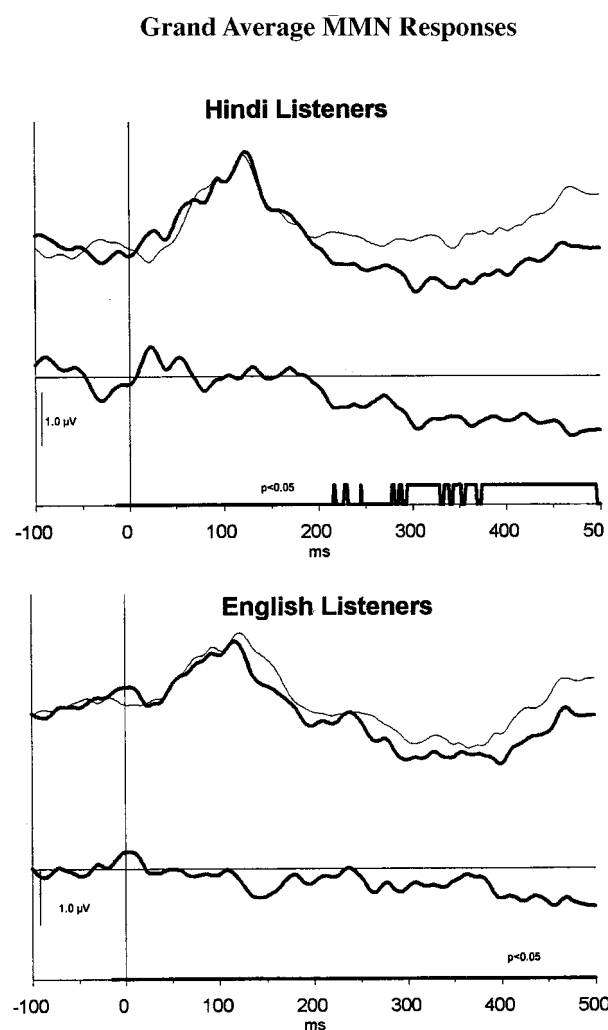


FIG. 4. Grand-averaged responses at the Fz scalp location for the Hindi listeners (left panel) and the English listeners (right panel). The top thin line is the response to the deviant stimulus when it was presented alone. The thick line is the response to the deviant stimulus when it occurred in the mismatch condition. The MMN response is seen in the difference wave (lower thick line) as a negativity. The boxes on the x-axis indicate the region of a significant mismatch response ($p < 0.05$).

viant stimulus when it was presented alone. A “difference wave” was computed by subtracting the response to the deviant stimulus from the deviant-alone stimulus presentation. In individual subjects, the morphologies of the alone, deviant, and difference waves were assessed relative to previously described morphologies of speech-evoked potentials (Kraus *et al.*, 1992). In individual subjects, the MMN was apparent in the difference waveform while the N1 was apparent in the alone and deviant waveforms. The MMN was identified visually as a relative negativity following the N1 peak. The point of maximum negativity of the MMN component was noted and the adjacent relative positive peaks were selected as the MMN onset and offset. To determine the area of the MMN, a line was drawn from the onset to the offset of the MMN in the difference wave. The area was defined as a summation of the point by point multiplication ($\text{ms} \times \mu\text{V}$) of the enclosed difference wave. The area of the MMN was computed for individual subjects from the Fz electrode since the most robust MMNs were observed at that electrode site.

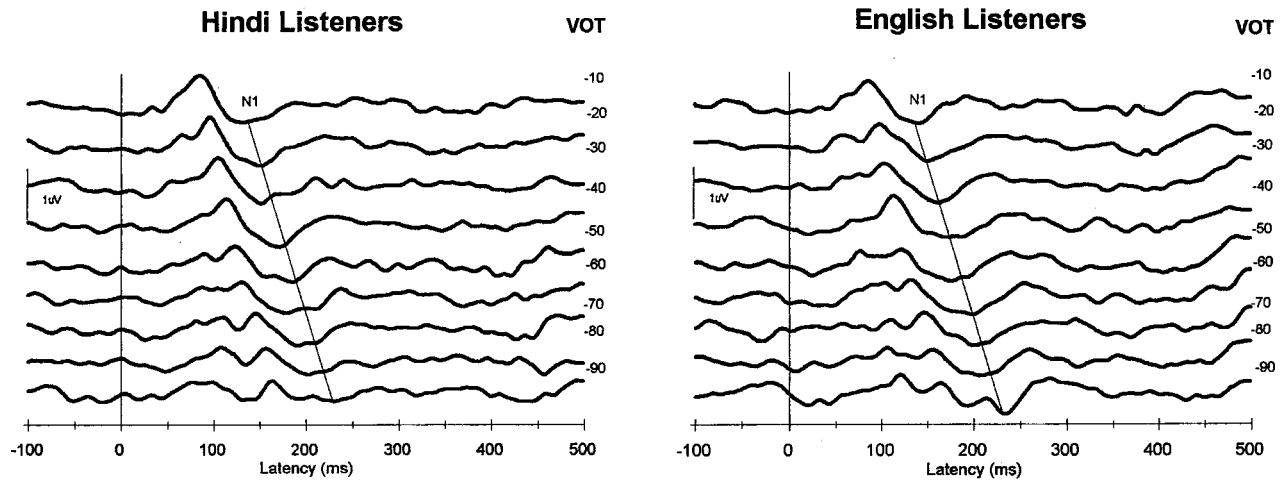


FIG. 5. Grand-averaged N1 responses for Hindi listeners (left panel) and English listeners (right panel). N1 latency was positively correlated with VOT for both groups of listeners [(Hindi group $r=0.79$; $p<0.0001$); (English group $r=0.8$; $p<0.0001$)].

Group averaged difference waves were computed by averaging individual subject difference waves. A point-to-point t -test of the values contributing to the group averaged difference waveform determined the period over which the grand averages were significantly different from zero at the $p < 0.05$ level. A significant negativity in the mean difference wave was defined as the group MMN. The disadvantage of this analysis is that the number of t -tests performed and the high correlation between adjacent points on a waveform increases chances of spurious significant values. To counter this difficulty, Guthrie and Buchwald (1991) based on multiple regression techniques on P300 waveforms, concluded that if a continuous interval of at least 12 sampling points shows significance then the power of the statistical test is sufficient. Because the A/D sampling rates and filter settings for the MMN are essentially similar to the P300, a similar significance interval should give appropriate statistical power (Kraus *et al.*, 1995a; McGee *et al.*, 1997).

b. N1. For individual subjects, sweeps were averaged to compute an individual average waveform. All waveforms were digitally high pass filtered off-line at 4 Hz (filter slope 12 dB/octave). In the group mean waveforms, N1 was identified visually as prominent negative peak within the first half of the time window. In order to aid in peak identification and measurement in data from individual subjects, response windows of ± 25 ms were created around the peak in the group mean waveforms. Peak latencies were detected based on the recordings from the electrode site Fz, where the response amplitudes were the largest in the group mean waveforms. The latency was typically marked at the center. Group averaged waveforms were computed by averaging across the individual average waveforms.

B. Results

1. MMN

The group averaged waveforms from the Fz electrode site are shown in Fig. 4. A robust and statistically significant MMN was obtained for Hindi listeners. However, a statistically significant MMN was not elicited in English listeners. Furthermore, a t -test revealed that the MMN area was sig-

nificantly larger for the Hindi compared to the English group ($t=2.3$; $df=18$; $p<0.05$). These results are consistent with those of Kraus *et al.* (1995b) and Tremblay *et al.* (1997) who reported that only minimal MMN responses were elicited to Hindi VOT contrasts in untrained English listeners.

To summarize, the native (Hindi) speakers who could behaviorally discriminate the stimulus contrast exhibited a robust MMN to the contrast, while in the nonnative (English) listeners whose perceptual discrimination of the same contrast was at a chance level, a significant MMN in response to the stimulus contrast was not observed.

2. N1

The grand average waveforms showing the variation in N1 latency across the stimulus continuum are shown in Fig. 5. A two way (VOT \times listener group) ANOVA revealed a significant main effect of VOT ($F=66.45$, $p>0.0001$) but not of listener group ($f=2.15$, $p>0.18$). Correlational analysis revealed that the N1 latency was positively correlated with VOT for both groups of listeners [(Hindi group $r=0.79$; $p<0.0001$); (English group $r=0.80$; $p<0.0001$)], suggesting that the latency of the N1 reflects the VOT of the syllable. Therefore, to summarize, N1 latencies systematically reflected the acoustic change from the pre-voiced portion to the voiced portion within the syllables equally well for *both* the English and Hindi groups.

III. GENERAL DISCUSSION

We have examined behavioral and neurophysiologic correlates of the perception of pre-voicing which provides a phonetically relevant change in Hindi but not in English. The results of the N1 experiment showed that the neurophysiologic representation of the duration of pre-voicing and voicing onset was equally robust in native and nonnative speakers. In contrast, a robust MMN was elicited only for native speakers of Hindi.

As seen in Fig. 5, changes in latency of the N1 component reflected the acoustic change from pre-voicing to voicing onset occurring within a syllable. This result is consistent

with previous reports based on animal studies (Stein-schnieder *et al.*, 1995; Eggermont, 1995; McGee *et al.*, 1996) and human experiments (Kaukaranta *et al.*, 1987; Ostroff *et al.*, 1998; Sharma and Dorman, 1999) that changes in morphology and latency of the AEP waveform reflect acoustic events ongoing within the speech stimulus. For example, Sharma and Dorman (1999) reported that changes in N1 latency and morphology reflected sequential changes in acoustic events within a /da/-/ta/ stimulus continuum such as the interval between burst and onset of voicing. Similarly, Kaukaranta *et al.* (1987) and Ostroff *et al.* (1999) reported that the N1 latency and morphology reflected the onset of friction relative to vowel onset in a fricative vowel. However, these studies did not specifically examine whether speech-elicited changes in the AEP waveform indicate that the subject can discriminate or accurately identify the speech sounds. The results of the present study sheds some light on this issue. In our stimulus continuum, the duration of pre-voicing was the acoustic feature which cued the change in phonetic categories (i.e., from /ba/ to /pa/) for Hindi listeners. For Hindi listeners then, as expected, the onset latency of the N1 response systematically reflected the change from pre-voicing to voicing onset across the stimulus continuum (Fig. 5). However, as can be seen in the same figure, changes in N1 latency also encoded the systematic changes in duration of pre-voicing in English listeners who were unable to behaviorally discriminate short versus long VOTs. This result suggests that neural encoding of acoustic changes ongoing in a speech signal as reflected by the N1 provides necessary but not sufficient information for behavioral discrimination of speech sounds.

In contrast to the N1 component, the MMN elicited in response to a stimulus contrast differed depending on whether that contrast constituted a phonemically relevant change in the listeners native language or not. In response to VOT differences which constituted phonetic categories in Hindi, but not in English, robust MMNs were elicited in Hindi listeners, but not in English listeners. These findings are consistent with those of Näätänen *et al.* (1997) and Winkler *et al.* (1999) who reported that the MMNs in response to nonnative vowel changes were reduced compared to MMNs elicited by native vowel contrasts in Finnish and Hungarian. The present results are further supported by preliminary results from Phillips *et al.* (1999) who reported that Japanese listeners did not demonstrate a significant MMN response to a /r/-/l/ contrast and by the results of Dehaene-Lambertz (1997) who reported that an MMN-like response (which was recorded in an active rather than a passive listening paradigm) was absent in response to a Hindi retroflex versus dental contrast which was perceptually discriminated at chance by French listeners. Future studies should systematically examine the changes in MMN that accompany re-learning of nonnative contrasts in adult second-language learners.

The different pattern of results obtained for the N1 and the MMN (i.e., similar encoding of the duration of pre-voicing for both Hindi and English listeners by the N1, and, the presence of an MMN for Hindi listeners but not for English listeners) suggests that the N1 and MMN components

of the AEP reflect functionally different levels of processing. The N1 reflects stimulus processing which occurs at a sensory level that is not modified by exposure to the phonetic categories of a language. The MMN reflects a level of processing in which language specific categories play a role. The characterization of the MMN as reflecting an entirely sensory level of processing is not tenable in light of the present results.

ACKNOWLEDGMENTS

We would like to thank the associate editor and two anonymous reviewers for their comments.

- Best, C. T. (1994). "The emergence of native-language phonological influences in infants: A perceptual assimilation model," in *The Development of Speech Perception*, edited by J. C. Goodman and H. C. Nusbaum (MIT Press, Cambridge, MA), pp. 167–224.
- Burnham, D. (1986). "Developmental loss of speech perception: Exposure to an experience with a first language," *Applied Psycholinguistics* **7**, 207–240.
- Csepe, V. (1995). "On the origin and development of the Mismatch Negativity," *Ear Hear.* **16**, 91–104.
- Dehaene-Lambertz, G. (1997). "Electrophysiological correlates of categorical phoneme perception in adults," *NeuroReport* **8**, 919–924.
- Eggermont, J. (1995). "Representation of a voice-onset time continuum in primary auditory cortex of the cat," *J. Acoust. Soc. Am.* **98**, 911–920.
- Guthrie, D., and Buchwald, J. (1991). "Significance of testing difference potentials," *Psychophysiology* **28**, 240–244.
- Kaukaranta, E., Hari, R., and Lounasmaa, O. (1987). "Responses of the human auditory cortex to vowel onset after fricative consonants," *Exp. Brain Res.* **69**, 19–23.
- Kraus, N., McGee, T., Carrell, T., and Sharma, A. (1995a). "Neurophysiologic bases of speech discrimination," *Ear Hear.* **16**, 19–37.
- Kraus, N., McGee, T., Carrell, T., King, C., Tremblay, K., and Nicol, T. (1995b). "Central auditory system plasticity associated with speech discrimination training," *J. Cogn. Neurosci.* **7**, 27–34.
- Kraus, N., McGee, T., Sharma, A., Carrell, T., and Nicol, T. (1992). "Mismatch negativity event-related potential elicited by speech stimuli," *Ear Hear.* **13**, 158–164.
- Logan, J. S., and Pruitt, J. S. (1995). "Methodological issues in training listeners to perceive nonnative phonemes," in *Speech Perception and Linguistic Experience: Issues in Cross-Language Research*, edited by W. Strange (York Press, Baltimore, MD), pp. 351–378.
- Maiste, A., Wiens, A., Hunt, M., Schery, M. and Picton, T. (1995). "Event-related potentials and categorical perception of speech sounds," *Ear Hear.* **16**, 67–89.
- McGee, T., Kraus, N., and Nicol, T. (1997). "Is there really a mismatch negativity? An assessment of methods for determining response validity in individual subjects," *Electroencephalogr. Clin. Neurophysiol.* **104**, 359–368.
- Näätänen, R. (1992). *Attention and Brain Function* (Erlbaum, Hillsdale, NJ).
- Näätänen, R., Lehtokoski, A., Lennes, M., Cheour, M., Huotilainen, M., Livonen, A., Vainio, M., Alku, P., Ilmoniemi, R., Luuk, A., Allik, J., Sinkkonen, J., and Alho, K. (1997). "Language-specific phoneme representations revealed by electric and magnetic brain responses," *Nature (London)* **385**, 432–434.
- Näätänen, R., and Picton, T. (1987). "The N1 wave of the human electric and magnetic response to sound: A review and analysis of the component structure," *Psychophysiology* **24**, 375–425.
- Ostroff, J. M., Martin, B. A., and Boothroyd, A. (1998). "Cortical evoked response to acoustic change within a syllable," *Ear Hear.* **19**, 290–297.
- Phillips, C., Marantz, A., McGinnis, M., Pesetsky, K., Wexler, E. and Yellin. (1999). "Brain mechanisms of speech perception: a preliminary report," MIT working papers in linguistics **26**, 513–191 (MIT Press, Cambridge, MA).
- Pisoni, D. B., Lively, S. E., and Logan, J. S. (1994). "Perceptual learning of nonnative speech contrasts: Implications for theories of speech perception," in *The Development of Speech Perception*, edited by J. C. Goodman and H. C. Nusbaum (MIT Press, Cambridge, MA), pp. 121–166.
- Polka, L. (1991). "Cross-language speech perception in adults: Phonemic, phonetic, and acoustic contributions," *J. Acoust. Soc. Am.* **89**, 2961–2976.

- Sams, M., Aulanko, R., Aaltonen, O., and Näätänen, R. (1990). "Event-related potentials to infrequent changes in synthesized phonetic stimuli," *J. Cogn. Neurosci.* **2**, 344–357.
- Sharma, A., and Dorman, M. F. (1999). "Cortical auditory evoked potential correlates of categorical perception of voice-onset-time," *J. Acoust. Soc. Am.* **106**, 1078–1083.
- Sharma, A., Kraus, N., McGee, T., Carrell, T., and Nicol, T. (1993). "Acoustic versus phonetic representation of speech as reflected by the mismatch negativity event-related potential," *Electroencephalog. Clin. Neurophysiol.* **88**, 64–71.
- Steinschneider, M., Schroeder, C., Arezzo, J., and Vaughan, Jr., H., (1995). "Physiological correlates of the voice onset time boundary in the primary auditory cortex (A1) of the awake monkey: Temporal response pattern," *Brain Lang.* **48**, 326–340.
- Tees, R. C., and Werker, J. F. (1984). "Perceptual flexibility: Maintenance of recovery of the ability to discriminate nonnative speech sounds," *Can. J. Psychol.* **38**, 579–590.
- Tremblay, K., Kraus, N., Carrell, T., and McGee, T. (1997) "Central auditory system plasticity: Generalization to novel stimuli following listening training," *J. Acoust. Soc. Am.* **102**, 3762–3773.
- Werker, J. F. (1994). "Cross-language speech perception: Developmental change does not involve loss," in *The Development of Speech Perception*, edited by J. C. Goodman and H. C. Nusbaum (MIT Press, Cambridge, MA), pp. 93–120.
- Winkler, I., Lehtoksoki, A., Alku, P., Vainio, M., Czigler, I., Csepe, V., Aaltonen, O., Raimo, I., Alho, K., Lang, H., Iivonen, A., and Näätänen, R. (1999). *Cognitive Brain Research* **7**, 357–369.

Speech recognition with reduced spectral cues as a function of age

Laurie S. Eisenberg,^{a)} Robert V. Shannon, Amy Schaefer Martinez, and John Wygonski
House Ear Institute, Los Angeles, California 90057

Arthur Boothroyd

Graduate School, City University of New York, Department of Speech and Hearing Sciences, New York, New York 10016

(Received 8 September 1999; accepted for publication 20 December 1999)

Adult listeners are able to recognize speech even under conditions of severe spectral degradation. To assess the developmental time course of this robust pattern recognition, speech recognition was measured in two groups of children (5–7 and 10–12 years of age) as a function of the degree of spectral resolution. Results were compared to recognition performance of adults listening to the same materials and conditions. The spectral detail was systematically manipulated using a noise-band vocoder in which filtered noise bands were modulated by the amplitude envelope from the same spectral bands in speech. Performance scores between adults and older children did not differ statistically, whereas scores by younger children were significantly lower; they required more spectral resolution to perform at the same level as adults and older children. Part of the deficit in younger children was due to their inability to utilize fully the sensory information, and part was due to their incomplete linguistic/cognitive development. The fact that young children cannot recognize spectrally degraded speech as well as adults suggests that a long learning period is required for robust acoustic pattern recognition. These findings have implications for the application of auditory sensory devices for young children with early-onset hearing loss. © 2000 Acoustical Society of America. [S0001-4966(99)00105-9]

PACS numbers: 43.71.Ft, 43.71.Ky, 43.66.Sr, 43.66.Ts [CWT]

INTRODUCTION

Speech recognition is a complex process that requires detailed spectro-temporal information from the sensory periphery as well as robust central pattern recognition. For such a complex cognitive task there is likely to be a tradeoff between the complexity of the listening conditions and the quality of the sensory information. Recent studies conducted with normally hearing adults have shown excellent speech recognition for sentences in quiet with as few as three to four spectral channels of auditory information (Dorman *et al.*, 1997; Shannon *et al.*, 1995). Significantly, asymptotic levels of performance have been achieved by profoundly hearing-impaired adults fitted with cochlear implants when listening to speech delivered through as few as four to six electrodes representing spectral channels (Dorman and Loizou, 1998; Fishman *et al.*, 1998). A greater number of channels is required to achieve the same level of recognition in noise in normally hearing and implanted adults (Dorman *et al.*, 1998a; Fu *et al.*, 1998), indicating that more detailed sensory information is necessary under more difficult listening conditions.

The spectral reduction technique originated as a model for the study of speech recognition in listeners with auditory prosthetic devices like cochlear implants (Shannon *et al.*, 1995). To implement this technique, the time-intensity envelope is extracted from bandpass filtered speech and used to modulate contiguous noise bands. This type of processing

has been shown to preserve the normal temporal variations of speech while systematically manipulating spectral information (Dorman *et al.*, 1997; Rosen, 1992; Shannon *et al.*, 1995; Turner *et al.*, 1995; Van Tassel *et al.*, 1987). In addition to investigating the number of spectral channels required to understand speech, this minimalist representation of the speech signal may also prove to be extremely useful as a tool in assessing the developmental robustness of speech pattern recognition in children.

Infants are born with adult-like cochlear function but acquire competency in their native language and speech recognition ability over years. Investigations into the developmental time course of speech recognition have demonstrated systematic improvements in performance as children mature to adolescence (Boothroyd, 1968; Elliott *et al.*, 1979; Fior, 1972; Hnath-Chisolm *et al.*, 1998; Siegenthaler, 1969). The underlying sources that explain these improvements are far from clear; they have been attributed to such factors as growth of vocabulary or increase in phonemic categories, maturation in decision-making processes, or articulation improvements (Boothroyd, 1970). Even when vocabulary can be controlled, developmental differences in speech recognition remain evident (Elliott *et al.*, 1979). Anatomical studies demonstrate that the human auditory cortex continues to develop up until adolescence (Moore *et al.*, 1997), which may partially explain age-related improvements in speech recognition.

Inasmuch as auditory perceptual abilities mature over the first 10–12 years, central pattern recognition for speech may be much less robust to sensory distortion as that ob-

^{a)}Electronic mail: leisenberg@hei.org

served for adults (e.g., Elliott, 1979; Marshall *et al.*, 1979; Nagafuchi, 1976; Neuman and Hochberg, 1983). With the exception of preliminary data reported by Dorman *et al.* (1998b), little effort has been extended to the investigation of number of spectral channels required by children to understand speech. This initial report suggested that normally hearing children as young as 3 and 4 years of age require a greater number of spectral channels than adults in order to understand a limited set of multisyllabic words appropriate for children of this age.

The present experiment investigates the developmental time course of pattern recognition for speech degraded in spectral resolution, using the noise-band processing technique described by Shannon *et al.* (1995). Specifically, we seek to determine the number of spectral channels required by children to recognize speech accurately and the extent to which cognitive and linguistic factors influence these findings.

I. METHODS

A. Subjects

Performance was measured on two groups of English-speaking, normally hearing children (5–7 and 10–12 years) and one group of adults. Inclusion into the study required that at least one ear demonstrate audiometric thresholds of 15 dB HL (ANSI, 1996) or better for the octave intervals between 250 and 8000 Hz. Tympanometry was performed to ensure normal middle-ear status.

Sixteen children, 5–7 years of age, were assigned to two groups: group A consisted of ten children tested on 4-, 6-, and 8-band processors; group B consisted of six children tested on 8-, 16-, and 32-band processors. Thus, no child received a repetition of the stimulus material. The ages of the younger children (nine females and seven males) ranged from 5 years, 9 months to 7 years, 7 months (mean=6 years, 5 months). Ten children, 10–12 years of age, were tested on 4, 6, and 8 bands only. The ages of the older children (six females and four males) ranged from 10 years, 4 months to 12 years, 11 months (mean=11 years, 5 months). Twenty-three of the 26 children exhibited receptive vocabulary age equivalents, as measured by the Peabody Picture Vocabulary Test—Third Edition (Dunn and Dunn, 1997), that were within 2 standard deviations of chronological age; the remaining three children were within 3 standard deviations. Children's results on the experimental measures were compared to those of ten normally hearing adult females between the ages of 21 and 55 years (mean=29 years) for the 4-, 6-, and 8-noise-band conditions.

B. Test materials

Age-appropriate stimulus materials including sentences, words, nonsense syllables, and digits were spectrally manipulated using the noise-band processing technique. The Hearing In Noise Test for Children, or HINT-C (Gelnett *et al.*, 1995), comprised the sentence material. Recorded at the House Ear Institute by a male talker, the HINT-C is composed of 130 sentences that are identifiable by normally hearing children as young as 5 and 6 years of age. These 130

sentences were derived from the original 250 HINT sentences (Nilsson *et al.*, 1994). Six 10-sentences lists were selected at random from the 13 lists, resulting in 20 sentences per noise-band condition. Three additional ten-sentence lists were selected for practice, resulting in one practice list per noise-band condition.

The word stimuli selected for this study were the Phonetically Balanced Word Lists—Kindergarten (PBK, Haskins, 1949). Derived from the spoken vocabulary of kindergarten-age children, these monosyllabic word lists are widely used in testing children (Kirk *et al.*, 1997). Scores of 96%–98% would be expected for normally hearing children 5 to 7 years of age listening under optimal (i.e., unprocessed) conditions (Sanderson-Leepa and Rintelmann, 1976). Three 50-item word lists used in the study were obtained from the Auditec of St. Louis compact disk. Presented with carrier phrase, the words were spoken by a male talker. Five words from each list were selected for practice; thus, 45 words remained for each experimental condition.

The Video Speech Pattern Contrast Test (VIDSPAC, Boothroyd, 1997) was selected to assess speech feature contrasts. The test was originally developed for use with children as young as 3 years of age (Boothroyd, 1991). Using a video game format, the VIDSPAC requires the subject to indicate when a phonetic change is detected during a series of repeated syllables (e.g., ti, ti, ti, ti, tu). The stimulus tokens, spoken by a woman talker, include nonsense syllables that contrast vowel place and height (ti vs tu, ti vs ta), consonant manner (da vs za, sa vs ta), consonant voicing (da vs ta, sa vs za), and consonant place (da vs ga, sa vs sha). An auditory/visual subtest is included for practice, contrasting vowel place (ti vs tu).

The digit span subtest from the Wechsler Intelligent Scale for Children (WISC III, Wechsler, 1991) was selected to test immediate recall for an auditory stimulus. Comparing digit span performance between unprocessed and noise-band processed conditions may elucidate cognitive effects on the ability to recall a sequence of numbers under degraded spectral conditions. To administer the test, a sequence of digits is presented and the listener is instructed to repeat the digits in forward or backward order. The number of digits is systematically increased (e.g., 2–9, 3–8–6, 3–4–1–7, etc.) until the listener is unable to repeat two sequences accurately. For the present study, the stimulus items were recorded by a woman talker and stored as digital files. The digit span test was administered in the unprocessed and 8-band processed conditions. Prior to testing, subjects were required to repeat each digit to ensure that the stimulus items were identifiable.

C. Noise-band processing and instrumentation

Spectral manipulation of the speech signal was accomplished digitally using filtered noise bands modulated by the speech amplitude envelope from the same spectral band (Shannon *et al.*, 1995). As noted earlier, this type of processing has been shown to preserve the temporal cues of speech while systematically reducing the spectral cues. To accomplish the noise-band processing, the acoustic signal was digitized at a sampling rate of 22 KHz with 16-bit resolution and input to a pre-emphasis filter followed by a set of 4-, 6-, 8-,

TABLE I. Crossover frequencies for the 4-, 6-, 8-, 16-, and 32-noise-band conditions.

4-band:	300	722	1528	3066	6000										
6-band:	300	550	936	1528	2440	3842	6000								
8-band:	300	477	722	1061	1528	2174	3066	4298	6000						
16-band:	300	382	477	590	722	878	1061	1276	1528	1825	2174	2584	3066	3632	4298
	5080	6000													
32-band:	300	339	382	427	477	531	590	653	722	797	878	966	1061	1164	1276
	1397	1528	1671	1825	1992	2174	2371	2584	2815	3066	3338	3632	3952		
	4298	4673	5080	5522	6000										

16-, and 32-bandpass filters spanning a frequency range of 300 to 6000 Hz. The frequency response of the pre-emphasis filter was uniform above 1200 Hz, with a -6 -dB/octave roll-off below 1200 Hz. This filter was implemented as a first-order Butterworth infinite impulse response (IIR) filter. The frequency spacing of the bank of filters was computed from the formula derived by Greenwood (1990) relating frequency to position along the basilar membrane. Assuming a total processor bandwidth of 300–6000 Hz, the distance corresponding to the difference, 5700 Hz, was divided by the number of processor channels determined for investigation. The crossover frequencies for the processed conditions are shown in Table I.

Each of the bandpass filters was implemented using a cascade of second-order elliptic IIR filters. The output from each filter was input to an envelope detector consisting of a full-wave rectifier followed by a low-pass filter (160 Hz). The extracted envelopes modulated a pseudorandom noise, which was input to corresponding bandpass filters with the same frequency bandwidths as those for the original filters. The bands were summed and the signal converted to analog form and played out through a professional audio quality PC sound card (Turtle Beach Fiji Board). The processed signal was input to an audiometer (Grason Stadler 16) for amplification, and presented monaurally at 70 dBA through an ER-2 insert earphone. Because the presentation levels varied as a function of noise-band condition and test material, the rms level of each set of processed materials was specified in a DB-100 occluded ear simulator.

D. Procedures

Each subject was tested in a single-walled, sound-attenuating booth. With the exception of two children tested in the left ear (due to either a flat tympanogram or mild hearing loss in the right ear), all other subjects were tested in the right ear.

The noise-band conditions were presented in random order as were the test measures within a condition (HINT-C, PBK, and VIDSPAC) and the stimulus items within a test (HINT-C and PBK only). Order of VIDPAC subtests was not changed during each administration. Because pilot testing on the VIDSPAC revealed learning effects for younger children, the test was first administered in the unprocessed condition to the children only. The forward/backward digit span test conditions (unprocessed and 8-band processed) were counterbalanced across subjects.

For the HINT-C, PBK, and digit span, the subjects were instructed to repeat the stimulus item (or series of items) and

encouraged to guess. For the VIDSPAC, the younger children were instructed to raise their hand whenever a contrast was perceived. The older children and adults were permitted to respond via computer mouse. A tester was seated in the booth with the subject throughout testing both to instruct and to record responses.

Testing was divided into two, 2-h sessions for the children and one, 2 1/2-h session for the adults. Rest periods were provided as needed between noise-band conditions. All subjects were compensated for their participation in the experiment.

II. RESULTS

A. Measures of speech recognition

Figure 1 displays the mean percent-correct scores for the four groups as a function of the number of noise bands. Mean sentence scores for the HINT-C (20 sentences/condition) improved as the number of noise bands increased up to 8 noise bands. Performance by older children and adults was similar, with mean scores of 93% and 94%, respectively, for 8 noise bands. The younger children (group A) demonstrated lower scores than adults and older children, recognizing only 82% of sentences correctly in the 8-band condition. Even for 32 noise bands, only two of the six younger children (group B) attained scores similar to those of adults and older children in the 8-band condition, indicating that sentence scores generally did not improve after 8 bands for the younger children. When analyzed by number of words correct within a sentence, the older children and adults approached asymptote by 6 bands, with a mean word score of 96%. The younger group (group A) achieved a mean word score of 94% by 8 bands. Word scores were at ceiling for 16 and 32 noise bands (group B).

Statistical analysis of the arc-sine transformed sentence scores for the 4-, 6-, and 8-band conditions revealed main effects for number of noise bands, $F(2,54)=105.97$, $p<0.001$, and age group, $F(2,27)=671.51$, $p<0.001$, with no significant interaction. Pairwise comparisons (using a Bonferroni adjustment) revealed significant differences between all noise-band conditions, and the adults and older children achieved significantly higher sentence scores than the younger children, $p<0.05$.

Mean percent-correct scores for the PBK words (45 words/condition) improved for each group as the number of noise bands increased. None of the groups, however, exhibited a mean score that exceeded 90% by 8 bands. With 16 and 32 noise bands, scores by group B approached or exceeded those of adults and older children for the 8-band con-

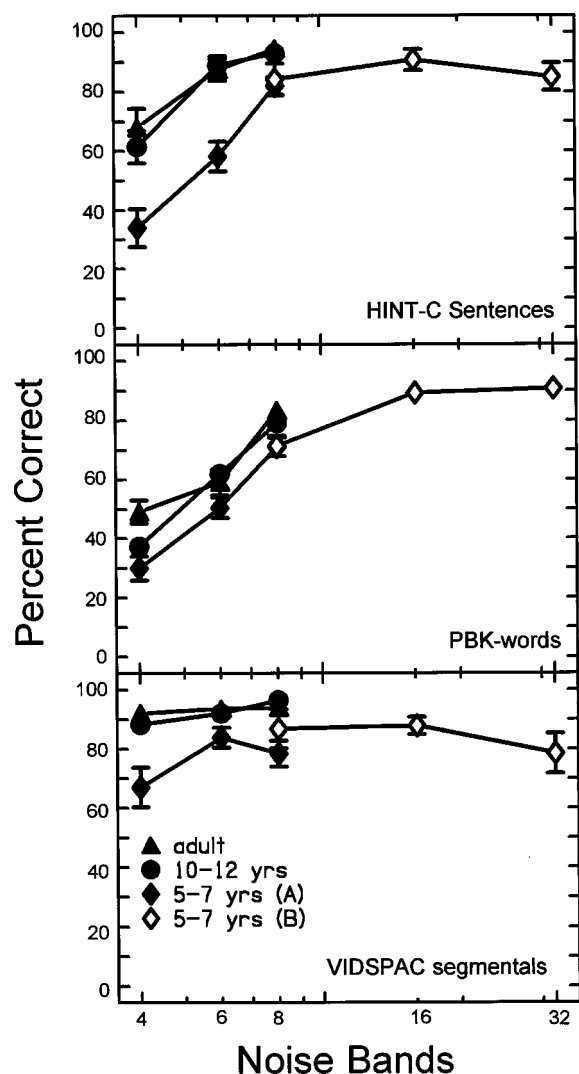


FIG. 1. Percent-correct scores for HINT-C sentences, PBK words, and VIDSPAC segmentals for three age groups across all noise-band conditions. The bars represent 1 standard error of the mean.

dition, implying that the adults and older children would have achieved higher scores had they been tested with a greater number of spectral channels. When analyzed according to the number of phonemes correct within each word, the adults and older children produced mean phoneme scores of 92.6% and 90.6%, respectively, in the 8-band condition only. The younger children (group A) achieved a mean phoneme score of 85.3% (range=78% to 96%) in the 8-band condition. Phoneme scores were 92% or higher for group B tested in the 16- and 32-band conditions.

For the 4-, 6-, and 8-band conditions, main effects from word scores were evident for number of noise bands, $F(2,54)=157.4$, $p<0.001$, and age group, $F(2,27)=9.68$, $p=0.001$, with no significant interaction. Pairwise comparisons (with Bonferroni adjustment) indicated that the mean word score for the 4-band condition was significantly lower than that for the 6-band, and both were significantly lower than that for the 8-band condition, $p<0.05$. The younger children yielded a significantly lower mean word score relative to the older children and adults, $p<0.05$.

The VIDSPAC segmental score represented a composite

of eight subtests: two vowel and six consonant contrasts (including manner, voicing, and place). Near-maximum-level performance ($>92\%$) was achieved with as few as 4 bands for the adults and 6 bands for the older children. The younger children (group A) did not approach this level of performance, even with 8 noise bands, exhibiting mean scores of 67%, 84%, and 78% in the 4-, 6-, and 8-band conditions, respectively; only two children achieved scores of 90% or higher for 8 bands or higher (group B).

Statistical analysis for the 4, 6, and 8 noise bands indicated main effects for noise-band condition, $F(2.54)=7.38$, $p=0.001$, and age group, $F(2,27)=10.13$, $p<0.001$, with a significant interaction between factors, $F(4,54)=2.64$, $p<0.05$. The younger children (group A) yielded significantly lower scores than the older children and adults on the 4- and 8-band conditions only, $p<0.05$. The younger children were found to be more variable in their performance than were adults and older children, suggesting cognitive or task-related factors. Given the number of times the test was administered through the course of the study, it is probable that these results were affected in part by fatigue or lack of attention.

Inasmuch as the results were influenced by the different features of speech, descriptive analysis was conducted on the vowels and consonants, including manner, voicing, and place contrasts. The high performance scores achieved by the adults and older children were weighted primarily by vowels and consonant manner, with the older children and adults obtaining mean scores that ranged from 95% to 100% across 4-, 6-, and 8-band conditions. The younger children achieved mean scores of 90% or higher by the 6-band condition for vowels and 89% or higher for consonant manner across all noise-band conditions. For consonant place, the older children and adults obtained mean scores that exceeded 90% by the 6-band condition; the younger children did not reach this level of performance for any of the noise-band conditions. The consonant voicing contrast was clearly the most troublesome for all subjects. Although one of the two voicing contrasts was relatively easy to discriminate (da vs ta), the other voicing contrast (sa vs za) was extremely difficult for all subjects, which adversely affected performance for this feature.

B. Context effects: Computation of the j factor

Boothroyd has developed a method of quantifying context effects in language perception by comparing recognition probabilities for parts and wholes (Boothroyd, 1985; Boothroyd and Nittrouer, 1988). This analysis results in a j factor, representing the effective number of independent parts in a whole (i.e., phonemes in a word or words in a sentence). When j equals the actual number of parts in the whole, it may be assumed that the recognition of one part (i.e., one phoneme in a word or one word in a sentence) has no influence on the probability of recognition of other parts. The value of j falls as the listener uses the recognition of one part to aid in the recognition of other parts. In the extreme, j approaches unity, at which point the recognition of one part is enough to guarantee recognition of the whole. Figure 2 illustrates application of the j -factor analysis to the current

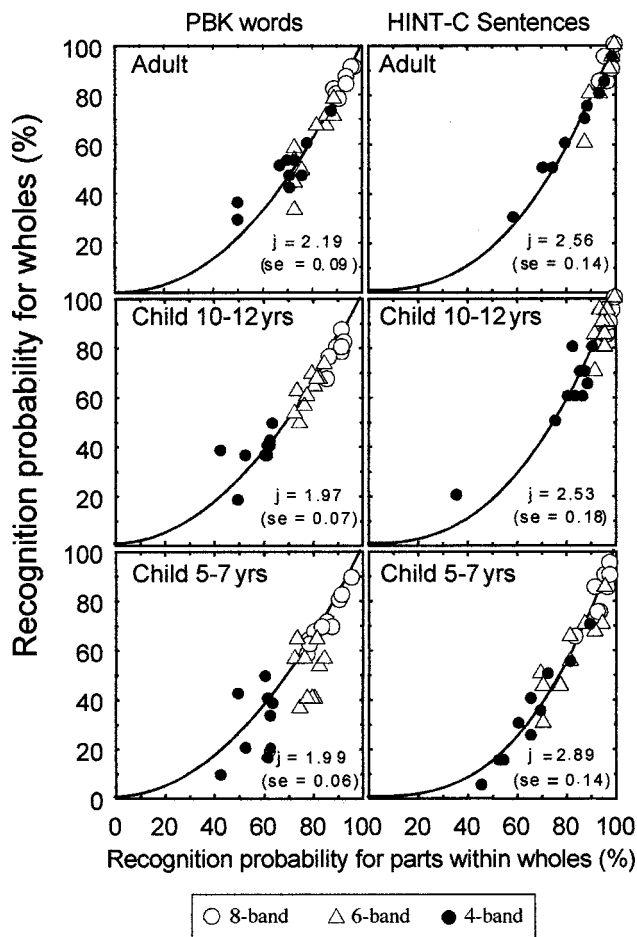


FIG. 2. Recognition probability for wholes (words and sentences) as a function of recognition probability for constituent parts (phonemes in words and words in sentences) for the 4-, 6-, and 8-band conditions for each age group. On the left are recognition probability data for the PBK words as a function of PBK phonemes. On the right are recognition probability data for the HINT-C sentences as a function of HINT-C words. Included in each graph is the j -factor computation and the standard error of the mean.

data for phoneme, word, and sentence recognition. For PBK word stimuli, the recognition probability of wholes (words) is plotted as a function of the recognition probability of the parts (phonemes in words). For HINT-C sentence stimuli, the recognition probability of wholes (sentences) is plotted as a function of the recognition probability of the parts (words in sentences). Values of j were obtained for each group using least-squares fits. For the PBK stimuli, there was no evidence of significant group differences of the j factor, implying that the three groups were taking equal advantage of word context to aid in phoneme recognition. For the HINT-C stimuli, however, the 5–7-year-olds (group A) gave a higher j factor than the other two older groups, implying that younger children were taking less advantage of sentence context to aid in word recognition. These findings support the conclusion that immaturity of both acoustic and linguistic factors contributed to the greater effect of sensory degradation found in the 5–7-year-olds.

C. Digit span

Mean digit span scores for the adults, older children, and younger children (group A) are displayed in Fig. 3 as a func-

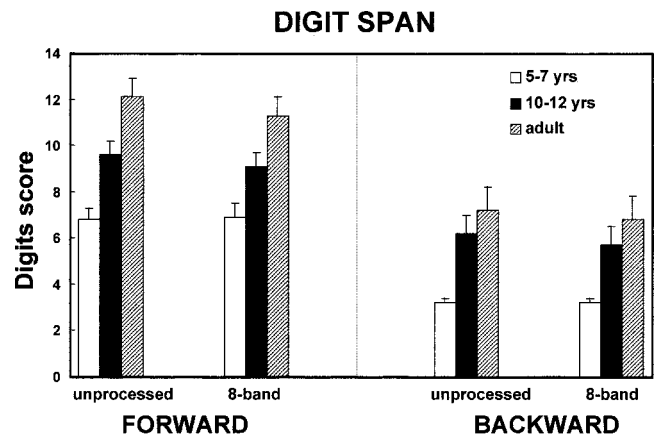


FIG. 3. Mean digit span scores for the forward and backward tests in the unprocessed and 8-band processed conditions for the three age groups. The bars represent 1 standard error of the mean.

tion of processing (unprocessed vs 8-band processed) and test (forward versus backward). Higher mean scores were observed for forward digit span ($M=9.3$) compared with backward ($M=5.9$), and performance improved with age (younger children $M=5.0$, older children $M=7.7$, and adults $M=9.4$). Mean scores in the unprocessed condition ($M=7.5$) were slightly higher than those in the 8-band processed condition ($M=7.2$). These findings were generally confirmed by statistical analysis, indicating main effects for test, $F(1,27) = 97.82$, $p < 0.001$, age group, $F(2,27) = 13.5$, $p < 0.001$, and processing, $F(1,27) = 5.63$, $p < 0.05$. There were no significant interactions. Pairwise comparisons using a Bonferroni criterion indicated significantly lower scores for the younger children when compared to those of the older children and adults ($p < 0.05$).

Pearson product moment correlation coefficients were computed between total score on digit span (unprocessed and 8-band processed) and scores on the HINT-C sentences, PBK words, and VIDSPAC (4-band and 8-band processed). In general, most coefficients were low ($r < 0.30$) when controlled for age, indicating a weak relationship between measures. Coefficients were also computed between total score on digit span (unprocessed and 8-band processed) and language quotient (receptive vocabulary age divided by chronological age) for the children only. A statistically significant coefficient was found between digit span (in the 8-band processed condition) and language quotient, $r = 0.503$, $p < 0.05$, when controlled for age.

III. DISCUSSION

The results from this investigation demonstrated that younger children were significantly less accurate than older children and adults in recognizing speech reduced in spectral content. With the exception of PBK word scores, the older children and adults obtained scores that exceeded 90% with 4 to 8 modulated noise bands, depending on the measure. Such findings corresponded with results from other studies conducted on adults using different stimulus materials. The younger children either approached or exceeded adult-like performance for PBK words with 16 noise bands but did not

reach this level for HINT-C sentences or VIDSPAC segmentals, even with 32 noise bands. These results were generally in agreement with earlier findings that documented the developmental effects of speech reduced in spectral content (Dorman *et al.*, 1998b). Analytical treatment of context effects demonstrated that younger children were less skilled than adults or older children in utilizing context to recognize words on the HINT-C sentences, although this age difference was not observed for recognition of phonemes on the PBK words. The lack of evidence of an age effect on the use of context in the recognition of PBK words may be due, in part, to the fact that a kindergarten-level vocabulary was used. Had a more adult vocabulary been used, it is reasonable to assume that age-related differences of the j factor would have been found. In essence, performance by younger children was influenced by sensory as well as linguistic factors, dependent upon the stimulus materials under investigation.

Noise-band processing adversely affected immediate recall of a sequence of numbers, verifying that degraded sensory information acts as a cognitive load on the speech recognition process (Cleary, 1997). Correlation coefficients relating digit span to speech recognition performance were found to be weak ($r < 0.30$), a finding not consistent with results shown for children with cochlear implants. Pisoni and Geers (1998) reported moderate to strong correlation coefficients between digit span performance and speech recognition scores among a large sample of pediatric implant users, suggesting that working memory may help explain individual differences in speech perception for this population. Further study is warranted to better understand the discrepancy in findings between the two studies.

The statistically significant correlation found between digit span in the 8-band processed condition and language quotient (receptive vocabulary) is compatible with the phonological loop model promoted by Baddeley *et al.* (1998), which suggests an association between working memory and receptive language. According to this model, working memory serves as the link between the initial sensory input and the stored knowledge that listeners have about their language. Because working memory is finite, there is a limited capacity to process sensory and perceptual information. Difficulty in rehearsing and maintaining the phonological representations of speech in working memory may adversely affect vocabulary acquisition. Consequently, it is not surprising that the ability to recall a sequence of digits under spectrally degraded conditions was related to word knowledge in the present study.

The results from this investigation indicate that speech pattern recognition is still maturing in children 5 to 7 years of age. In contrast to adults and older children who are able to recognize speech with remarkable accuracy when the information conveyed is primarily temporal, young children are less robust in their ability to recognize this type of signal. Similar long-term developmental learning effects have been observed in other complex pattern recognition tasks, such as size constancy in vision (Liebowitz, 1971).

The findings from this study have implications for the application of auditory sensory devices, such as hearing aids or cochlear implants, for young children with early-onset

hearing loss. An immediate aim in providing an auditory sensory aid to young children with hearing loss is to ensure access to the acoustic-phonetic patterns of speech that will form the basis for further phonological and linguistic development. Children with early-onset hearing loss in the severe to profound range are particularly at a disadvantage, having been deprived of significant auditory input in the first months (and sometimes years) of life. Given that such children will still receive an impoverished signal once a sensory aid is fitted, optimization of device design and processing characteristics remains an important goal for this population. If normally hearing children require more than 7 years to develop robust auditory pattern recognition, it should be expected that a similar or longer training time would be required in children who receive imperfect sensory input and are delayed in auditory perceptual development. Even older children with more fully developed cognitive/linguistic capabilities may require many years of auditory experience with a sensory aid before they can integrate the sensory, cognitive, and linguistic information to achieve robust pattern recognition for the acoustic patterns of speech.

ACKNOWLEDGMENTS

This research was supported in part by research Grants Nos. DC03668 and DC01526 from NIDCD-NIH. The authors wish to thank Qian-Jie Fu and Mark Robert for technical support and Karen Iler Kirk for providing insightful comments about working memory.

- ANSI (1996). ANSI S3.6-1996, "Specifications for audiometers" (American National Standards Institute, New York).
- Baddeley, A., Gathercole, S. E., and Papagno, C. (1998). "The phonological loop as a language learning device," *Psychol. Rev.* **105**, 158–173.
- Boothroyd, A. (1968). "Developments in speech audiometry," *Sound* **3**, 3–10.
- Boothroyd, A. (1970). "Developmental factors in speech recognition," *Int. Audiol.* **98**, 30–38.
- Boothroyd, A. (1985). "Evaluation of speech production in the hearing-impaired: Some benefits of forced-choice testing," *J. Speech Hear. Res.* **28**, 185–196.
- Boothroyd, A. (1991). "Speech perception measures and their role in the evaluation of hearing aid performance," in *Pediatric Amplification*, edited by J. A. Feigin and P. G. Stelmachowicz (Boys Town National Research Hospital, Omaha), pp. 77–92.
- Boothroyd, A. (1997). "A video game test of speech pattern contrast perception," Graduate Center, City University of New York.
- Boothroyd, A., and Nittrover, S. (1988). "Mathematical treatment of context effects in phoneme and word recognition," *J. Acoust. Soc. Am.* **84**, 101–114.
- Cleary, M. (1997). "Measures of phonological memory span for sounds differing in discriminability: Some preliminary findings," *Research on Spoken Language Processing Progress Report No. 21* (Indiana University, Bloomington), pp. 93–140.
- Dorman, M. F., and Loizou, P. C. (1998). "The identification of consonants and vowels by cochlear implant patients using a 6-channel continuous interleaved sampling processor and by normally hearing subjects using simulations of processors with two to nine channels," *Ear Hear.* **19**, 162–166.
- Dorman, M. F., Loizou, P. C., Fitzke, J., and Tu, Z. (1998a). "The recognition of sentences in noise by normally hearing listeners using simulations of cochlear-implant signal processing," *J. Acoust. Soc. Am.* **104**, 3583–3585.
- Dorman, M. F., Loizou, P. C., Kirk, K. I., and Svirsky, M. (1998b). "Channels, children and the Multisyllabic Lexical Neighborhood Test (MLNT)," NIH Neural Prosthesis Workshop, Bethesda, MD, October, 1998.

- Dorman, M. F., Loizou, P. C., and Rainey, D. (1997). "Speech intelligibility as a function of the number of channels of stimulation for signal processors using sine-wave and noise-band outputs," *J. Acoust. Soc. Am.* **102**, 2403–2411.
- Dunn, L. M., and Dunn, L. M. (1997). "*Peabody Picture Vocabulary Test—Third Edition* (American Guidance Service, Circle Pines, MN).
- Elliott, L. L. (1979). "Performance of children aged 9 to 17 years on a test of speech intelligibility in noise using sentence material with controlled word predictability," *J. Acoust. Soc. Am.* **66**, 651–653.
- Elliott, L. L., Conners, S., Kille, E., Levin, S., Ball, K., and Katz, D. (1979). "Children's understanding of monosyllabic nouns in quiet and in noise," *J. Acoust. Soc. Am.* **66**, 12–21.
- Fior, R. (1972). "Physiological maturation of auditory function between 3 and 13 years of age," *Audiol.* **11**, 317–321.
- Fishman, K., Shannon, R. V., Slattery, W. A. (1997). "Speech recognition as a function of the number of electrodes used in the SPEAK cochlear implant speech processor," *J. Speech Lang. Hear. Res.* **40**, 1201–1215.
- Fu, Q.-J., Shannon, R. V., and Wang, X. (1998). "Effects of noise and spectral resolution on vowel and consonant recognition: Acoustic and electric hearing," *J. Acoust. Soc. Am.* **104**, 3586–3596.
- Gelnett, D., Sumida, A., Nilsson, M., and Soli, S. D. (1995). "Development of the Hearing In Noise Test for Children (HINT-C)," Presented at the American Academy of Audiology, Dallas, Texas, April 1995.
- Greenwood, D. D. (1990). "A cochlear frequency-position function for several species—29 years later," *J. Acoust. Soc. Am.* **87**, 2592–2605.
- Haskins, H. (1949). "A phonetically balanced test of speech discrimination for children," Unpublished master's thesis, Northwestern University, Evanston, IL.
- Hnath-Chisolm, T. E., Laipply, E., and Boothroyd, A. (1998). "Age-related changes on a children's test of sensory-level speech perception capacity," *J. Speech Lang. Hear. Res.* **41**, 94–106.
- Kirk, K. I., Diefendorf, A. O., Pisoni, D. B., and Robbins, A. M. (1997). "Assessing speech perception in children," in *Audiologic Evaluation and Management and Speech Perception Assessment*, edited by L. L. Mendel and J. L. Danhauer (Singular, San Diego), pp. 101–132.
- Liebowitz, H. (1971). "Sensory, learned, and cognitive mechanisms of size perception," *Ann. (N.Y.) Acad. Sci.* **188**, 47–62.
- Marshall, L., Brandt, J. F., Marston, L. E., and Ruder, K. (1979). "Changes in number and types of errors on repetition of acoustically distorted sentences as a function of age in normal children," *J. Am. Aud. Soc.* **4**, 218–225.
- Moore, J. K., Guan, Y.-L., and Wu, B. J. (1997). "Maturation of human auditory cortex: Laminar cytoarchitecture and axonal ingrowth," 20th ARO Midwinter Meeting, 28 (abstract), St. Petersburg, FL (Association for Research in Otolaryngology, Des Moines) p. 28.
- Nagafuchi, M. (1976). "Intelligibility of distorted speech sounds shifted in frequency and time in normal children," *Audiol.* **15**, 326–337.
- Neuman, A. C., and Hochberg, I. (1983). "Children's perception of speech in reverberation," *J. Acoust. Soc. Am.* **73**, 2145–2149.
- Nilsson, M., Soli, S. D., and Sullivan, J. A. (1994). "Development of the Hearing in Noise Test for the measurement of speech reception thresholds in quiet and in noise," *J. Acoust. Soc. Am.* **95**, 1085–1099.
- Pisoni, D. B., and Geers, A. (1998). "Working memory in deaf children with cochlear implants: Correlations between digit span and measures of spoken language processing," Seventh Symposium on Cochlear Implants in Children, Iowa City, IA, June 1998.
- Rosen, S. (1992). "Temporal information in speech and its relevance for cochlear implants," *Philos. Trans. R. Soc. London, Ser. B* **336**, 336–373.
- Sanderson-Leepa, M. E., and Rintelmann, W. F. (1976). "Articulation functions and test-retest performance of normal-hearing children on three speech discrimination tests: WIPI, PBK-50, and NU Auditory Test No. 6," *J. Speech Hear. Disord.* **41**, 503–519.
- Shannon, R. V., Zeng, F.-G., Wygonski, J., Kamath, V., and Ekelid, M. (1995). "Speech recognition with primarily temporal cues," *Science* **270**, 303–304.
- Siegenthaler, B. M. (1969). "Maturation of hearing abilities in children," *Int. Audiol.* **8**, 59–71.
- Turner, C. W., Souza, P. E., and Forget, L. N. (1995). "Use of temporal envelope cues in speech recognition by normal and hearing-impaired listeners," *J. Acoust. Soc. Am.* **97**, 2568–2576.
- Van Tassel, D. J., Soli, S. D., Kirby, V. M., and Widin, G. P. (1987). "Speech waveform envelope cues for consonant recognition," *J. Acoust. Soc. Am.* **82**, 1152–1161.
- Wechsler, D. (1991). "WISC-III Manual" (The Psychological Corporation, San Antonio).

An investigation of current models of second language speech perception: The case of Japanese adults' perception of English consonants

Susan G. Guion^{a)}

The University of Oregon, Department of Linguistics, Eugene, Oregon 97403-1290

James E. Flege

The University of Alabama at Birmingham, Department of Rehabilitation Sciences, VH 503, Birmingham, Alabama 35294

Reiko Akahane-Yamada and Jesica C. Pruitt

ATR Human Information Processing Research Labs, Soraku-gun, Kyoto, 619-0288 Japan

(Received 18 June 1999; accepted for publication 5 January 2000)

This study reports the results of two experiments with native speakers of Japanese. In experiment 1, near-monolingual Japanese listeners participated in a cross-language mapping experiment in which they identified English and Japanese consonants in terms of a Japanese category, then rated the identifications for goodness-of-fit to that Japanese category. Experiment 2 used the same set of stimuli in a categorial discrimination test. Three groups of Japanese speakers varying in English-language experience, and one group of native English speakers participated. Contrast pairs composed of two English consonants, two Japanese consonants, and one English and one Japanese consonant were tested. The results indicated that the perceived phonetic distance of second language (L2) consonants from the closest first language (L1) consonant predicted the discrimination of L2 sounds. In addition, this study investigated the role of experience in learning sounds in a second language. Some of the consonant contrasts tested showed evidence of learning (i.e., significantly higher scores for the experienced than the relatively inexperienced Japanese groups). The perceived phonetic distance of L1 and L2 sounds was found to predict learning effects in discrimination of L1 and L2 sounds in some cases. The results are discussed in terms of models of cross-language speech perception and L2 phonetic learning. © 2000 Acoustical Society of America.

[S0001-4966(00)02404-8]

PACS numbers: 43.71.Hw, 43.71.Es, 43.71.An [JMH]

INTRODUCTION

A great deal of recent research has examined the perception of vowels and consonants (or “sounds,” for short) in a second language. In many but not all instances, adults who learn a second language (L2) perceive L2 sounds differently than monolingual native speakers of the target L2 do. The general aim of this study was to provide insight into the perception of English consonants by native speakers of Japanese differing in English-language experience. The more specific aim of the study was to address questions relating to a model of cross-language perception, the perceptual assimilation model (PAM) developed by Best and her colleagues (e.g., Best, 1995), and a model of L2 speech acquisition, the speech learning model (SLM) developed by Flege and his colleagues (e.g., Flege, 1995).

A. Previous research

Many studies have examined Japanese learners' acquisition of English /ɹ/ and /l/. Adult Japanese learners have great difficulty in distinguishing /ɹ/ from /l/ (Goto, 1971; Miyawaki *et al.*, 1975; MacKain, Best, and Strange, 1981; Mochizuki, 1981). This is because Japanese does not have an

/ɹ/–/l/ contrast, and neither English liquid is similar phonetically to any Japanese consonant. At a more abstract phonological level, the single liquid consonant found in Japanese might be considered similar to both English /ɹ/ and /l/. However, reports on the Japanese liquid (often referred to as an “r”) suggest that it is usually produced as an apico-alveolar tap /ɾ/ (Miyawaki *et al.*, 1975; Vance, 1987). In some instances it is produced with a lateral or retroflex articulation, although these variants do not occur in any predictable phonological context.¹

The perceptual relationship between English /ɹ/ and /l/ and the Japanese /ɾ/ is uncertain. However, it appears that English /l/ is perceived as being phonetically more similar to Japanese /ɾ/ than English /ɹ/ is. In a cross-language mapping study, Takagi (1993) found that inexperienced Japanese listeners identified English word-initial /ɹ/ and /l/ tokens as instances of Japanese /ɾ/. However, in a rating experiment, the English /l/ tokens were judged to be more similar to Japanese /ɾ/ than the English /ɹ/ tokens were. Sekiyama and Tohkura (1993) found that Japanese listeners identified word-initial tokens of English /ɹ/ most often as English /ɹ/ but also as Japanese /ɾ/, /ɰ/ (an unrounded velar approximant), and /g/. Conversely, English listeners identified syllable-initial tokens of Japanese /ɾ/ most often as /l/. Best and Strange (1992) suggested that both English /ɹ/ and /l/ may be identi-

^{a)}Electronic mail: guion@oregon.uoregon.edu

fied as a poor exemplar of Japanese /u/ because English liquids are approximants, and because the closest Japanese consonant, in articulatory terms, is /u/. Support for this suggestion was obtained in studies by Yamada and Tohkura (1992) and Mochizuki (1981). Japanese listeners identified the middle stimuli in synthetic /ɪ-to-/l/ continua as “w.” However, the synthetic stimuli did not contain all the information found in natural stimuli. Therefore, the “w” identifications do not guarantee that naturally occurring English /ɪ/ and /l/ tokens will be heard as Japanese /u/.

The difficulty adult Japanese learners of English have in identifying English /ɪ/ and /l/ is well-known. However, there is evidence that Japanese learners’ discrimination of /ɪ/ and /l/ may improve as they gain experience in English. MacKain *et al.* (1981) found that a group of Japanese participants with little exposure to native-produced English showed near-chance performance on /ɪ/ and /l/ identification and discrimination, whereas Japanese participants with intensive English conversational training showed categorical perception of /ɪ/ and /l/. Flege, Takagi, and Mann (1996) found that Japanese adults who had been in the United States for an average of 21 years identified /ɪ/ and /l/ tokens more accurately than a group of participants who had been in the United States for just 2 years. Best and Strange (1992) also found an effect of English-language experience on the perception of the English contrasts /ɪ-/l/ and /ɪ-w/. These authors found a significant difference in identification and discrimination tasks between an inexperienced Japanese group who had been in the United States less than 7 months and had no conversational training and an experienced Japanese group who had been in the United States for 18–48 months and had English conversation training. The experienced group performed more like a native English group in labeling and discriminating English contrasts than the inexperienced group of Japanese participants.

Other studies have investigated the effect of laboratory training on /ɪ-/l/ identification and discrimination (Logan, Lively, and Pisoni, 1991; Lively, Logan, and Pisoni, 1993; Lively *et al.*, 1994; Bradlow *et al.*, 1997). These studies have shown that perceptual training using highly varied stimuli from multiple talkers in multiple phonetic environments, but lasting only a few weeks, can yield a small but significant improvement in /ɪ-/l/ identification.

Although most studies investigating Japanese learners of English have focused on /ɪ/ and /l/, a few studies have examined other consonants as well. Japanese learners of English are reported to produce English /θ/ as /s/ (Lado, 1957; Ritchie, 1968). There is also evidence that Japanese listeners often misidentify voiceless English fricatives. Lambacher *et al.* (1997) found that Japanese listeners had the most difficulty distinguishing between /θ/ and /s/. When presented with a syllable containing /θ/, 28% of the participants chose /s/, and when presented with a syllable containing /s/, nearly 25% chose /θ/. The number of /f-for-/θ/ responses was also quite high (13%). Yoshida and Hirasaka (1983) investigated the identification of minimal-pair contrasts between English /b-/v/, /ɪ-/l/, and /s-/θ/. Some pairs consisted of real words and some consisted of nonwords. The results obtained from 96 Japanese listeners indicated that, overall, the rate of

errors decreased from /ɪ-/l/ (27%) to /b-/v/ (23%) to /s-/θ/ (16%). An effect of lexical status was also observed in that more errors were found for consonants in nonwords than real words (see also Flege, Takagi, and Mann, 1996). As mentioned above, the English /w-/r/ contrast has also been investigated. Best and Strange (1992) found that Japanese listeners correctly labeled and discriminated English /w-/r/ at higher rates than /ɪ-/l/. Listeners with more English-language experience responded to both contrasts more like native English speakers than those with less experience. Recently, Pruitt *et al.* (1999) used a training procedure similar to that used by Lively *et al.* (1994) for /ɪ-/l/ perception to train other English consonants. They found a significant increase in pretest to post-test accuracy scores for the /b-/v/, /s-/θ/, and /z-/ð/ contrasts.

The results obtained in the studies just cited suggest that certain English consonant contrasts are more readily learned than others (either during the process of naturalistic acquisition or laboratory training). This raises the issue of why this might be. English /ɪ/ and /l/ appear to be especially difficult because these liquids do not map well onto any Japanese category(s) (although English /l/ may be phonetically more similar to the Japanese /r/ than English /ɪ/ is). The likelihood of hearing two English sounds in terms of one Japanese sound may be the cause of the difficulty Japanese learners of English have perceiving and producing English liquids. Contrasts involving other English consonants (e.g., /θ/, /w/, and /v/) may also pose a perceptual challenge for Japanese adult learners of English. Taken together, the literature reviewed here leads us to formulate the following questions: Why are certain English consonants more difficult for Japanese learners of English than others? Will experience with English affect the discrimination of various English consonants by Japanese adult learners differentially, and if so, why?

B. Theoretical models

Two existing models relate to the differential learnability of second language (L2) consonants. The speech learning model (SLM) developed by Flege (1995) and the perceptual assimilation model (PAM) developed by Best and colleagues (Best, McRoberts, and Sithole, 1988; Best, 1993, 1995) both model the degree of success listeners will have in perceiving non-native sounds. These models posit that success will depend on the perceived relationship between phonetic elements found in the first language (L1) and the L2 systems. The models make predictions about performance in non-native segmental perception based on the perceived phonetic distance between L1 and L2 sounds.

PAM starts with the observation that certain pairs of sounds from an unknown foreign language are easier to discriminate than other pairs are. In fact, certain foreign contrasts are easy to discriminate, even for listeners who have never heard them before. Other contrasts, on the other hand, are quite difficult to discriminate. PAM proposes that sounds in a foreign language are perceived according to their similarities to, or discrepancies from, native-language sounds that are closest articulatorily. PAM proposes that listeners will detect similarities and dissimilarities to native sounds based

on perceived articulatory properties (e.g., constriction locations, active articulators, constriction degree, and phasing). The perceived distance between the unknown foreign sounds and the closest L1 sound (if any) leads to differences in discriminability.

The SLM differs from PAM in that it focuses on L2 *learning*. The purpose of the SLM is to account for changes across the life span in L2 speech learning. Hypotheses of the SLM can generate predictions concerning the accuracy with which highly experienced learners will produce and perceive L2 sounds. The SLM hypothesizes that basic speech learning mechanisms, including the ability to establish long-term memory representations for speech sounds (“phonetic categories”), remain intact across the life span. It also hypothesizes that L2 learners can establish new L2 phonetic categories if they detect phonetic differences between an L2 sound and the nearest L1 sound. The SLM predicts that the greater the perceived phonetic distance between an L2 sound and the closest L1 sound is, the more likely it is that phonetic differences between the sounds will be detected and a phonetic category eventually established. The acquisition of phonetic categories is thought to make L2 segmental perception more native-like because it enables the learner to base perception on L2 phonetic input without interference from prior learning.

The extent to which foreign (or L2) sounds resemble sounds in the naive listener’s (or L2 learner’s) L1 phonetic inventory plays a crucial role in both PAM and the SLM. According to PAM the degree of perceived phonetic distance is based on the perceived resemblance of articulatory gestures used to produce the foreign phones being discriminated and those used to produce the closest L1 sound. The SLM does not take a specific position regarding how cross-language phonetic distance is gauged by speakers of an L2 (see, e.g., Flege, 1995, p. 264). However, both models agree that perceived phonetic distance must be assessed empirically through cross-language mapping experiments. For predictions to be generated by either PAM or the SLM, cross-language phonetic distance data are needed. Degree of phonetic distance has been examined using an identification and rating methodology (see, e.g., Schmidt, 1996; Strange *et al.*, 1998). The foreign (or L2) sounds are first classified as instances of a phonetic category(s) in the listener’s L1, then rated for goodness-of-fit to the L1 category.

Unfortunately, very few studies thus far have provided the needed phonetic distance data *and also* obtained the relevant discrimination data. Two such studies examined vowels in an unknown foreign language; one other examined consonants. Best, Faber, and Levitt (1996) showed that cross-language mapping can predict vowel discrimination for unknown foreign languages. Native speakers of English used English keywords to classify vowels drawn from three unknown foreign languages. The same participants also participated in a categorial AXB discrimination experiment using the same foreign vowel stimuli. The study showed that the identification data predicted accuracy in discriminating the foreign vowel contrasts. Polka (1995) found that monolingual English listeners differed in their accuracy in discriminating German vowels. Discrimination accuracy was related

to the observed cross-language mapping pattern. Two vowels that were both considered to be good instances of the same L1 category received lower discrimination scores than two vowels identified in terms of a single L1 category but differing in goodness ratings did. In the case of consonants, Best (1990) obtained cross-language identification data for ejectives examined earlier in a discrimination study. Again, discrimination accuracy was related to the observed cross-language mapping pattern.

The results from the studies just cited provide important insight into how naive listeners perceive the vowels and consonants of an unknown foreign language. One important question is whether the relationship between cross-language mapping patterns and discrimination in an unknown foreign language will also apply to individuals who are learning a second language. As far as we know, no previous study has examined both perceived cross-language phonetic distance and the discrimination of a wide range of L2 consonants. Therefore, the present study addressed two questions pertaining to the models discussed earlier. The first question was whether the PAM framework can be extended to naturalistic L2 acquisition. The second was whether the SLM framework can be extended to the acquisition of an L2 by relatively inexperienced L2 learners.

C. Present study

The primary aim of the present study was to examine the relation between the perceived phonetic distance of L2 and L1 consonants and discrimination of those consonants. In order to address questions pertaining to PAM and the SLM (see above), cross-language mapping data were needed. Thus, in experiment 1, a group of Japanese listeners who had little experience with spoken English identified English and Japanese consonants in terms of Japanese categories, then rated the consonants for goodness-of-fit to that Japanese category. In experiment 2, a categorial discrimination experiment was carried out. Given that the SLM predicts learnability, three groups of Japanese learners of English who differed in English experience were recruited, as well as a native English control group. The stimuli examined in the cross-language mapping experiment were used again in the discrimination experiment, which tested English–English, English–Japanese, and Japanese–Japanese consonant contrasts.

The organization of the paper is as follows: The results from the cross-language mapping experiment will be presented first, then the results from the categorial discrimination test. Finally, the research questions pertaining to PAM and the SLM (see above) will be addressed.

I. EXPERIMENT 1

The purpose of this experiment was to assess the perceived relation between English and Japanese consonants. Native speakers of Japanese identified English and Japanese consonant stimuli in terms of Japanese consonant categories,

then rated the same stimuli for goodness-of-fit to the Japanese category. As an experimental control, they also identified and rated Japanese consonants.

A. Method

1. Speech materials

The speech materials were produced by eight male native speakers each of English and Japanese. The Japanese speakers (mean age=37 years) had been living in the United States for an average of 3.3 years, and were from a variety of cities in Japan (Tokyo, Fukuoka, Oita, Kobe, Chiba). The native English speakers (mean age=33 years) were from several places in the United States (Alabama, Ohio, South Carolina, Illinois, Virginia, Wisconsin, Georgia, Washington). These speakers produced English or Japanese consonants followed by /a/ in a carrier phrase (“Then I saw ____ there,” for English; “Korewa ____ desu,” meaning “This is ____” for Japanese). The speakers produced each phrase at a relatively slow speech rate, then at a faster rate.

The following method was used to obtain two different speech rates: First, the speakers listened to a tape that modeled the task. They heard isolated consonant + /a/ combinations (/Ca/) then another voice saying, “Then I saw /Ca/ there” (or “Korewa/Ca/desu”). Three such examples were given. The rate modeled was fairly slow and careful. After practicing, the speakers were given a written list of /Ca/ syllables to repeat in the carrier phrase following an auditory model. After the slow tokens were recorded, the faster tokens were collected. The speakers were asked to repeat the task, but this time they were asked to speak more rapidly. (They heard some examples of the task modeled at a faster rate, practiced, and then repeated the task.) The talkers were also encouraged to speak more rapidly by a smaller interstimulus interval (ISI) at the faster rate (1.6 vs 2.2 s in the earlier block). It was assumed that the increase in the speaking rate that was modeled would lead to a somewhat less careful speech style. Productions of most English and Japanese consonants (and vowels) were elicited in this way. However, only a subset of the consonant stimuli (English /b v w θ t s ɹ l/, Japanese /b ɯ t d s r h/) was used in the present study.

As expected, the duration of the stimuli varied as a function of the modeled speaking rate. The /a/ in the English tokens averaged 184 ms in the faster condition vs 237 ms in the slower condition. The /a/ in the Japanese stimuli averaged 86 and 129 ms in duration, respectively. The Japanese /a/ was consistently shorter than the English /a/ (mean = 108 vs 210 ms). To prevent vowel length from being used as a cue in consonant discrimination, all vowels were first normalized to 50% of peak intensity, then truncated to the same duration (50 ms for the faster rate tokens, and 75 ms

for the slower rate tokens). To minimize a “clipped” percept at the vowel ending, the intensity of the last 20 ms of the vowel ending was ramped off from 100% to 0%.

The following procedure was used to ensure that only good examples of the English and Japanese consonants were used as stimuli. Four native English speakers (two female, two male) from urban areas in the southern United States and four native Japanese speakers who had been in the United States for less than a year (two female, two male), all from the Tokyo area, judged consonants from their native language. The stimuli were presented simultaneously with a text stimulus (a key word for English, a *Katakana* transcription for Japanese). The listeners rated the stimuli for goodness on a scale ranging from “bad example” (1) to “very good example” (5). The best five (out of the eight) tokens, as judged by native speakers, were selected for use as stimuli. The ratings for both the Japanese and English stimuli were high, with means around the 4.0 level.

2. Participants

Nine native Japanese speakers living in Japan participated. All were college students (mean age=20.1 years) who had never lived outside of Japan. It is difficult to find normal young adults in Japan who have not been exposed to English. However, the participants selected for the present study had the minimum possible exposure to English. They began to study English at about the age of 12 ($M=11.7$ years) at school. Most of their exposure to English had taken place in the classroom. No participant reported a history of hearing or speech disorders, and all passed a pure-tone hearing screening at octave frequencies from 250 to 4000 Hz at 15 dB SPL in both ears. The participants were recruited in the same location (Kyoto) as the groups tested in the discrimination experiment (experiment 2, below).

3. Procedure

The English and Japanese speech stimuli (/Ca/ tokens) were presented to the participants for two kinds of auditory evaluation. They were first asked to identify each token as an instance of some Japanese consonant category. Then, immediately after, they were asked to rate the token for goodness-of-fit to the (just-selected) Japanese category. The opportunity to use all Japanese consonants might have overwhelmed the participants. Therefore, a set of likely consonant response categories was determined based on the results of a pilot experiment. The pilot used free transcription of the stimuli to determine which Japanese orthographic symbols should be presented to the participants as possible response alternatives. The choices of Japanese *Katakana* orthography (IPA representations given) were as follows:

ワ/ɰa/, ラ/ra/, ウラ/uɾa/, リヤ/rʲa/, ヤ/ja/, ハ/ha/, パ/pa/, バ/ba/, ヲア
/vaʲ/, ファ/ɸa/, タ/ta/, ダ/da/, サ/sa/, ザ/za/, シャ/ʃa/, ツァ/tsa/, ア/a/.

TABLE I. Mean percent identification and goodness rating (in parentheses) of Japanese consonant stimuli in terms of Japanese categories. Boldfaced values indicate the modal identification response. The goodness ratings are based on a scale that ranged from “bad example” (1) to “very good example” (7).

Consonant Stimuli	Percent Identification and Rating																
	ワ	ラ	ウラ	リャ	ヤ	ハ	ハ ^o	ハ ^u	ウ ^u ア	フア	タ	ダ	サ	ザ	シャ	ツア	ア
	u ^a	ra	ura	r ⁱ a	ja	ha	pa	ba	va	φa	ta	da	sa	za	ʃa	tʃa	a
/b/							8 (1.7)	81 (5.5)	11 (4.8)								
/u/	62 (4.2)	28 (4.0)	4 (1.8)			1 (2.0)		1 (1.0)	4 (2.8)								
/t/											99 (5.9)	1 (1.0)					
/d/												88 (5.1)		12 (3.7)			
/s/													85 (5.0)		1 (1.0)	14 (3.8)	
/r/		99 (4.8)			1 (3.0)												
/h/						100 (5.5)											

After identifying each stimulus using one of these labels, the participants rated the goodness-of-fit to the selected Japanese consonant category using a scale ranging from bad example (1) to very good example (7).

Each of the nine participants responded to 150 trials, 10 tokens (5 fast and 5 slow) each of the 15 consonant types (8 English and 7 Japanese), for a total of 1350 responses. The average identification and goodness-of-fit rating for each consonant type was based on 90 responses (9 participants × 10 tokens).

B. Results

Table I presents the results for the Japanese consonant stimuli. Two types of data are presented. The percentages indicate the frequency with which various Japanese consonant categories were used to classify the Japanese stimuli. The numbers in parentheses indicate the average ratings for the stimuli receiving a particular classification. The boldfaced percentages indicate the modal (i.e., most frequently used) classification. For example, the Japanese /b/ stimuli were classified as /b/ in 81% of instances (the modal classification), as /p/ in 8% of instances, and as /v/ in 11% of instances. The average ratings given to stimuli receiving the modal classification were always higher than the ratings obtained for stimuli receiving other classifications.

The Japanese consonant stimuli were not identified as intended in 100% of instances. However, the correct identification rates were high for most of the Japanese consonants, which indicates that the participants understood the task and could reliably perform it. The Japanese listeners identified the Japanese consonants correctly (i.e., as intended) 88% of the time, on average. The correctly identified Japanese consonants were given an average goodness rating of 5.1. As

shown in Table I, /t/, /r/, and /h/ were identified correctly at near-perfect rates. The consonants /b/, /d/, and /s/ were identified correctly in more than 80% of instances. The Japanese consonant correctly identified least often (62%) was /u/ (an unrounded velar approximate), which was frequently identified as /r/.

Table II presents the results for the English consonants. Unlike the Japanese consonants, there was not a “correct” classification for the English stimuli. However, the data in Table II allow us to determine which Japanese consonant category (or categories) would be used most often to classify the English consonant stimuli. Of the eight English consonants examined, five were consistently (>75%) classified as instances of a single Japanese consonant category. English /t/ was heard as Japanese /t/ in 91% of instances, English /s/ as Japanese /s/ in 87% of instances, English /b/ as Japanese /b/ in 84% of instances, English /v/ as Japanese /v/ in 80% of instances, and English /w/ was heard as Japanese /u/ in 79% of instances. These consonants received a mean goodness rating of 4.3 for the corresponding Japanese category.

The grouped data in Table II indicate that some English consonants, on the other hand, were identified in terms of two Japanese consonants. It was *not* the case that some participants consistently classified an English consonant type in terms of one Japanese category, while other participants consistently classified the same consonant type in terms of another Japanese category. A close inspection of the data revealed that individual listeners identified these consonants as examples of two Japanese sounds. That is, individual participants heard these consonants as *intermediate* between two Japanese categories. English /θ/ was heard as Japanese /s/ 39% of the time, and as /φ/ (a labial fricative) 38% of the time. English /ɹ/ was heard as Japanese /r/ 46% of the time

TABLE II. Mean percent identification and goodness rating (in parentheses) of English consonant stimuli in terms of Japanese categories. Boldfaced values indicate the modal identification response. The goodness ratings are based on a scale that ranged from "bad example" (1) to "very good example" (7).

Consonant Stimuli	Percent Identification and Rating															
	ワ wa	ラ ra	ウラ ura	リャ rya	ヤ ya	ハ ha	パ pa	バ ba	ヴァ va	ファ fa	タ ta	ダ da	サ sa	ザ za	シャ sha	ツァ tsa
/b/								84 (5.3)	16 (4.8)							
/v/	2 (4.5)						1 (6)	17 (4.2)	80 (4.4)							
/w/	79 (3.5)	10 (3.0)	6 (2.2)			2 (3.0)			3 (2.0)							
/θ/								2 (2.0)	2 (4.5)	38 (3.4)			39 (3.8)	7 (2.7)		12 (3.3)
/t/											91 (3.9)					8 (2.4)
/s/													87 (4.5)			13 (3.6)
/ɹ/		46 (3.4)	50 (3.3)	4 (1.2)												
/l/	7 (3.5)	50 (3.2)	37 (3.0)	1 (1.0)					5 (3.6)							

and as /uɹ/ (a high back unrounded vowel+tap) 50% of the time. Finally, English /l/ was heard as Japanese /ɹ/ 50% of the time and as /uɹ/ 37% of the time. As mentioned above, the /uɹ/ response alternative was determined by free transcription of the stimuli in the pilot experiment. The /uɹ/ before the /ɹ/ might be the result of the approximate production of English /ɹ/, which is more vowel-like than a tap. The identifications of the /θ/, /ɹ/, and /l/ stimuli received an average goodness rating of 3.3.

The responses to the English stimuli were analyzed further in terms of overall fit to a Japanese category in order to provide what will be called a "fit index." The fit index used here combined both the identification and the goodness-of-fit data into a single metric.³ The fit indexes calculated here will be used in Sec. III to investigate the relations between cross-language mapping and discrimination. For the five English consonants that were consistently classified (>75%) as one Japanese consonant (*viz.*, /b/, /v/, /w/, /t/, /s/), the modal response was considered. For the three English consonants that were classified in terms of two Japanese consonants (*viz.*, /θ/, /ɹ/, /l/), both Japanese classifications were considered. The proportion of classification as a particular Japanese consonant was weighted by the mean goodness rating for stimuli receiving that identification. For example, the fit index for English /b/ stimuli was obtained by multiplying the proportion of responses receiving the modal identification (0.84) by the goodness rating of that identification (5.3). This resulted in a fit index of 4.5 for English /b/ to Japanese /b/. To take another example, the English /θ/ stimuli were classified in terms of two Japanese categories, /s/ and /φ/. Fit indexes were calculated for both Japanese response categories. The

proportion of /s/ identifications (0.39) was multiplied by the mean rating for the /s/ response (3.8), giving a fit index of 1.5. The proportion of /φ/ identifications (0.38) was multiplied by the mean rating for the /φ/ response (3.4), giving a fit index of 1.3.

The fit indexes derived for the eight English consonants are shown in Table III. (As mentioned earlier, two fit indexes were derived for the three English consonants that were identified in terms of two different Japanese consonants.) The fit indexes spanned a wide range, from a low value of 1.3 (the fit of English /θ/ to Japanese /φ/) to a high of 4.5 (the fit of English /b/ to Japanese /b/). It seems reasonable to suppose that the English consonants with relatively high fit indexes would be readily accepted as instances of a Japanese consonant category, whereas those with relatively low fit indexes would be heard either as "foreign" or as distorted instances of a Japanese category.

An important empirical question is whether variation in the fit indexes just described is relevant to the discriminability of English consonants (see experiment 2 below). As a working hypothesis, we divided the English consonants into subclasses based on the fit indexes using a standard deviation (s.d.) criterion. The mean fit index obtained for the seven Japanese consonants was 4.5 (s.d. = 1.1). The English consonants receiving a fit index that fell within 1.0 s.d. of the mean fit index obtained for the Japanese consonants (not shown in Table III) were classified as "good" instances of a Japanese category. (Thus, a good fit index for an English consonant was considered to be 3.4 and over.) The good-

TABLE III. Fit indexes derived for English consonants in terms of Japanese categories. The fit index was derived from the proportion of identifications and goodness ratings (see the text). Only identifications that were more than 30% are included.

English consonant	Most common identification	Proportion of identifications	Goodness rating	Fit index	
/b/	/b/	0.84	5.3	4.5	good /b/
/s/	/s/	0.87	4.5	3.9	good /s/
/t/	/t/	0.91	3.9	3.5	good /t/
/v/	/v/	0.80	4.4	3.5	good /v/
/w/	/w/	0.79	3.5	2.8	fair /w/
/ɹ/	/wɹ/	0.50	3.3	1.7	poor /wɹ/
/ɹ/	/ɹ/	0.46	3.4	1.6	poor /ɹ/
/l/	/ɹ/	0.50	3.2	1.6	poor /ɹ/
/l/	/wɹ/	0.37	3.0	1.1	poor /wɹ/
/θ/	/s/	0.39	3.8	1.5	poor /s/
/θ/	/φ/	0.38	3.4	1.3	poor /φ/

fitting English consonants were /b/, /s/, /t/, and /v/. English consonants that received fit indexes within 2 s.d.'s of the Japanese mean (2.3 to 3.3) were considered to have a "fair" fit index. English /w/ was considered to have fair fit index to Japanese /w/. Three other English consonants, /ɹ/, /l/, and /θ/, received a fit index that was more than 2.0 s.d.'s below the mean fit index computed for the Japanese consonants (*viz.*, values < 2.3). These consonants have been designated the "poor"-fitting English consonants.⁴ These classifications will be used in later sections to evaluate the relation between perceived cross-language phonetic distance and the discrimination of L2 consonants.

II. EXPERIMENT 2

The purpose of this experiment was to examine the discrimination of word-initial consonants by native speakers of Japanese and English. The stimuli were the Japanese and English consonants used in the cross-language mapping experiment (experiment 1). These stimuli were presented in a categorial discrimination test examining three types of contrasts. As shown in Table IV, four contrasts between two English consonants (the "*E-E*" contrasts) were examined. Seven contrasts between an English and a Japanese consonant (the "*E-J*" contrasts) were examined. Finally, three contrasts between two Japanese consonants (the "*J-J*" contrasts) were examined. As discussed in the Introduction, there was evidence that the *E-E* contrasts /ɹ/-/l/, /s/-/θ/,

TABLE IV. Three kinds of consonant contrasts examined in experiment 2. The English consonant is listed first, the Japanese consonant second for all English-Japanese contrasts.

English-English	English-Japanese	Japanese-Japanese
/b/-/v/	/ɹ/-/ɹ/	/ɹ/-/d/
/ɹ/-/l/	/ɹ/-/w/	/s/-/h/
/s/-/θ/	/l/-/ɹ/	/s/-/d/
/ɹ/-/w/	/θ/-/s/	
	/v/-/b/	
	/t/-/t/	
	/b/-/b/	

and /ɹ/-/w/ would prove difficult for at least some of the native Japanese participants. As far as we know, no previous study has examined *E-J* contrasts.

Four groups of individuals participated. One group consisted of native speakers of English; the other three consisted of native speakers of Japanese with varying amounts of English-language experience. We had two general expectations concerning how participants in these groups would perform. The first expectation was that the native speakers of Japanese would discriminate some *J-J* contrasts more successfully than the native speakers of English, whereas the reverse would hold true for the *E-E* contrasts. The second expectation was that the Japanese participants who were relatively experienced in English would discriminate at least some of the *E-E* and *E-J* contrasts more successfully than those who were relatively inexperienced in English.

A. Method

1. Speech materials

The stimuli used in experiment 1 were used again here.

2. Participants

As summarized in Table V, 30 native speakers of Japanese varying in experience with English participated. Ten native Japanese speakers living in the United States comprised the "high-experience" group. These participants had

TABLE V. Characteristics of the four groups of ten participants in experiment 2.

Group	Age ^a	AOA ^b	LOR ^c
Native English	24.6 (3.1) ^d	...	24.6 (3.1)
High-experience Japanese	29.7 (5.1)	12.3 (0.9)	3.1 (1.4)
Mid-experience Japanese	28.5 (4.3)	11.8 (1.4)	...
Low-experience Japanese	19.5 (1.2)	12.5 (0.5)	...

^aAge=mean chronological age at the time of testing in years.

^bAOA=mean age of acquisition of English.

^cLOR=mean length of residence in an English-speaking country in years.

^dStandard deviations are in parentheses.

resided in the United States for an average of 3.1 years (range 1.8 to 5.5 years). Ten native Japanese speakers matched to the United States group for age and education made up the “mid-experience” group. The participants in this group had never lived outside of Japan, but used English often in their jobs. The “low-experience” group consisted of ten Japanese college students who had never lived outside of Japan. Most of their exposure to English had taken place in the classroom and consisted largely of written English. (In Japan, English education focuses on reading and writing, while conversational practice with a native English speaker is rare.) Ten monolingual native speakers of American English made up the comparison group.

There were five male and five female participants in each group. All of the native Japanese participants had begun learning English at about 12 years of age, and had studied English through middle school and high school. All of the participants had at least some college and many (in the high- and mid-experience groups) had advanced degrees. No participant reported a history of hearing or speech disorders; all passed a pure-tone hearing screening from 250 to 4000 Hz at 15 dB SPL in both ears. All of the participants were paid.

3. Procedure

A categorical discrimination test was used to assess consonant perception. The test used here is similar to an ABX or AXB discrimination test (Gottfried, 1984; Best *et al.*, 1988) in that listeners heard three stimuli per trial and were asked to pick the odd item. However, the test used here differed from traditional oddity tasks in that it incorporated “catch” trials consisting of three physically different tokens of the same consonant. This encouraged the participants to respond only to phonetically relevant differences, not to any auditorily detectable difference. To successfully discriminate consonants, the participants had to recognize the *categorical* identity of a set of physically different tokens of the same consonant category while ignoring acoustic/auditory differences among instances of the category, which were phonetically irrelevant to their categorical identity.

Each consonant contrast investigated was tested by eight catch trials (comprised of three physically different tokens of the same stimulus type) and by eight “different” trials (in which there was an odd item among the three stimuli). The odd item appeared equally in all three possible positions. There were four catch trials for each of the two consonants being contrasted. In addition, an instance of each consonant was the odd item four times each in the different trials. For example, to test the /ɪ/-/ɪ/ contrast, four /ɪ/-/ɪ/-/ɪ/ and four /ɪ/-/ɪ/-/ɪ/ catch trials were administered. Four different trials in which /ɪ/ was the odd item, and four trials in which /ɪ/ was the odd item, were also presented. All trials consisted of tokens spoken by three different speakers. The three tokens were played at a 1.2 s ISI. The participants could replay a trial as often as they wished, but could not change a response once given.

The participants were tested individually in a sound booth and heard the stimuli at a comfortable listening level over headphones. They were told that the three stimuli in each trial were always spoken by different talkers, and that

they should ignore differences in speakers’ voices as much as possible. The participants selected “1,” “2,” or “3” if they judged a stimulus in one of those three serial positions to be different from the other two stimuli. They selected “no” if all three examples were considered to be instances of the same consonant.

All 448 trials were presented in two blocks during a 1-h session. The rate at which the stimuli had been produced (i.e., fast or slow) was counterbalanced across the listeners. To begin, the participants were familiarized with the task using catch trials and different trials made up of stimuli drawn from the experiment (but which did not test any contrasts found in the experiment). The participants received feedback during the practice, but not during the experiment.

4. Analysis

A-prime (A') scores were calculated for each of the 14 consonant contrasts examined. These scores were derived from the proportion of “hits” (correct selection of the odd item in different trials) and “false alarms” (incorrect selections of an odd item in catch trials) obtained for each contrast, using the formula provided by Snodgrass, Levy-Berger, and Haydon (1985).⁵ The A' scores provide an unbiased measure of perceptual sensitivity by taking into account the responses to the different trials and the catch trials. An A' score of 1.0 indicates perfect discrimination of a contrast, and an A' score of 0.5 or lower indicates insensitivity to a contrast.

B. Results

A series of analyses revealed that discrimination of the faster-vs-slower-rate stimuli did not vary according to group. A rate (2) × group (4) × contrast (4) analysis of variance (ANOVA) examined A' scores obtained for the $E-E$ contrasts. Similar ANOVAs examined the $E-J$ and the $J-J$ contrasts. The overall effect of rate was not significant for any of the three contrast types (F -values = 0.6 to 0.9). Nor did the rate factor interact significantly with group in any of the two-way interactions (F -values = 0.9 to 2.2) or three-way interactions (F -values = 0.4 to 1.5). This indicated that stimulus rate had a similar (non-)effect for the native English group and the three Japanese groups. Accordingly, the decision was made to pool the results obtained for the two sets of stimuli. New A' scores were computed based on the 16 catch and 16 different trials available for each contrast. These new A' scores represent a robust measure of the participants’ perception of the contrasts investigated.

Figure 1 displays the results of the four $E-E$ consonant contrasts. There was a general trend for the native English group to receive the highest scores, followed by the high-experience Japanese, then the mid-experience Japanese, and finally the low-experience Japanese. In addition, the /v/-/b/ and /ɪ/-/w/ contrasts showed more between-group differences for the three Japanese groups than the other two $E-E$ contrasts.

The A' scores for the $E-E$ contrasts were submitted to a group (4) × contrast (4) ANOVA. The main effects of group, $F(3,26) = 25.9$, $p < 0.01$, and contrast, $F(3,108)$

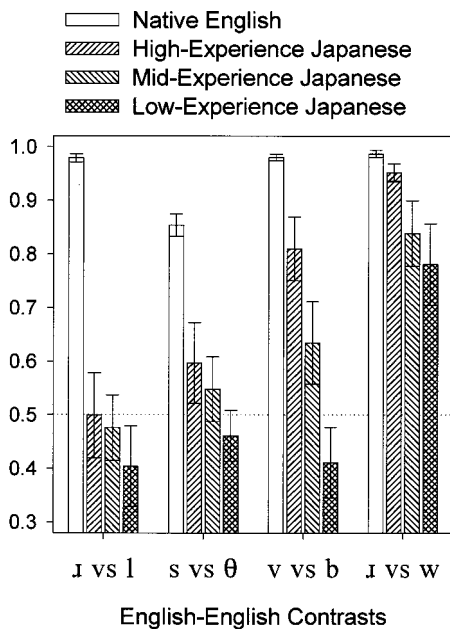


FIG. 1. Mean A-prime (A') scores obtained in experiment 2 for the four groups for four English–English contrasts. Error bars represent standard errors. A score of 1.0 indicates perfect discrimination, and a score of 0.5 or below indicates insensitivity to a contrast.

=35.4, $p < 0.01$, as well as the two-way interaction, $F(9,108) = 4.7$, $p < 0.01$, were significant. The simple effect of group was significant for all four contrasts [$/ɪ/-/w/$: $F(3,36) = 19.6$, $p < 0.01$; $/s/-/θ/$: $F(3,36) = 10.5$, $p < 0.01$; $/v/-/b/$: $F(3,36) = 19.2$, $p < 0.01$; and $/ɪ/-/w/$: $F(3,36) = 4.2$, $p = 0.01$]. A Tukey's test ($\alpha = 0.01$) revealed that the native English group received higher scores than all three Japanese groups for $/ɪ/-/ɪ/$ and $/s/-/θ/$. On these contrasts, none of the Japanese groups differed significantly from one another. For $/v/-/b/$, the native English and high-experience Japanese groups did not differ significantly from one another, but received higher scores than both the

mid- and low-experience Japanese groups. In turn, the mid-experience group received higher scores than the low-experience group. Finally, for $/ɪ/-/w/$, the native English group received significantly higher scores than the low-experience group. The other two Japanese groups did not differ from the native English listeners.

Figure 2 shows the results from the $E-J$ consonant contrasts. For three consonant contrasts, $/v/-/b/$, $/θ/-/s/$, $/ɪ/-/r/$, the more experienced Japanese groups tended to receive higher A' scores than the less experienced groups. Between-group differences for the remaining four consonant contrasts, $/b/-/b/$, $/l/-/r/$, $/t/-/t/$, $/ɪ/-/u/$, were smaller.

The A' scores for the $E-J$ contrasts were submitted to a group (4) \times contrast (7) ANOVA, which yielded significant main effects of group, $F(3,26) = 7.3$, $p < 0.01$, and contrast, $F(6,216) = 51.3$, $p < 0.01$, and a significant two-way interaction, $F(18,216) = 7.3$, $p < 0.01$. The simple effect of group was significant for three contrasts [$/θ/-/s/$: $F(3,36) = 9.1$, $p < 0.01$; $/v/-/b/$: $F(3,36) = 15.4$, $p < 0.01$; $/ɪ/-/r/$: $F(3,36) = 9.5$, $p < 0.01$], but not for four others ($/b/-/b/$, $/l/-/r/$, $/t/-/t/$, $/ɪ/-/u/$, $p > 0.01$). All four groups received scores of about 0.5 for $/b/-/b/$, $/l/-/r/$, and $/t/-/t/$ ($M = 0.43$, 0.54 , and 0.59 , respectively), indicating a lack of sensitivity to these contrasts. Relatively high scores ($M = 0.87$) were obtained for $/ɪ/-/u/$ from all four groups, indicating a high level of discriminability for this contrast.

A Tukey's test ($\alpha = 0.01$) revealed that the native English and high-experience Japanese groups did not differ significantly from one another for the $/v/-/b/$ contrast. However, both of these groups received significantly higher scores than the mid- and low-experience Japanese groups for $/v/-/b/$. The mid-experience group, in turn, received significantly higher scores than the low-experience group. The native English, high-experience, and mid-experience groups received significantly higher scores than the low-experience group for $/ɪ/-/r/$. For $/θ/-/s/$, the native English group received significantly higher scores than all three Japanese

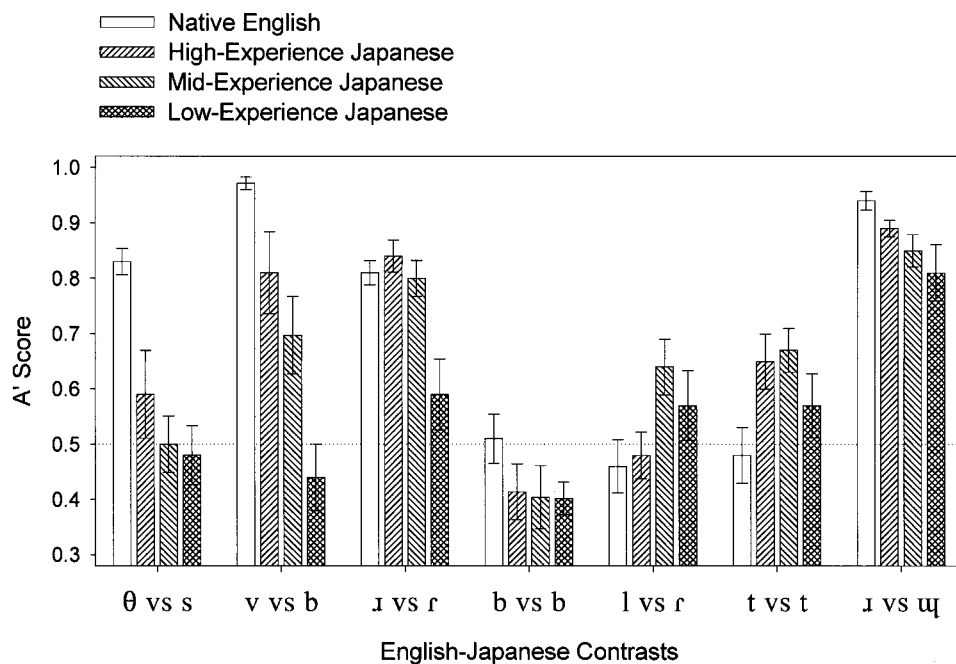


FIG. 2. Mean A-prime (A') scores obtained in experiment 2 for four groups for the eight English–Japanese contrasts. Error bars represent standard errors. A score of 1.0 indicates perfect discrimination, and a score of 0.5 or below indicates insensitivity to a contrast.

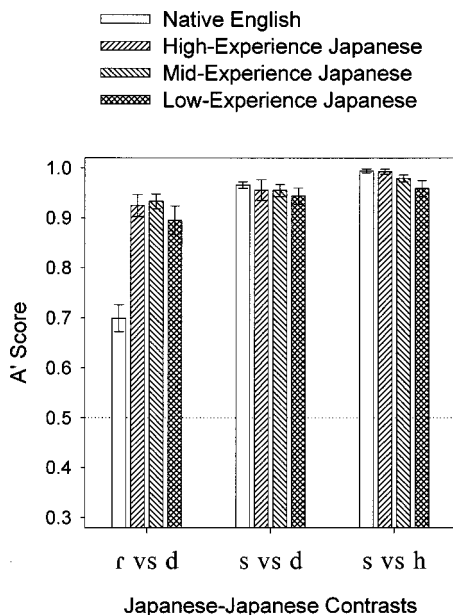


FIG. 3. Mean A' scores obtained in experiment 2 for four groups for the three Japanese-Japanese contrasts. Error bars represent standard errors. A score of 1.0 indicates perfect discrimination, and a score of 0.5 or below indicates insensitivity to a contrast.

groups, whereas the Japanese groups did not differ from one another.

Figure 3 shows the results of $J-J$ contrasts. All four groups received relatively high A' scores for the $/s-/d/$ and $/s-/h/$ contrasts. However, the native English group received lower scores for $/r-/d/$ than the three Japanese groups did.

The A' scores for the $J-J$ consonants were submitted to a group (4) \times contrast (3) ANOVA, which yielded significant main effects of group, $F(3,26) = 6.5$, $p < 0.01$, and contrast, $F(2,72) = 98.8$, $p < 0.01$, and a significant two-way interaction, $F(6,72) = 28.3$, $p < 0.01$. The simple effect of group was significant for $/r-/d/$, $F(3,36) = 22.5$, $p < 0.01$, but not for $/s-/h/$ or $/s-/d/$, $p > 0.01$. A Tukey's test ($\alpha = 0.01$) revealed that the native English listeners received significantly lower scores than all three native Japanese groups for $/r-/d/$. For $/s-/h/$ and $/s-/d/$, all four groups received high scores ($M = 0.98$ and 0.96 , respectively) and no significant between-group differences were obtained.

C. Discussion

The results obtained here suggest that some $E-E$ and $E-J$ consonant contrasts were more learnable than others. The Japanese participants with more English-language experience obtained higher discrimination (A') scores than the Japanese participants with less English experience for two $E-E$ contrasts ($/v-/b/$ and $/ɪ-/w/$), but not for two other $E-E$ contrasts ($/ɪ-/ɪ/$ and $/s-/θ/$). Of the eight $E-J$ contrasts examined, two ($/v-/b/$ and $/ɪ-/r/$) showed comparable effects of experience. It is interesting to note that the $/ɪ-/r/$ contrast showed improvement with experience, whereas the $/ɪ-/r/$ contrast did not. All four groups received low A' scores for $/ɪ-/r/$, suggesting that English $/ɪ/$ is closer to (and thus more difficult to discriminate from) Japanese $/r/$ than English $/ɪ/$.

All three native Japanese groups, and also the native English group, obtained very low A' scores for $/b-/b/$ and $/t-/t/$. This indicated a lack of sensitivity for these $E-J$ contrasts. However, the findings for the $J-J$ contrasts revealed that the Japanese participants could perform the discrimination task at high levels of accuracy. In fact, the Japanese participants obtained higher A' scores for some $J-J$ contrasts than the native English participants. This demonstrates that the lower A' scores obtained by the Japanese participants for some $E-E$ and $E-J$ contrasts was based on discrimination ability, not on test-taking ability or some type of cultural factor.

III. PREDICTION OF DISCRIMINABILITY

In this section we examine the relationship between the cross-language mapping data obtained in experiment 1 and the consonant discrimination scores from experiment 2. The purpose was to determine if PAM and the SLM can be extended to early stages of naturalistic L2 acquisition. (As mentioned in the Introduction, PAM usually focuses on the discrimination of sounds in an unknown foreign language; and the SLM usually focuses on highly experienced learners of an L2.)

A. Perceptual assimilation model

Best (Best *et al.*, 1988; Best, 1993, 1995) formulated the perceptual assimilation model (PAM) of cross-language speech perception. This model predicts the discriminability of non-native sounds based on their perceived relation to sounds in the L1. Best (1995, pp. 194–195) proposes the following patterns of perceptual assimilation for non-native sounds: They can either be (1) assimilated to a native category, (2) assimilated as uncategorizable speech sound, or (3) not assimilated to speech and heard as a nonspeech sound. If (1), then the non-native sound is clearly assimilated to a particular native segmental category, or perhaps to a cluster or string, in which case it may be heard either as a good exemplar of that category, an acceptable but not ideal exemplar of the category, or a notably deviant exemplar of the category. If (2), then the non-native sound is assimilated within native phonological space, but not as a clear exemplar of any particular native category (i.e., it falls within native phonological space, but in between specific native categories). If (3), then the non-native sound is not assimilated into native phonological space at all, but heard, instead, as some sort of nonspeech sound. PAM proposes that the patterns of perceptual assimilation will predict the discriminability of non-native consonants (and vowels).

Each of the assimilation patterns just mentioned makes a specific prediction regarding degree of discriminability. Based on the cross-language mapping data collected in experiment 1, the $E-E$ contrasts provide examples of the following assimilation pattern types: As shown in Table VI, the $/ɪ-/ɪ/$ contrast is an example of the “both uncategorizable” type. Both consonants fall within native phonological space, but “in between” (Best, 1995, p. 194) specific L1 categories. That is, both $/ɪ/$ and $/ɪ/$ fall in between Japanese $/ɯr/$ and $/r/$. The $/s-/θ/$ and $/ɪ-/w/$ contrasts are examples of the “uncategorized vs categorized” type. English $/s/$ was assimi-

TABLE VI. Categorization of English–English contrasts according to PAM.

Contrast	Fit index	PAM category type	Performance ^a		
			Low-experience	Mid-experience	High-experience
/ɪ/–/I/	poor /ɯr/ or /r/	both	poor	poor	poor
	poor /ɯr/ or /r/	uncategorizable			
/ɪ/–/w/	poor /ɯr/ or /r/	uncategorized vs	moderate	moderate	good
	fair /ɯ/	categorized			
/s/–/θ/	good /s/	uncategorized vs	poor	poor	poor
	poor /s/ or /φ/	categorized			

^aPerformance = A' scores obtained by the three Japanese groups (low-experience, mid-experience, and high-experience) on the English–English contrasts tested in experiment 2.

lated to Japanese /s/, and English /θ/ fell in between Japanese /s/ and /φ/. English /w/ was assimilated to Japanese /ɯ/, and English /ɪ/ fell in between Japanese /ɯr/ and /r/. The remaining *E–E* contrast, /b/–/v/ is more difficult to interpret in terms of the PAM typology. Both /b/ and /v/ were identified as good examples of Japanese categories, seeming to illustrate a case of “two-category assimilation.” However, the phonemic status of /v/ in Japanese is uncertain, rendering it questionable to interpret /v/’s perceptual assimilation pattern as a case of assimilation to a native category.

Given the cross-language mapping data obtained in experiment 1, PAM makes a number of testable predictions. These will be evaluated for the *E–E* contrast data obtained from the Japanese groups. Discrimination of the both uncategorizable type is expected to range from poor to very good, depending on the proximity of English consonants to each other and to Japanese categories. The /ɪ/–/I/ contrast is predicted to show poor discrimination because both sounds fall in between the same Japanese categories. This was indeed the case. All Japanese groups discriminated the /ɪ/–/I/ contrast at low rates.

Discrimination of the uncategorized vs categorized type was expected to be good. This prediction was borne out in the case of /ɪ/–/w/, which the Japanese participants discriminated with a moderate to high level of accuracy. It is interesting to note that the results of this type of contrast (uncategorized vs categorized), which is predicted to show good discrimination, improved with experience in the L2. In contrast, the results for the both uncategorized type of contrast, which is predicted to be poor, indicated that none of the three groups showed any sensitivity to the contrast.

The results for /s/–/θ/ were not as predicted. All three Japanese groups obtained low A' scores for this contrast. A closer look at the /s/–/θ/ contrast reveals some interesting facts. English /θ/ was classified as uncategorizable because it fell in between two native sounds. This is, it was heard as Japanese /s/ or as Japanese /φ/. English /θ/ was contrasted with English /s/ (heard as a good example of Japanese /s/), so there was likely to be overlap in category assimilation between the two English consonants. These results point to a possible revision of PAM in the predicted discrimination of uncategorized vs categorized non-native sounds. It might be useful to make provision for cases where the uncategorized sound is close in phonological space to the categorized sound.

Discrimination of the two-category assimilation type

was expected to be excellent. However, as outlined above, the /b/–/v/ contrast did not provide a good test of this prediction. This prediction has nonetheless been tested, and supported, by an investigation into English speakers’ assimilation patterns of Ethiopian ejectives /p’/ and /t’/ (Best, 1990). Native English participants identified the ejectives as examples of two English categories /p/ and /t/ (two-category assimilation) and discriminated the Ethiopian contrast at high rates.

In summary, the data from the cross-language mapping allowed us to test some predictions that might be derived for PAM if this model were extended to L2 learning. The predictions were supported with one exception. A contrast containing an uncategorized vs categorized sound was discriminated poorly, contrary to prediction. A revision of PAM was proposed to allow for poor discrimination of this contrast type when the uncategorized sound is in close phonological space to the categorized sound.

B. Speech learning model

The speech learning model (SLM) proposes that the likelihood of category formation for “sounds” is affected importantly by at least two factors. The likelihood of category formation is hypothesized to be inversely related to the age of first exposure to the sound system of an L2, but positively related to the perceived phonetic distance of an L2 sound from the closest native sound (Flege, 1995). This leads to the following two predictions for highly experienced speakers of an L2. The first is that, all else being equal (including degree of perceived phonetic distance), individuals learning an L2 early in life will be more likely to have established a category for an L2 consonant than those who began learning the L2 later in life. The native Japanese participants examined in this study were all exposed to English in school at about the same age (12 years old). Therefore, the first prediction, which addresses the effect of age of L2 learning, cannot be tested here.

A second prediction of the SLM is that, all else being equal (including the age of first exposure to the L2), L2 learners will be more likely to have developed a phonetic category for sounds that are perceptually distant from the closest native category than for sounds that are perceptually close to a native category. This prediction seems to hold true for highly experienced speakers of an L2. For example, Flege, Takagi, and Mann (1995, 1996) provided evidence

TABLE VII. Fit indexes derived for the English–Japanese contrasts (see the text).

English–Japanese contrast	Fit index ^a	Learning effect? ^b
/t/–/t/	/t/ 3.5 (good)	No
/b/–/b/	/b/ 4.5 (good)	No
/ɹ/–/ɹ/	/ɹ/ 1.6 (poor)	Yes
/θ/–/s/	/θ/ 1.5 (poor)	No
/l/–/r/	/l/ 1.6 (poor)	No

^aFit index=Fit index of the English consonant to the Japanese consonant to which it is being compared (see Sec. I B).

^bLearning effect?=results of pairwise comparisons for the Japanese groups' *A'* scores from experiment 1. A “yes” indicates that the high-and/or mid-experience groups received significantly higher *A'* scores than the low-experience group (see Sec. II B).

that category formation may require many years of native speaker input, at least for adult learners. Native Japanese participants showed evidence of category formation for /ɹ/ after 21 years in the United States, whereas those with two years of residence in the United States years did not. Thus, the results obtained for the relatively “experienced” group of native Japanese participants in the present study, who had lived in the United States for an average of 3.1 years, may not provide a fair test of the SLM’s prediction regarding the effect of perceived phonetic distance on phonetic category formation. However, if category formation is related to perceived phonetic distance from an L1 category, then one might expect to see some evidence of learning, even after a relatively limited exposure to the L2. Evidence for this would be more accurate perception by the relatively experienced than the relatively inexperienced Japanese learners of English examined here. In other words, learning effects should be more likely for contrasts between English and Japanese consonants if the English consonant is phonetically distant from the closest Japanese consonant than if the English consonant is similar to the closest Japanese consonant.

The SLM proposes that L2 learners must detect phonetic differences between L2 and L1 sounds before beginning to establish a new L2 category. This motivated a closer examination of the results obtained for the *E–J* contrasts. Our specific aim was to investigate a prediction that could be derived from the SLM if it were extended to relatively early stages of L2 acquisition. The prediction is that amount of L2 perceptual learning will depend on the perceived similarity of an English (L2) sound to the closest Japanese (L1) sound.

Table VII summarizes the predictions generated in this way for five *E–J* contrasts. The table shows the fit index for the English member of each English–Japanese contrast, along with its adjectival classification (see experiment 1). The prediction is that, when the English consonant represents a good fit to the Japanese consonant it was contrasted with, there will be little, if any, evidence of learning. However, evidence of learning should be seen for contrasts in which the English member of the contrast represented a poor fit to the Japanese consonant to which it was contrasted. Table VII also summarizes the effects of English-language experience reported in experiment 2. “No” indicates that there was no difference between groups differing in English–language

experience, whereas “yes” indicates that a relatively experienced group obtained significantly higher *A'* scores than a less experienced group. As predicted, the two contrasts involving “good-fitting” English consonants, /t/–/t/ and /b/–/b/, showed no effect of learning.

Contrary to prediction, just one of the three contrasts involving a “poor-fitting” English consonant showed evidence of learning, namely the /ɹ/–/r/ contrast. The Japanese groups did not differ significantly for the /θ/–/s/ contrast; however, a nonsignificant trend in the expected direction is evident (see Fig. 2). The /l/–/r/ contrast showed no effect of learning; all the Japanese participants received low *A'* scores for this contrast.

The asymmetry between English /ɹ/ and /l/ as compared to Japanese /r/ is of interest. Even though both /l/ and /ɹ/ received poor fit indexes to Japanese /r/, there was some evidence that the English /ɹ/ is more perceptually distant from the Japanese /r/ than the English /l/ is. As mentioned earlier, Takagi (1993) found that inexperienced Japanese listeners judged English /l/ tokens to be more similar to Japanese /r/ than English /ɹ/ tokens. Takagi used a different methodology to assess the perceived cross-language relations. He had one group of listeners identify the English stimuli and then another group of listeners rate those identifications. This procedure was adopted to avoid any biasing effect the identification responses might have on the ratings. Perhaps a rating procedure such as that used by Takagi would have revealed a difference in the /ɹ/ and /l/ ratings in our stimuli, as well. Further investigation into the /ɹ/ and /l/ ratings using Takagi’s methodology is warranted.

The remaining two *E–J* contrasts, /ɹ/–/ɰ/ and /v/–/b/, could not be used to test predictions derived from the SLM. This is because the English consonants were not considered to be an example of the Japanese categories to which they were being compared. The cross-language mapping data showed that English /ɹ/ was mapped to Japanese /ɰr/ or /r/ and that English /v/ mapped onto Japanese /v/. At the time the experiment was designed, we thought it possible that English /ɹ/ would be considered an example of Japanese /ɰ/ and that English /v/ would be considered an example of Japanese /b/ (see the Introduction), but that was not the case.

IV. GENERAL DISCUSSION

Experiment 2 examined the discriminability of English–English (*E–E*), English–Japanese (*E–J*), and Japanese–Japanese (*J–J*) consonant contrasts. Three groups of native Japanese speakers, who differed in English–language experience, and a native English group, participated. The Japanese groups obtained high discrimination (*A'*) scores for the *J–J* contrasts, indicating that they fully understood and were able to perform the discrimination test. However, they received low scores on some *E–E* and *E–J* contrasts, indicating that these contrasts were difficult to discriminate. In addition, some contrasts were discriminated more accurately by Japanese groups with more English experience than by groups with less English experience, showing an effect of learning. This indicated that some *E–E* and *E–J* contrasts were more learnable than others.

A careful assessment of the experiment 1 results provided insight into how the perceived phonetic distance of English and Japanese consonants affected discrimination of English consonants, as well as how phonetic distance affected learning of English consonants over time. Experiment 1 provided the identification and goodness-of-fit of English to Japanese consonants as heard by native speakers of Japanese. The cross-language mapping data obtained in experiment 1 made it possible to explore possible extensions of current models of non-native discrimination and L2 learning.

The question pertaining to the perceptual assimilation model (PAM) (Best, 1995) addressed here was: Can the PAM framework be extended to the naturalistic acquisition of English as an L2? The results obtained here suggest an affirmative answer to this question. PAM was able to predict the discrimination of L2 consonants based on the perceived relationship between English (the L2) and Japanese (the L1) consonants. The results from our study also indicated a minor revision to PAM.

The question pertaining to the speech learning model (SLM) addressed here was: Can the SLM framework be extended to the acquisition of L2 by relatively inexperienced learners of an L2? The results obtained here suggest that the SLM cannot be extended without revision. The SLM was able to predict learnability effects for some *E-J* contrasts but not others. Only one out of the three contrasts in which the English consonant was perceptually distant from the Japanese consonant to which it was being compared showed learning effects. Perhaps the amount of English experience of the most experienced group (3 years spent living in the United States) was insufficient for a learning effect to emerge for some contrasts. Or, perhaps a more sensitive cross-language mapping methodology might have supported the predictions. Investigation into whether *degree* of perceptual distance (among sounds considered to be poor fits to the L1 sound) is related to the emergence of learning effects could suggest an extension of the SLM. Perhaps those L2 sounds that are more distant from the L1 sounds will be learned earlier than those sounds that are not as distant (although still considered poor fits to the L1 category).

In summary, the results obtained here indicated that certain English consonant contrasts are more difficult for Japanese adults to discriminate than others. The degree of perceptual difficulty seemed to depend on the extent to which the two members of a consonant contrast would be identified as instances of a single Japanese consonant category. Thus, the results obtained here suggest that the PAM framework can be extended to early stages of naturalistic L2 speech learning. However, the relative "learnability" of the consonant contrasts did not seem to depend lawfully on the perceived cross-language similarity of English and Japanese consonants. This finding suggests that the SLM cannot be readily extended to early stages of L2 speech acquisition without further investigation.

ACKNOWLEDGMENTS

This research was supported by a grant from the National Institute for Deafness and Other Communicative Dis-

orders (No. DC00257). The authors thank Thorsten Piske, Rattree Wayland, and Jonathan Loftin for their many helpful comments.

¹The symbol /ɾ/ will be used to represent the Japanese liquid.

²Japanese is not usually analyzed as having a /v/ phoneme; however, there is an orthographic representation for /v/ that is used primarily in writing loan words.

³It is possible that some identifications were highly agreed upon because they were in fact good tokens of the Japanese category, whereas others were highly agreed upon because there were no other good competitors. The goodness-of-fit scores were collected to help distinguish between these two possibilities. Weighting the identification scores by the goodness-of-fit data served to raise the scores of those identifications that were indeed considered good tokens of the category and to lower the scores of those identifications that were selected because they had no good competitors. The data in Table III show that the identification ratings are *not* perfectly correlated with the fit indexes; the goodness-of-fit data have done the expected job and modified the identifications.

⁴It is noteworthy that when the same procedure was applied to the Japanese consonants, none of them received the designation of "poor."

⁵ A' was calculated by the following formula provided by Snodgrass *et al.* (1985, p. 451), where H = hit rate (i.e., the proportion of different trials in which the odd item was correctly selected), FA = false alarm rate (i.e., the proportion of catch trials in which an odd item was incorrectly selected).

(1) If $H > FA$ then $A' = 0.5 + [(H - FA)(1 + H - FA)]/[4H(1 - FA)]$;

(2) If $H = FA$ then $A' = 0.5$;

(3) If $H < FA$ then $A' = 0.5 - [(FA - H)(1 + FA - H)]/[4FA(1 - H)]$.

⁶The English consonant is listed first, the Japanese consonant second for all *E-J* contrasts.

Best, C. (1990). "Adult Perception of Nonnative Contrasts Differing in Assimilation to Native Phonological Categories," *J. Acoust. Soc. Am.* **88**, S177.

Best, C. (1993). "Emergence of Language-Specific Constraints in Perception of Non-Native Speech: A Window on Early Phonological Development," in *Developmental Neurocognition: Speech and Face Processing in the First Year of Life*, edited by B. de Boysson-Bardies, S. de Schonen, P. Jusczyk, P. MacNeilage, and J. Morton (Kluwer, Dordrecht, The Netherlands).

Best, C. (1995). "A Direct Realist View of Cross-Language Speech Perception," in *Speech Perception and Linguistic Experience: Issues in Cross-Language Research*, edited by W. Strange (York, Timonium, MD), pp. 171-204.

Best, C., Faber, A., and Levitt, A. (1996). "Assimilation of Non-Native Vowel Contrasts to the American English Vowel System," *J. Acoust. Soc. Am.* **99**, 2602(A).

Best, C., and Strange, W. (1992). "Effects of Phonological and Phonetic Factors on Cross-Language Perception of Approximants," *J. Phonetics* **20**, 305-350.

Best, C., McRoberts, G., and Sithole, N. (1988). "The Phonological Basis of Perceptual Loss for Non-Native Contrasts: Maintenance of Discrimination Among Zulu Clicks by English-Speaking Adults and Infants," *J. Exp. Psychol. Hum. Percept. Perform.* **14**, 345-360.

Bradlow, A., Pisoni, D., Akahane-Yamada, R., and Tohkura, Y. (1997). "Training Japanese Listeners to Identify English /ɾ/ and /l/. IV. Some Effects of Perceptual Learning on Speech Production," *J. Acoust. Soc. Am.* **101**, 2299-2310.

Flege, J. (1995). "Second-Language Speech Learning: Theory, Findings and Problems," in *Speech Perception and Linguistic Experience: Issues in Cross-Language Research*, edited by W. Strange (York, Timonium, MD), pp. 233-273.

Flege, J., Takagi, N., and Mann, V. (1995). "Japanese Adults Can Learn to Produce English /ɪ/ and /I/ Accurately," *Lang Speech* **38**, 35-55.

Flege, J., Takagi, N., and Mann, V. (1996). "Lexical Familiarity and English-Language Experience Affect Japanese Adults' Perception of /ɪ/ and /I/," *J. Acoust. Soc. Am.* **99**, 1161-1173.

Goto, H. (1971). "Auditory Perception by Normal Japanese Adults of the Sounds L and R," *Neuropsychologia* **9**, 317-323.

Gottfried, T. (1984). "Effects of Consonant Context on the Perception of French Vowels," *J. Phonetics* **12**, 91-114.

Lado, R. (1957). *Linguistics Across Cultures* (University of Michigan Press, Ann Arbor, MI).

- Lambacher, S., Marten, W., Nelson, B., and Berman, J. (1997). "Perception of English Voiceless Fricatives by Native Speakers of Japanese," in *New Sounds 97: Proceedings of the Third International Symposium on the Acquisition of Second-Language Speech*, edited by J. Leather and A. James (University of Klagenfurt Press, Klagenfurt, Austria), pp. 186–195.
- Lively, S., Logan, J., and Pisoni, D. (1993). "Training Japanese Listeners to Identify English /r/ and /l/. II. The Role of Phonetic Environment and Talker Variability in Learning New Perceptual Categories," *J. Acoust. Soc. Am.* **94**, 1242–1255.
- Lively, S., Pisoni, D., Yamada, R., Tohkura, Y., and Yamada, T. (1994). "Training Japanese Listeners to Identify English /r/ and /l/. III. Long-Term Retention of New Phonetic Categories," *J. Acoust. Soc. Am.* **96**, 2076–2087.
- Logan, J., Lively, S., and Pisoni, D. (1991). "Training Japanese Listeners to Identify English /r/ and /l/: A First Report," *J. Acoust. Soc. Am.* **89**, 874–886.
- MacKain, K., Best, C., and Strange, W. (1981). "Categorical Perception of English /r/ and /l/ by Japanese Bilinguals," *Appl. Psycholing.* **2**, 369–390.
- Miyawaki, K., Strange, W., Verbrugge, R., Liberman, A., Jenkins, J., and Fujimura, O. (1975). "An Effect of Linguistic Experience: The Discrimination of /r/ and /l/ by Native Speakers of Japanese and English," *Percept. Psychophys.* **18**, 331–340.
- Mochizuki, M. (1981). "The Identification of /r/ and /l/ in Natural and Synthesized Speech," *J. Phonetics* **9**, 283–303.
- Polka, L. (1995). "Linguistic Influences in Adult Perception of Non-Native Vowel Contrasts," *J. Acoust. Soc. Am.* **97**, 1286–1296.
- Pruitt, J., Akahane-Yamada, R., Strange, W., Kubo, R., and Takada, T. (1999). "Perceptual Training with Native Japanese Speakers on Various L2 Sounds," *J. Acoust. Soc. Am.* **105**, 1096(A).
- Ritchie, W. (1968). "On the Explanation of Phonic Interference," *Language Learning* **18**, 183–197.
- Schmidt, A.-M. (1996). "Cross-Language Identification of Consonants. Part 1. Korean Perception of English," *J. Acoust. Soc. Am.* **99**, 3201–3211.
- Sekiyama, K., and Tohkura, Y. (1993). "Inter-Language Differences in the Influence of Visual Cues on Speech Perception," *J. Phonetics* **21**, 427–444.
- Snodgrass, J., Levy-Berger, G., and Haydon, M. (1985). *Human Experimental Psychology* (Oxford University Press, New York).
- Strange, W., Akahane-Yamada, R., Kubo, R., Trent, S., Nishi, K., and Jenkins, J. (1998). "Perceptual Assimilation of American English Vowels by Japanese Listeners," *J. Phonetics* **26**, 311–344.
- Takagi, N. (1993). "Perception of American English /r/ and /l/ by Adult Japanese Learners of English: A Unified View," unpublished Univ. of California, Irvine, Ph.D. dissertation.
- Vance, T. (1987). *An Introduction to Japanese Phonology* (State University of New York Press, Albany, NY).
- Yamada, R., and Tohkura, Y. (1992). "The Effects of Experimental Variables on the Perception of American English /r/ and /l/ by Japanese Listeners," *Percept. Psychophys.* **52**, 376–392.
- Yoshida, K., and Hirasaka, F. (1983). "The Lexicon in Speech Perception," *Sophia Linguistica* **11**, 105–116.

Construction and interpretation of equal-tempered scales using frequency ratios, maximally even sets, and P-cycles

Richard J. Krantz^{a)}

Department of Physics, Metropolitan State College of Denver, Denver, Colorado 80217-3362

Jack Douthett

Department of Arts and Sciences, Albuquerque Community College and Technical-Vocational Institute, Albuquerque, New Mexico 87106

(Received 15 January 1999; revised 31 January 2000; accepted 1 February 2000)

Using recent developments in music theory, which are generalizations of the well-known properties of the familiar 12-tone, equal-tempered musical scale, an approach is described for constructing equal-tempered musical scales (with “diatonic” scales and the associated chord structure) based on good-fitting intervals and a generalization of the modulation properties of the circle of fifths. An analysis of the usual 12-tone equal-tempered system is provided as a vehicle to introduce the mathematical details of these recent music-theoretic developments and to articulate the approach for constructing musical scales. The formalism is extended to describe equal-tempered musical scales with nonoctave closure. Application of the formalism to a system with closure at an octave plus a perfect fifth generates the Bohlen–Pierce scale originally developed for harmonic properties similar to traditional chords but without the perceptual biases of these familiar chords. Subsequently, the formalism is applied to the group-theory-based 20-fold microtonal system of Balzano. It is shown that with an appropriate choice of nonoctave closure (6:1 in this case), determined by the formalism combined with continued fraction analysis, that this group-theoretic-generated system may be interpreted in terms of the frequency ratios 21:56:88:126. Although contrary to the spirit of the group-theoretic approach to generating scales, this analysis may be applicable for discovering the ratio basis of unusual tunings common in non-Western music. © 2000 Acoustical Society of America. [S0001-4966(00)01305-9]

PACS numbers: 43.75.Bc [WJS]

INTRODUCTION

Musical scales, of necessity, are built on compromise among three competing expectations. First and foremost, we have come to expect music to be based on harmonious sounds. With this in mind, historically, musical scales were constructed of intervals generated by the ratios of partial frequencies.^{1,2} This expectation was formally established when, in the 19th century, Helmholtz³ showed that our perception of consonance is strongly influenced by the presence of these ratios. Second, we expect music to have variety and versatility. This is usually accomplished in Western music by modulation to closely related keys.⁴ Third, we expect a certain amount of convenience regarding modulation and transposition. Equal-tempered scales, that preserve (or close) the octave, in which all half-steps are equally spaced logarithmically, make modulation and transposition straightforward.⁵ It is well-known that the expectations of a harmonious scale and the ease of modulation and transposition afforded by equal-tempered scales are incompatible.^{1,2,5–11} As a result, many authors^{6,12–15} have discussed various mathematical approaches for approximating harmonious scales in the context of equal temperament.

Contemporary music is usually written with 12 equal-tempered semitones to the octave. The reasons for this system is directly related to the musical expectations referred to

above. First, the octave, which is considered the most harmonious interval,⁵ is preserved. Second, 12 equal-tempered notes to the octave are used because the just (harmonious) intervals of the perfect fifth (frequency ratio 3/2), the major third (frequency ratio 5/4), and the minor third (frequency ratio 6/5) are reasonably approximated by their closest equal-tempered interval.⁶ Therefore, our expectation of harmonious sounds is met for the octave and approximately so for the other three most important musical intervals. Third, 12 is the first relatively small number of notes to the octave that allows for a reasonable approximation of these just intervals and allows for enough variety by modulation through 12 different keys. Fourth, because it is an equal-tempered system transposition is relatively easy.

In the first three subsections (IA, IB, and IC) of Sec. I we analyze the 12-tone equal-tempered system in use today in light of recent developments in music theory.^{6,16,17} We do this in order to introduce these mathematical developments in the context of a familiar example. In Sec. ID the chord structure of our usual 12-tone system is discussed in the context of the mathematical notation developed in the first three subsections. In Secs. IE and IF we articulate and summarize our approach for constructing equal-tempered musical scales based on these recent developments.

In Sec. IIA we extend our formalism to describe nonoctave musical systems and then apply the extended formalism to the Bohlen–Pierce scale in Sec. IIB. As another example, the formalism is shown to generate all the details of the

^{a)}Electronic mail: krantzr@mscd.edu

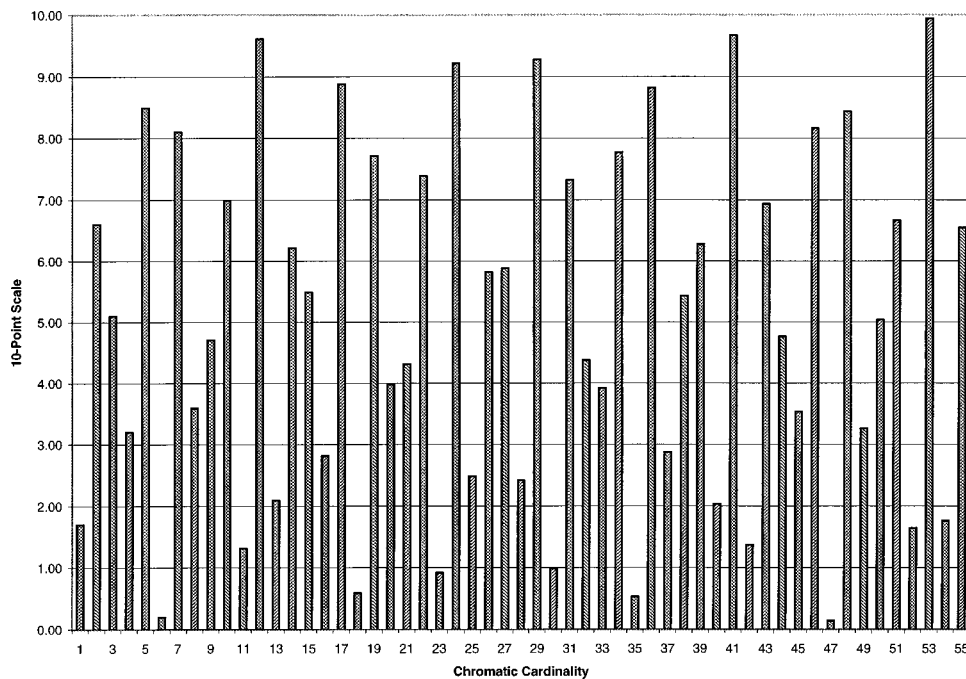


FIG. 1. Ten-point desirability function (closure of the octave by the pure fourth).

group-theoretic 20-fold system of Balzano in Secs. II C 1 and 2. The interpretive power of the formalism is shown in Sec. II C 3, where we show that Balzano's 20-fold scale may be analyzed in terms of frequency ratios in the appropriate non-octave system.

I. A MATHEMATICAL INTERPRETATION OF 12-TONE EQUAL-TEMPERED MUSIC

A. Equal-temperament, approximation of just intervals, and the desirability function⁶

Recently, Krantz and Douthett⁶ developed a single interval ten-point *desirability function*, based on the concept of octave closure, to assess the ability of c -tone equal-tempered

musical scales to best approximate just (harmonious) musical intervals. In this notation, c refers to the chromatic cardinality (number of notes) of the scale. A generalization of this desirability function, based on the concept of a generalized comma, was developed to assess the ability of c -tone equal-tempered scales to best approximate multiple just intervals simultaneously. Furthermore, an extension of this generalized desirability function was developed in which individual single intervals could be weighted according to the relative preference of one interval over another. The weighted, multiple-interval, ten-point desirability function is

$$D(c, N) = 10 - 20 \sum_{i=1}^N p_i \left| \left\{ c \log_2(R_i) + \frac{1}{2} \right\} - \frac{1}{2} \right|, \quad (1)$$

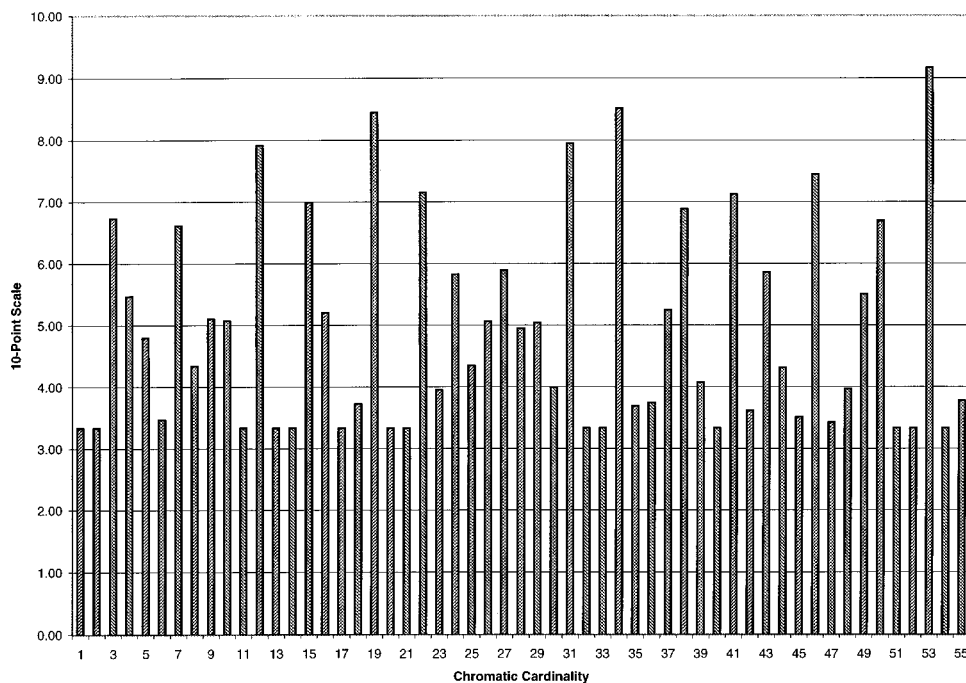


FIG. 2. Generalized ten-point desirability function (closure of the octave by the pure fourth, major third, and minor third, simultaneously, with equal weighting).

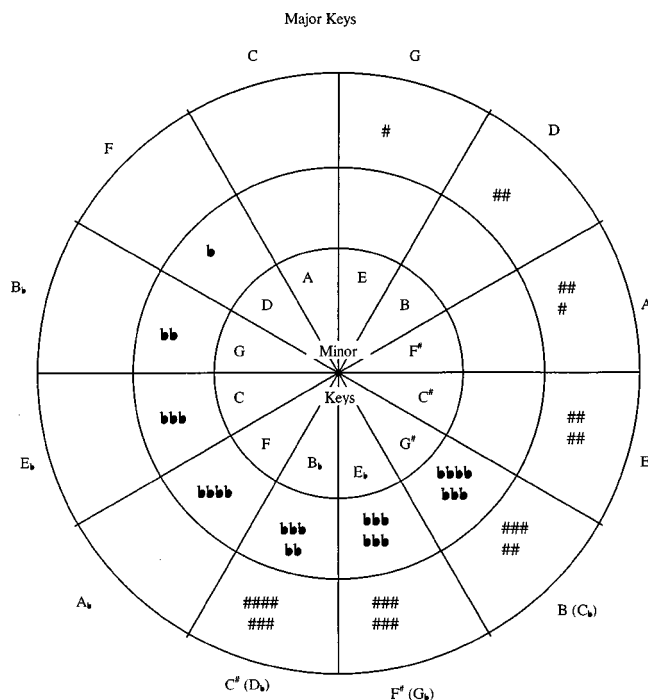


FIG. 3. Circle of fifths.

where $\{x\}$ is the fractional part of x , N is the number of intervals to be approximated, and c , the chromatic cardinality, is the number of equal-tempered intervals (i.e., number of notes) per octave. The R_i 's are the frequency ratios of the individual intervals, and the p_i 's are the respective normalized weights of the R_i 's. If the intervals are weighted equally (i.e., $p_i = 1/N$ for all i), then the desirability of the intervals collectively is the same as the average of the desirability of the individual intervals.

Shown in Fig. 1 are the results of applying Eq. (1) to the single interval of the pure fourth ($R = 4/3$), or equivalently, the inversion of the pure fourth—the perfect fifth ($R = 3/2$), for chromatic cardinalities up through 55. A value near 10 on the vertical axis indicates that the pure fourth is well approximated by the chromatic cardinality on the horizontal axis. As shown in the figure, for relatively small chromatic cardinalities, 12 is the best equal-tempered system which best approximates the just interval of the pure fourth. Shown in Fig. 2 are the results of applying Eq. (1) to the just intervals of the pure fourth ($R_1 = 4/3$), major third ($R_2 = 5/4$), and minor third ($R_3 = 6/5$) simultaneously with equal weighting ($p_1 = p_2 = p_3 = 1/3$). As shown, 12 notes is the best choice for an equal-tempered scale with relatively few notes to the octave which provides enough variety and versatility, as discussed above. Historically, 12 notes to the octave was the preferred choice due, in part, to the well-approximated perfect fifth (or, equivalently, the pure fourth). Other examples are discussed in Ref. 6.

B. Modulation

1. Circle of fifths and unidirectional P-cycles

As pointed out in the *New Harvard Dictionary of Music*,⁴ modulation is “the process of changing from one key to another, or the result of such a change.” The impor-

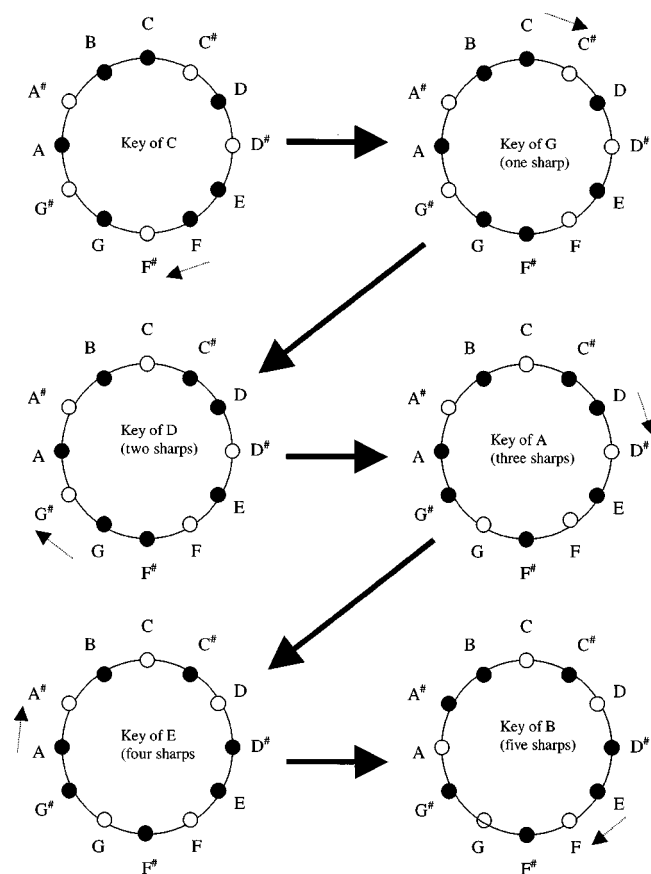


FIG. 4. Progression through a portion of the circle of fifths (see the text for details).

tance of modulation in Western music is stated thus, “The capability for modulation, even more than the establishment of key, is the most distinctive and powerful property of the tonal system in Western music, especially since the community of twelve major and twelve minor keys was made internationally practical by equal temperament.” As pointed out in Ref. 4, the most common modulations are between closely related keys which are those keys adjacent to each other on the circle of fifths.

A diagram of the circle of fifths is shown in Fig. 3. The keys on the outside represent major keys and keys on the inside represent minor keys. Starting on any key (major or minor), the next key clockwise (or counterclockwise) represents an interval of a fifth. For example, starting on F^\sharp major, C^\sharp major is a fifth above and these two keys are so-called closely related keys. Again, from the *New Harvard Dictionary of Music*; “Because of the way in which sharps and flats are added to key signatures along the circle, the number of pitches in common between the starting key and each successive key outward in either direction decreases by one.”

This property is most easily seen by reference to Fig. 4. We progress through a portion of the circle of fifths starting with the key of C major. The filled circles represent the 7 notes, out of the 12, used in the major scale. The heavy arrows show the progression through the circle of fifths. The dotted arrows indicate which note is moved going to the next key in the circle of fifths. As seen in the figure, each succes-

sive change of key has six of seven notes in common with the previous key. As we progress further and further from the starting key, each successive key has one less note in common with the starting key.

In recent works in music theory, microtonal versions of the circle of fifths have become known as *unidirectional P-cycles* (P for proximity).¹⁸ More technically, a *P-cycle* is a cycle of three or more pairwise distinct sets (scales or chords) from the same set class (sets equivalent under transposition and inversion) such that there is a map between each pair of adjacent sets that leaves all but one of its notes fixed. Moreover, the note that moves is changed by precisely a half-step. If the sets in the cycle exhaust the set class then the cycle is said to be unidirectional, otherwise it is a *toggling P-cycle* (these names came about as a result of particular properties inherent in these two types of cycles). Thus, the circle of fifths is a *unidirectional P-cycle*, while the 6-cycle of consonant triads e, E, g#, G#, c, C, and back to e is a *toggling P-cycle*. Since we will restrict our attention to unidirectional P-cycles, in what follows it will be understood that all P-cycles are unidirectional for our purposes.

2. Mathematical generalization of unidirectional P-cycles

Over the last 65 years other scale structures that have the properties of or related to unidirectional P-cycles have been constructed by Yasser,¹⁹ Mendalbaum,¹² Chalmers,^{20–22} Wilson,²³ Gamer,²⁴ Balzano,²⁵ Clough and Myerson,^{26,27} Mathews, Pierce, Reeves, and Roberts,²⁸ Agmon,²⁹ Clampitt,¹⁶ Clough and Douthett,¹⁷ Clough, Cuciurean, and Douthett,³⁰ and Zweifel.³¹ In this discussion we adopt much of the terminology from the seminal works of Clough and Myerson.^{26,27} These terms arise as a result of musical intervals being associated with two somewhat incompatible lengths as described below.

In the usual diatonic (major) scale, the musical third comes in one of two sizes; three half-steps (the minor third) and four half-steps (the major third), whereas six half-steps (the tritone) can either be a fourth (augmented) or a fifth (diminished). To generalize and relate these measures, let c be the size of the *chromatic universe* or the chromatic cardinality, as before. Similarly, for a given scale, let d be the *diatonic cardinality* which is the number of pitch classes (e.g., all C's belong to the same pitch class) that define the scale. We may denote a given scale by the set

$$S_{c,d} = \{s_0, s_1, s_2, \dots, s_{d-1}\}, \quad (2)$$

where the elements are organized so that $s_0 < s_1 < s_2 < \dots < s_{d-1}$. In our usual 12-tone chromatic universe, if we make the assignments C=0, C#/D \flat =1, D=2, etc. the C major scale may be represented by $S_{12,7} = \{0, 2, 4, 5, 7, 9, 11\}$. Then the *chromatic length from s_i to s_j* is $(s_j - s_i)$ expressed as the smallest non-negative integer modulo c , and the *diatonic length from s_i to s_j* is $(j - i)$ expressed as the smallest non-negative integer modulo d . For the diatonic scale the chromatic length from one note to another (reduced to within the octave) is the length of the interval in half-steps, and the diatonic length is *one less* (for mathematical convenience) than the usual musical interval (unison, second, third, etc.).

For example, the chromatic length from the note B to E is 5, and since this interval is a fourth, the diatonic length is 3.

Although their investigation was primarily to generalize other properties of the diatonic scale, Clough and Myerson^{26,27} were the first to construct an algorithm that generates, up to transposition, the complete family of sets that induces P-cycles. Also, Clough and Douthett¹⁷ extended and generalized this algorithm now known as the *maximally even (ME) algorithm*. This algorithm is discussed in the next subsection.

C. Maximally even (ME) sets

1. Diatonic scales

All of the seven note diatonic scales shown in the circle of fifths may be generated by the maximally even (ME) algorithm developed by Clough and Douthett:

Let c and d be the chromatic and diatonic cardinalities, respectively, and let n be any fixed integer such that $0 \leq n \leq c - 1$. Then, a *ME set* with these parameters is

$$J_{c,d}^n = \left\{ \left\lfloor \frac{ck + n}{d} \right\rfloor \right\}_{k=0}^{d-1}, \quad (3)$$

where $\lfloor x \rfloor$ is the smallest integer greater than or equal to x .^{17,30,32}

For our usual 12-tone system with a 7-note diatonic scale, $c = 12$ and $d = 7$. If we let the notes C, C#/D \flat , D, D#/E \flat , etc., in Fig. 4 be represented by the numbers 0, 1, 2, 3, etc., respectively, Eq. (3) generates all the scales of the circle of fifths. For example, the C major scale is represented by the ME set

$$J_{12,7}^5 = \left\{ \left\lfloor \frac{12k + 5}{7} \right\rfloor \right\}_{k=0}^6 = \{0, 2, 4, 5, 7, 9, 11\}. \quad (4)$$

Note that the index of the C major scale is $n = 5$. As the index increases incrementally (mod 12) the associated scales rotate clockwise around the cycle of fifths in Fig. 3. Because the ME algorithm reproduces the scales in the circle of fifths, in order, it preserves the properties of the *unidirectional P-cycles* referred to in the previous subsection.

2. Maximally even sets and P-cycles

Clampitt¹⁶ has shown that a set can induce a P-cycle if and only if the set is maximally even and the chromatic cardinality, c , and the diatonic cardinality, d , are coprime, or $(c, d) = 1$ in mathematical parlance. Clough and Douthett¹⁷ have shown that such sets are also *generated sets* with two distinct generators (e.g., the fourth and fifth generate the major scales in our familiar 12-tone system). Moreover, the chromatic and diatonic lengths of the generators can be determined by the linear Diophantine equation

$$cI_g - dg = \pm 1, \quad (5)$$

where g is the chromatic length of a generator and I_g is the corresponding diatonic length.³³

For example, consider the usual diatonic scale ($c = 12$ and $d = 7$). Then, Eq. (5) becomes

$$12I_g - 7g = \pm 1. \quad (6)$$

The only chromatic and diatonic lengths that satisfy Eq. (6) are $g=5$, $I_g=3$ (the fourth) and $g=7$, $I_g=4$ (the fifth). It is left to the reader to show that the fourth or the fifth also generates the pentatonic scale ($c=12$ and $d=5$).

Although the diatonic length of a generator plays an important role in the theory of scales, our subsequent analysis will consider only chromatic lengths of generators. Therefore, Eq. (5) simplifies to

$$dg \equiv \pm 1 \pmod{c}. \quad (7)$$

We will have occasion to use these generalizations in Sec. II.

D. Intervals and chords

In our usual 12-tone system the intervals of the fourth with a frequency ratio of 4/3, the major third with a frequency ratio of 5/4, and the minor third with a frequency ratio of 6/5 are of prime importance. The sequence of these ratios is 3:4:5:6, and the chromatic lengths of these intervals are 5, 4, and 3, respectively. [Note that the ratio of the first element in the sequence of ratios (3) to the last element (6) produces the frequency ratio 6:3 or 2:1, which is the ratio of the octave. More on this later.] The major and minor triad structure can now be determined from the chromatic lengths.

The sequence of these lengths arranged from largest to smallest is the step-interval sequence of the major triads. In this case the sequence is (5,4,3) and the class of major triads (equivalent under transposition) is

$$M = \{\{0,5,9\}, \{1,6,10\}, \{2,7,11\}, \{0,3,8\}, \dots\}. \quad (8)$$

The sequence is cyclic; therefore (4,3,5) and (3,5,4) are in the same class. These sequences just generate chord inversions that belong to the same class. Thus, this family consists of major triads F, F \sharp , G, G \sharp , ..., E. Similarly, the minor triads are the triads whose step-interval sequence is the sequence of the chromatic lengths arranged from smallest to largest. The sequence is (3,4,5) and the class of minor triads is

$$m = \{\{0,3,7\}, \{1,4,8\}, \{2,5,9\}, \dots\}. \quad (9)$$

The union of M and m constitutes the set class of consonant triads.

In each diatonic scale there are six major/minor triads (three of each). For example, the triads G, C, D, a, b, and e are all *embedded* in the G major scale. The only embedded (third-generated) triad not of major or minor quality is the F \sharp dim triad. Therefore, using the chromatic lengths of the frequency ratios, 3:4:5:6 which are best approximated by our usual 12-tone equal-tempered system, we have determined the chord structure of the system.

E. An approach to constructing musical scales

We may now formalize an approach for constructing equal-tempered musical scales. Given a sequence of frequency ratios, which may be chosen for a variety of reasons (e.g., because they are “harmonious,” they convey some “mood,” or they meet some other mathematical, acoustical, or perceptual criteria), an equal-tempered system based on the desirability function, Eq. (1), that best approximates these ratios may be determined. The chromatic lengths of

TABLE I. Individual desirabilities for the frequency ratios 4/3, 5/4, and 6/5 for a 53-note equal-tempered system with octave closure. Also shown are the chromatic generators and diatonic cardinalities for each ratio. For comparison, the equal-weight ten-point desirability for all three ratios is shown.

Ratios	$c = 53$		
	Desirability function base 2	Chromatic generator	Diatonic cardinalities
4/3	9.940	22, 31	12, 41
5/4	8.756	17, 36	25, 28
6/5	8.816	14, 39	19, 34
All	9.171

these ratios can then be determined and used as generators to identify scales that induce P-cycles; Eqs. (7) and (3). These scales then have the modulation properties associated with the familiar circle of fifths. Once the chromatic lengths of the chosen frequency ratios are known, the triad structure of the scale can be determined from the step-interval sequence; e.g., Eqs. (8) and (9) for the usual diatonic scale.

We invite the reader to apply the formalism, using the fourth (or fifth) as the “chosen” interval and verify that the usual pentatonic scale ($c=12, d=5$) is generated. The results should be the complement of the diatonic set and produce the open circles shown in Fig. 4. It is also left to the reader to verify that the ME algorithm [Eq. (3)] generates the P-cycle of pentatonic scales.

For another example, it is well-known that the 53-tone, equal-tempered system closely approximates the ratios of the fourth, major third, and minor third. The desirability function shown in Fig. 2 reflects this result and shows that the 53-note-to-the-octave, equal-tempered scale is considerably better at approximating these intervals simultaneously. In Table I we summarize some results, using Eqs. (1) and (7), for a chromatic cardinality of $c=53$.

Since we are weighting the intervals equally, the overall desirability is the same as the average of the desirabilities of the individual intervals. It is left to the interested reader to construct all the scales that induce P-cycles (There are six set classes, two for each ratio, whose sets induce P-cycles.) and determine the embedded major and minor triads in each.

F. Summary

In light of recent developments in music theory we have articulated an approach for constructing equal-tempered musical scales that: (1) closes the octave, (2) approximates multiple chosen intervals simultaneously, (3) preserves modulation properties generalized from those of the circle of fifths, and (4) determines the scale and chord structure for the system. For historical and pedagogical reasons we have shown that the mathematical patterns inherent in the familiar 12-tone, equal-tempered system may be generalized. Alternatively, we could have taken an axiomatic approach by describing the formalism, applying it to the sequence of harmonious ratios (3:4:5:6), and generating our usual 12-tone, equal-tempered system.

In the next section, we extend the formalism to account for nonoctave closure and show that some unusual scales

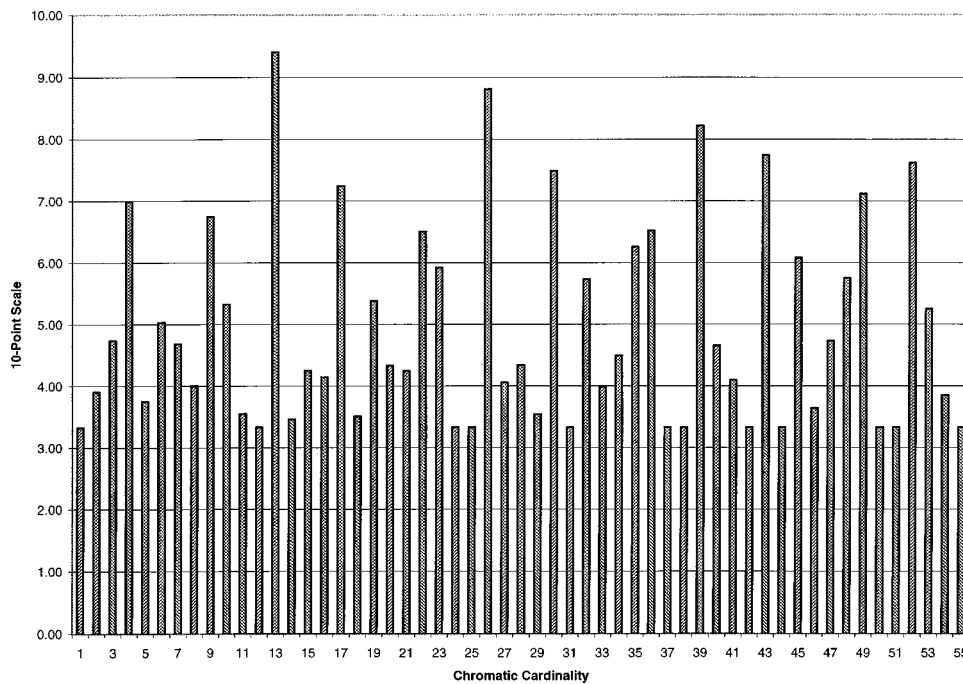


FIG. 5. GDF for the ratios 5/3, 7/5, and 9/7 (closure at the tritave; $b=3$).

generated by other workers^{25,28,31} based on very different criteria, that range from group-theoretic arguments to acoustical and perceptual reasons, are described by our approach.

II. ANALYSIS OF OTHER EQUAL-TEMPERED SYSTEMS

A. Extension to nonoctave systems

1. Nonoctave desirability function

The ten-point desirability function need not be restricted to octave (base 2) closure. A straightforward generalization of the weighted, multiple-interval desirability function yields the following:

$$D_b(c, N) = 10 - 20 \sum_{i=1}^N p_i \left[\left\{ c \log_b(R_i) + \frac{1}{2} \right\} - \frac{1}{2} \right], \quad (10)$$

where c , N , p_i , and R_i are defined as above. In the following, we refer to $D_b(c, N)$ as the generalized desirability function (GDF). The base, b , of the logarithm represents the interval of closure. For octave closure, the usual case, the base is 2; $D(c, N) = D_2(c, N)$.

2. Nonoctave example

In general, if the sequence of chosen ratios is $r_1:r_2:r_3:r_4$ the base for the desirability function is just $b = r_4/r_1$, as was the case for the pure fourth, major third, and minor third described above where the sequence of ratios was 3:4:5:6. In anticipation of the next section we chose the sequence of ratios 3:5:7:9 for our example. In this case, the base of the desirability function is 3 ($=9/3$), which means that our equal-tempered scale will close (repeat) every octave plus a fifth. Shown in Fig. 5 is the GDF with equal weighting for these ratios. It is clear from the figure that the best, reasonably small, chromatic cardinality is $c=13$ with a GDF of 9.40.³⁴ In fact, a better chromatic cardinality does not occur for these three ratios simultaneously, with equal weighting,

until $c=271$ with a GDF of 9.90. This desirability exceeds even that of the sequence of ratios 3:4:5:6 at 53 divisions to the octave (compare with Table I).

The desirability function, chromatic lengths, and diatonic cardinalities associated with this sequence of ratios are given in Table II. There are six scales (up to transposition), with diatonic cardinalities of 2, 3, 4, 9, 10, and 11, that induce P-cycles. Diatonic cardinalities $d=2, 3$, and 4 are too small to be of much interest (i.e., the scales would have too few pitch classes for much variation or variety). Therefore, the diatonic cardinalities $d=9, 10$, or 11 are better choices. Bohlen,³⁵ based on combination tones, and somewhat later Mathews, Pierce, Reeves, and Roberts,²⁸ based on the *intonation sensitivity measure* of Roberts and Mathews,³⁶ chose a diatonic cardinality $d=9$. This scale has become known as the Bohlen–Pierce scale. With this in mind we will use $d=9$ for our example and discuss these results in the next subsection.

B. The Bohlen–Pierce scale²⁸

Having chosen a chromatic cardinality of $c=13$ (according to our generalized desirability function) and a diatonic

TABLE II. Individual generalized desirabilities for the frequency ratios 5/3, 7/5, and 9/7 for a 13-note equal-tempered system with tritave (base 3) closure. Also shown are the chromatic generators and diatonic cardinalities for each ratio. For comparison, the equal-weight generalized desirability for all three ratios is shown.

$c=13$			
Ratios	Desirability function base 3	Chromatic generators	Diatonic cardinalities
5/3	9.107	6, 7	2, 11
7/5	9.630	4, 9	3, 10
9/7	9.477	3, 10	4, 9
All	9.405

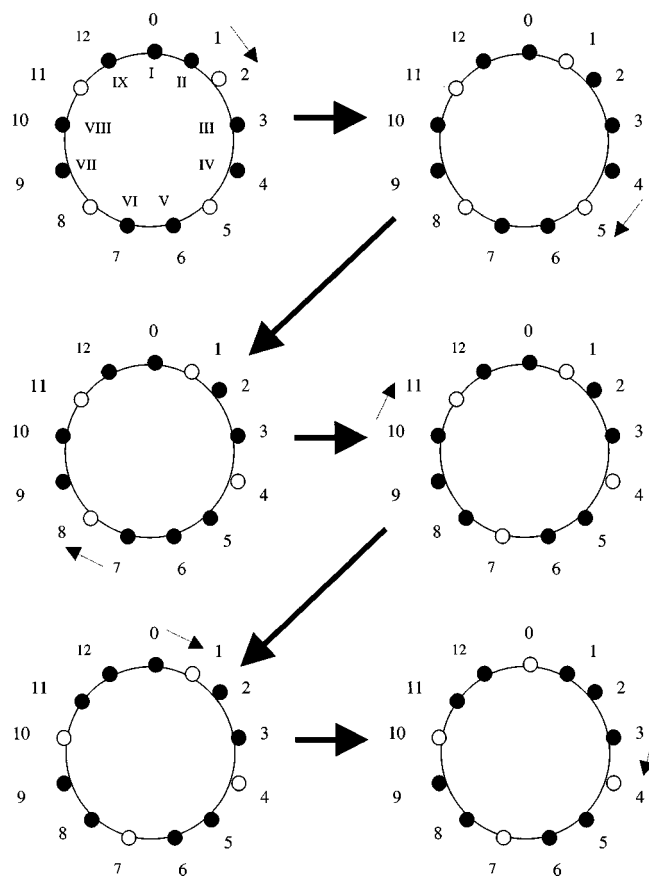


FIG. 6. A 13-note chromatic scale with 9-note embedded diatonic scale generated by the ME algorithm compatible with unidirectional P-cycles. The Bohlen–Pierce scale is superimposed on the first diagram. Progression through a few closely related keys is also shown.

cardinality of $d=9$ (in accordance with the modulation properties of *unidirectional* P-cycles) for the sequence of ratios 3:5:7:9 (chosen by previous workers^{28,35,36} for acoustic and perceptual reasons), with chromatic generators of 6 (or 7), 4 (or 9), and 3 (or 10), for an equal-tempered system with base 3 closure, we are ready to construct the “diatonic” scale and chord structure for the system. The ME algorithm for $c=13$, $d=9$, and $n=4$ yields the following for a typical scale:

$$J_{13,9}^4 = \left\{ \left\lfloor \frac{13k+4}{9} \right\rfloor \right\}_{k=0}^8 = \{0, 1, 3, 4, 6, 7, 9, 10, 12\}. \quad (11)$$

This scale is shown by the filled circles in the first diagram of Fig. 6. Superimposed on that diagram are the nine tones of the Bohlen–Pierce scale, denoted as I, II, III, etc., shown in Fig. 6 are the first few closely related keys generated by the ME algorithm, with indices; $n=4, 5, 6, 7, 8$, and 9. Because the chromatic cardinality and the diatonic cardinality are coprime these sets induce P-cycles, and therefore have the modulation properties of the generalized circle of fifths. Also shown in Fig. 6, as a dotted arrow, is the note that moves to generate the next closely related key. As before, the heavy arrows indicate the progression through the generalized circle of fifths.

The classes of major and minor triads are generated by arranging the step-interval sequence for the chosen ratios

from largest to smallest (6, 4, 3) and smallest to largest (3, 4, 6), respectively, as before. The class of Bohlen–Pierce major triads is

$$M_{BP} = \{\{0, 6, 10\}, \{1, 7, 11\}, \{2, 8, 12\}, \{0, 3, 9\}, \dots\}. \quad (12)$$

The class of Bohlen–Pierce minor triads is

$$m_{BP} = \{\{0, 3, 7\}, \{1, 4, 8\}, \{2, 5, 9\}, \dots\}. \quad (13)$$

For this scale, the set class of “consonant triads” is $M_{BP} \cup m_{BP}$. Although Bohlen³⁵ and Mathews, Pierce, Reeves, and Roberts²⁸ did not have the GDF, the ME algorithm, or the P-cycle algorithm, they constructed a scale consistent with our application of these tools for constructing scales given a chosen sequence of ratios.

C. A ratio interpretation of Balzano’s c -fold microtonal system²⁵

1. Introduction

Our next example of a nontraditional system is based on Balzano’s well-known c -fold microtonal system where $c = n(n+1)$ and n is an integer greater than or equal to 3. Balzano points out that “With the advent of the computer, the possibilities of exploring microtonal systems of octave division broadens considerably.” As an alternative, he bases his octave division constructs on mathematical group-theoretic properties rather than constructing microtonal systems based on good fits to frequency ratios.

The chromatic cardinalities considered by Balzano are the products of two consecutive integers (a generalization of $12=3 \cdot 4$ in our familiar 12-tone system) and the diatonic cardinalities are the sum of those two integers (a generalization of $7=3+4$). His “F to F# property” is equivalent to requiring scales that induce P-cycles and his “major” and “minor” triads have the interval sequence $(n^2-n-1, n+1, n)$ and $(n, n+1, n^2-n-1)$, respectively. Again, these sequences are the generalization of the usual $n=3$ case for the familiar 12-tone system. As a specific example, we will use Balzano’s 20-fold system for comparison.

2. P-cycles, ME sets, and triads

Balzano’s 20-fold system requires that $d=4+5=9$. With a diatonic cardinality of 9, the two generator lengths that satisfy Eq. (7), and hence generate scales that induce P-cycles are $g=9$ and $g=11$. (The length $g=11$ will also be important in the construction of the major and minor triads.) Since Balzano’s scale will induce a P-cycle, it can be determined via the ME algorithm

$$J_{20,9}^6 = \left\{ \left\lfloor \frac{20k+6}{9} \right\rfloor \right\}_{k=0}^8 = \{0, 2, 5, 7, 9, 11, 14, 16, 18\}. \quad (14)$$

This scale represents what Clough and Douthett¹⁷ call a *hyperpentatonic scale*, which is a ME set whose diatonic cardinality is 1 less than half the chromatic cardinality and the chromatic cardinality is divisible by 4. Note that if $c=12$ the hyperpentatonic scale reduces to the familiar pentatonic scale.

Balzano uses the diatonic length of 2 (the “third”) to generate triads. As seen in Eq. (14) the chromatic lengths of

TABLE III. For a small number of bases (that define closure), the principal-convergents-continued-fraction approximations for the step intervals (11, 5, and 4) in Balzano's 20-fold system are calculated.

Base	$c = 20$ and $d = 9$		
	$b^{11/20}$	$b^{5/20}$	$b^{4/20}$
2	3/2, 19/13, 41/28	6/5, 19/16, 25/21	7/6, 8/7, 23/20
3	2/1, 9/5, 11/6	4/3, 25/19, 229/174	5/4, 71/57, 147/118
4	2/1, 13/6, 15/7, 433/202	3/2, 7/5, 17/12	4/3, 29/22, 33/25
5	2/1, 5/2, 12/5, 17/7	3/2, 160/107, 643/430	3/2, 4/3, 7/5
6	2/1, 3/1, 8/3, 67/25	2/1, 3/2, 11/7, 25/16	3/2, 10/7, 83/58

the thirds are 4 and 5. Therefore, along with the chromatic length of the generator $g = 11$, the step-interval size for the third is 5. Along with a step interval of 11 for the generator, the step-interval sequences for major and minor triads become (11, 5, 4) and (4, 5, 11), respectively. The classes of Balzano 20-fold major and minor triads are

$$M_{B_{20}} = \{\{0, 11, 16\}, \{1, 12, 17\}, \{2, 13, 18\}, \dots\}, \quad (15)$$

$$m_{B_{20}} = \{\{0, 4, 9\}, \{1, 5, 10\}, \{2, 6, 11\}, \dots\}. \quad (16)$$

It is left to the reader to verify that there are four major triads and four minor triads embedded in Balzano's scale.

For a 20-fold system, sets with diatonic cardinality 11 and generators $g = 9$ or $g = 11$ also induce P-cycles. Recently, Zweifel³¹ has investigated the properties of these 11-note scales. These scales are the complements of Balzano's nine-note scales and belong to a larger class of scales first isolated by Agmon²⁹ and later studied in depth by Clough and Douthett¹⁷ and Clough, Cuciurean, and Douthett.³⁰ Scales from this class are embedded in chromatic universes where $c = 0 \pmod{4}$, and the diatonic cardinality is 1 more than half the chromatic cardinality. It is left to the reader to generate this class of scales and the resulting chord structure.

In the next subsection, we take an alternative, nonoctave approach to analyzing Balzano's 20-fold system.

3. Ratios and continued fractions

Schechter³⁷ and Douthett, Entringer, and Mullhaupt³⁸ have shown the utility of using continued-fraction-principal convergents³⁹ for generating equal-tempered musical scales with octave closures. We have taken a somewhat different approach in analyzing Balzano's 20-fold system. The step intervals of 11, 5, and 4 correspond to "frequency ratios" of the form: $b^{11/20}$, $b^{5/20}$, and $b^{4/20}$, where b is the base of the logarithm that defines closure. For example, $b = 2$ defines the usual case of octave closure and $b = 3$ defines closure at an octave plus a fifth, which was the case for the Bohlen-Pierce scale. We proceeded to choose a number of bases, $b = 2, 3, 4, 5$, and 6 and calculated the three frequency ratios given above, in each base. We then generated the first few principal-convergents-continued-fraction approximations for each frequency ratio in each base. The results of these calculations are summarized in Table III. For convenience, in Table III, we have excluded the musically trivial 1/1 ratio.

We then chose a candidate sequence of ratios in each base consistent with the principal-convergents-continued-fraction approximations and compatible with that base. For

TABLE IV. Analysis of Balzano's 20-fold system using candidate sequences compatible with principal-convergent-continued fractions (see the text for details).

Base	Sequence	$c = 20$ and $d = 9$			$D_b(20,3)$ Equal-weight GDF
		$b^{11/20}$ Fractions	$b^{5/20}$ Fractions	$b^{4/20}$ Fractions	
3	30:55:72:90	11/6	72/55	5/4	8.71
4	7:15:21:28	15/7	7/5	4/3	8.00
5	5:12:18:25	12/5	3/2	25/18	8.39
5	14:34:51:70	17/7	3/2	70/51	9.13
6	21:56:88:126	8/3	11/7	63/44	9.31
6	15:40:63:90	8/3	63/40	10/7	9.06

example, for base 2 we chose the ratios 10:15:18:(20) because $15:10 = 3:2$, $18:15 = 6:5$, and $20:18 = 10:9$ because $20:10 = 2$ (the base). In general, we cannot choose a base and a sequence of ratios in which all these ratios are consistent with the continued fraction convergents simultaneously. We chose the base for closure. The first ratio was chosen because it was the generator of the diatonic scale that induces P-cycles. In the cases examined, we chose either the second or third ratio to be compatible with a (relatively small integer) principal-convergent-continued-fraction approximation for that ratio. We allowed the remaining ratio to be compromised, i.e., it was not necessarily a continued-fraction convergent. For base 3, it turned out that the candidate sequence 5:9:12:15 yields the principal-convergent-continued fractions 9/5, 4/3, and 5/4. The ratio 15:5 yields the base. Similar results are obtained for other bases.

Each of these candidate ratio sequences was then used to calculate the equal-weight GDF. The candidate sequences, for bases through $b = 6$, leading to a GDF greater than or equal to 8.00 are shown in Table IV. Fractions shown in bold are the resulting principal-convergent-continued fraction approximation to the given step-interval ratio. The fractions given in normal type are the resulting compromised ratios. For example, for the base 3 ratio sequence 30:55:72:90 we get a ratio of 11/6 ($= 55/30$) for the generator of the diatonic scale, 5/4 ($= 90/72$) is the continued fraction approximation for $3^{4/20}$, and the middle ratio (72/55) is compromised. These ratios result in an equal-weight GDF of 8.71. As shown, based on the GDF, the best choice for closure that yields small-integer-principal-convergent-continued fraction ratios is base 6. This represents closure at two octaves plus a fifth. Alternatively, the ratio sequence that would generate Balzano's 20-fold system with a 9-note scale that induces P-cycles and yields a reasonable desirability function value is (21:56:88:126).

For comparison, Fig. 7 shows the equal-weight, multiple-interval desirability function, base 6, for these ratios. In the sequence of chromatic cardinalities $c = 20$ is the best small number chromatic cardinality. Note that although $c = 4$ has a relatively good value, it may be neglected as too small to be interesting.

4. Summary

We have extended Balzano's approach somewhat by suggesting, based on continued fractions and the GDF, a

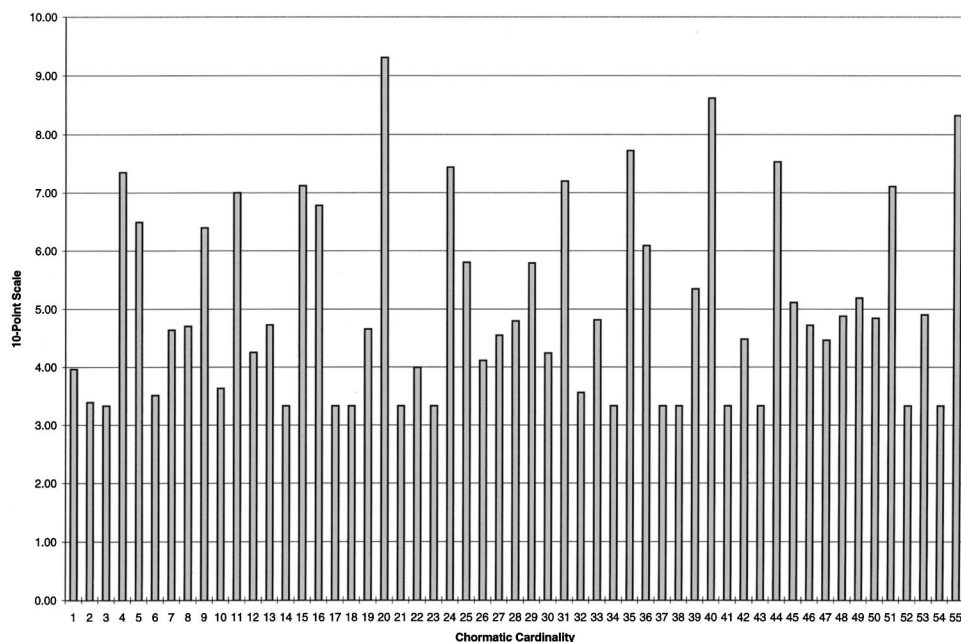


FIG. 7. GDF for the ratios 56/21, 88/56, and 126/88 (closure at the sextave; $b = 6$).

sequence of ratios (21:56:88:126) in a *nonoctave* system that may be used to generate his 20-fold, 9-note system. The ratios found to generate Balzano's system do not necessarily define "consonant" intervals. This is in the spirit of Balzano's approach that justified the system solely on group-theoretic grounds. If an *a priori* reason for choosing the above interval sequence could be justified, our formalism could be used to generate this system. Otherwise, the formalism can be used to interpret the system in terms of frequency ratios including nonoctave closure.

III. DISCUSSION

Based on good-fitting intervals, consistent with the multiple-interval desirability function⁶ and the modulation properties of unidirectional P-cycles,¹⁸ along with the application of maximally even set theory,¹⁷ we have developed a formalism for constructing equal-tempered scales, of any chromatic cardinality, with "diatonic" scales, and the associated chord structure, that have the modulation properties of the closely related keys in the usual 12-tone system. We have applied the formalism to the 12-tone, equal-tempered system to show that this approach describes the details of this traditional system, as well as to introduce the mathematical details of these recent music-theoretic developments in the context of a familiar example.

Also, the formalism has been extended to describe non-traditional scales that include nonoctave closure. As an example, the formalism was applied to the frequency ratios 3:5:7:9 of the Bohlen–Pierce scale. This scale was chosen originally for its similarity to the structural and acoustic properties of the usual diatonic scale. It is shown that our approach reproduces all the details of the Bohlen–Pierce scale. More generally, given a sequence of frequency ratios determined for their acoustical, structural, mathematical, or perceptual properties, our formalism can generate an equal-tempered system that: (1) has closure, (2) reasonably approximates the "chosen" frequency ratios simultaneously,

and (3) can yield a scale and chord structure that has the modulation properties of closely related keys.

As an example of the applicability of our approach, we show that the nontraditional 20-fold microtonal system of Balzano²⁵ can: (1) be generated by our formalism, and (2) can be interpreted in terms of the ratio sequence 21:56:88:126 in a 6:1 nonoctave system when the formalism is coupled with continued fraction analysis.

The above examples show the power of this approach for generating nonstandard musical systems. First, one can choose an interval for closure and assess the ability of various equal-tempered systems to approximate chosen intervals by application of the GDF, thereby determining an appropriate chromatic cardinality for the scale. Second, to ensure musical variety and variation, a diatonic cardinality that induces P-cycles, and therefore retains the modulation properties of closely related keys, can be chosen. Third, the ME algorithm can then be used to generate all the appropriate diatonic scales of the system. Fourth, the step-interval sequence, based on the chosen intervals, can be used to generate the chord structure for the system.

Alternatively, our approach can be used as an analysis tool in which a given scale may be interpreted in terms of ratios in an appropriate nonoctave system. Used in this way, we hope to discover a ratio basis of unusual tunings common in non-Western music.

¹J. M. Barbour, "The Persistence of Pythagorean Tuning Systems," *Scr. Math.* **1**, 286 (1933).

²J. M. Barbour, "Musical Logarithms," *Scr. Math.* **3**, 21 (1940).

³H. Helmholtz, in *On Sensation of Tone* (Dover, New York, 1954), Chaps. I and XVI (originally published in 1885).

⁴"Modulation," in *The New Harvard Dictionary of Music*, edited by D. M. Randell (Belknap, Cambridge, MA, 1986), p. 503.

⁵"Temperaments," in *The New Harvard Dictionary of Music* (Ref. 4), p. 837.

⁶R. J. Krantz and J. Douthett, "A Measurement of the Reasonableness of Equal-Tempered Musical Scale," *J. Acoust. Soc. Am.* **95**, 3642–3650 (1994).

- ⁷D. E. Hall, "The Objective Measure of Goodness-of-Fit for Tunings and Temperaments," *J. Music Theory* **17**, 274 (1973).
- ⁸D. E. Hall, "Quantitative Evaluation of Musical Scale Tunings," *Am. J. Phys.* **42**(7), 543 (1974).
- ⁹D. E. Hall, "Acoustical Numerology and Lucky Temperaments," *Am. J. Phys.* **56**(4), 329 (1988).
- ¹⁰D. E. Hall, "A Systematic Evaluation of Equal Temperaments Through $N=612$," *Interface (USA)* **14**, 61–73 (1985).
- ¹¹G. C. Hartmann, "A Numerical Exercise in Musical Scales," *Am. J. Phys.* **55**(3), 223 (1987).
- ¹²J. Mandelbaum, "Multiple Division of the Octave and Tonal Resources of 19-Tone Temperament," Ph.D. dissertation (Indiana University, Bloomington, IN, 1961).
- ¹³W. Stoney, "Theoretical Possibilities for Equally Tempered Musical Systems," in *The Computer and Music*, edited by H. B. Lincoln (Cornell University Press, Ithaca, NY, 1970), pp. 163–171.
- ¹⁴D. De Klerk, "Equal Temperament," *Acta Musicol.* **51**, 140 (1979).
- ¹⁵M. Yunik and G. Swift, "Tempered Music Scales for Sound Synthesis," *Comput. Music J.* **4**(4), 60 (1980).
- ¹⁶D. Clappitt, "Pair-Wise, Well-Formed Scales: Structured and Transformational Properties," Ph.D. dissertation (SUNY at Buffalo, Buffalo, NY, April 1997).
- ¹⁷J. Clough and J. Douthett, "Maximally Even Sets," *J. Music Theory* **35**, 93–173 (1991).
- ¹⁸In July of 1993 Clough (SUNY at Buffalo) assembled a group—known as the "SUNY Buffalo Working Group"—of music theorists, music psychologists, and mathematicians to investigate and expand on the ideas put forth by Richard Cohn (University of Chicago) on late 19th century voice leading. The term *P-cycles* was introduced by Cohn and was among the topics explored at the meeting. This group met again in July of 1997 to explore a related topic, *neo-Riemannian Transformations*. For more information on the above, including a bibliography of related topics, see the special topics edition of the *Journal of Music Theory* **42**(2) (1998).
- ¹⁹J. Yasser, *The Theory of Evolving Tonality* (American Library of Musicology, New York, 1932).
- ²⁰J. Chalmers, "Cycle Scales," *Xenharmonikon* **4**, 69–78 (1975).
- ²¹J. Chalmers, "Polychordal Matrices and MOS Scales," *Xenharmonikon* **7–8**, 156–167 (1979).
- ²²J. Chalmers, "Construction and Harmonization of Microtonal Scales in Non-Twelve-Tone Equal Temperaments," Proceedings of the 8th International Computer Music Conference, Venice, Italy, pp. 534–555, 1982 (unpublished).
- ²³E. Wilson, Private communication to J. Douthett (1996).
- ²⁴C. Gamer, "Some Combinatorial Resources of Equal-Tempered Systems," *J. Music Theory* **11**(1), 32–59 (1967).
- ²⁵G. Balzano, "The Group-Theoretic Description of 12-Fold Pitch Systems," *Comput. Music J.* **4**, 66–84 (1980).
- ²⁶J. Clough and G. Myerson, "Musical Scales and the Generalized Circle of Fifths," *Am. Math. Monthly* **93**(9), 695–701 (1986).
- ²⁷J. Clough and G. Myerson, "Variety and Multiplicity in Diatonic Systems," *J. Music Theory* **29**, 249–270 (1985).
- ²⁸M. V. Mathews, J. R. Pierce, A. Reeves, and L. A. Roberts, "Theoretical and Experimental Explorations of the Bohlen–Pierce Scale," *J. Acoust. Soc. Am.* **84**, 1214–1222 (1988).
- ²⁹E. Agmon, "A Mathematical Model of the Diatonic System," *J. Music Theory* **33**, 1–25 (1989).
- ³⁰J. Clough, J. Cuciurean, and Douthett, "Hyperscales and the Generalized Tetrachord," *J. Music Theory* **41**(2), 67–100 (1997).
- ³¹P. Zweifel, "Generalized Diatonic and Pentatonic Scales: A Group Theoretic Approach," *Perspect. New Music* **34**(1), 140–161 (1996).
- ³²To get a sense of why these sets are called *maximally even*, S. Block and J. Douthett, *J. Music Theory* **38**, 21 (1994) constructed a measure of the evenness of subsets in the same cardinal family (e.g., same chromatic cardinality or same diatonic cardinality). This measure is consistent with the ME algorithm.
- ³³J. Douthett, "The Theory of Maximally and Minimally Even Sets, the One-Dimensional Antiferromagnetic Ising Model, and the Continued Fraction Compromise of Musical Scales," Ph.D. dissertation (University of New Mexico, Albuquerque, NM, May 1999).
- ³⁴When c is a denominator of a continued fraction convergent of the \log_b of a ratio, then the ratio's desirability at c is better than at any smaller chromatic cardinality (Ref. 6). Remarkably, for the sequence of ratios 3:5:7:9, $c=13$ is a denominator of continued fraction convergents to \log_3 of all three ratios: $\log_3(5/3) \approx 6/13$, $\log_3(7/5) \approx 4/13$, and $\log_3(9/7) \approx 3/13$. This coincidence explains why this small chromatic cardinality has such an extraordinarily good desirability.
- ³⁵H. Bohlen, "13 Tonstufen in der Doudezeme," *Acustica* **39**, 76–86 (1978).
- ³⁶L. Roberts and M. Mathews, "Intonation Sensitivity for Traditional and Nontraditional Chords," *J. Acoust. Soc. Am.* **75**, 954–959 (1984).
- ³⁷M. Schechter, "Tempered Scales and Continued Fractions," *Am. Math. Monthly* **87**, 40–42 (1980).
- ³⁸J. Douthett, R. Entringer, and A. Mullhaupt, "Musical Scale Construction: The Continued Fraction Compromise," *Util. Mathemat.* **42**, 97–113 (1992).
- ³⁹A. Ya Khinchin, *Continued Fractions* (The University of Chicago Press, Chicago, 1992).

Optimization of a low-frequency ultrasonic technique to monitor the change in physical states in viscoelastic media: Gelation process

G. Nassar^{a)} and B. Nongaillard

I.E.M.N.-U.M.R. C.N.R.S. 9929-D.O.A.E.—Université de Valenciennes, B.P. 311-59313 Valenciennes Cedex 09, France

Y. Noël

INRA-Station de Recherches en Technologie et Analyses Laitières, B.P. 89-39801 Poligny Cedex, France

(Received 19 April 1999; accepted for publication 8 February 2000)

A new low-frequency ultrasonic device (50–100 kHz) in highly sharpened end sensors that behave as point sources were examined. The application of this new ultrasonic technique with two sensors coupled in the near field is to explore the relations between the physical properties measured through the evolution of the wave time of flight and structural changes during gel formation which is related to two factors: the ambient temperature and the mechanical resistance of the medium. The network evolution was interpreted by an approach based on the Flory model. The physical significance of this model was shown through a series of experiments using a low-frequency ultrasonic technique. Response curves demonstrate the different stages during gel formation. © 2000 Acoustical Society of America. [S0001-4966(00)03405-6]

PACS numbers: 43.80.Ev, 43.80.Jz [FD]

INTRODUCTION

The analysis of the different stages in the formation of macromolecular networks is of major importance, since understanding the structure and properties (physical or chemical) of gels requires the understanding of their organization process. In many physical, chemical, or biological processes, the union of small separate elements to form aggregates of different sizes and further macroscopic phases makes connectivity an essential characteristic of this type of process. Many models have been proposed to explain the aggregation phenomenon. The most important ones are those of Flory¹ and Stockmayer,² Case,³ Gupta,⁴ Eichinger,⁵ Allsopp,⁶ and San Biagio,⁷ describing the phenomenon in most cases by the classical theory as a “particular case of percolation” and the bidimensional growth of the network according to Cayley’s tree. Some other works like those of De Gennes⁸ and Stauffer^{9,10} describe the random aggregation phenomenon and the percolation and gelation problems. However, the different characteristics of the macromolecular chain-making system may be evaluated according to Clerc,¹¹ with simulation methods like Monte Carlo, foreseeing the influence of the characteristics of the starting solution and the gelation conditions on the structure and the arrangement of masses.

Moreover, the gelation process can be defined as a transition phenomenon in which a soluble suspension made of macromolecules (liquid phase) becomes insoluble when a giant mass forms.¹² Below the gelification threshold, the medium viscosity increases as the conversion rate p approaches the critical advancement rate p_c . This phenomenon is known as a critical connectivity transition.¹ After the threshold, the

medium ceases to flow and the gel phase develops some elasticity.

Basic macromolecules are synthesized by linking the monomer molecules (building units) with covalent bonds. This chemical reaction of the monomers is due to the presence of functional groups within their structure able to form chemical bonds with the functional groups of other molecules of the monomer.¹³ This phenomenon introduces structural changes in the physical properties, especially the mechanical behavior of the medium. These changes result in the transition from a liquid state to a viscoelastic solid state.

To study this phenomenon, several physics measurement techniques, i.e., optical, thermal, rheological, and acoustic, exist.¹⁴ Sampling and sensitivity to a limited range of physical properties of matter are often drawbacks. As a consequence, different techniques are required to explore a whole process with difficulty of bringing together the heterogeneous data provided by these techniques. There is, for example, the case of the optic methods, which is penalized by the opacity of the analyzed substances as well as the size of the formed molecules in relation to the wavelength. The thermal methods are insensitive to the mechanical features of the medium. The fragility of some gel (milk gel) in formation limits the rheological technique. The acoustic means, i.e., 10 MHz, has been used in extreme conditions reserved exclusively for laboratory use. In many cases, several analytical techniques exist, but they are only used in the laboratory.

To develop further instrumentation in order to understand and to quantify the process of changing media in a real conditions, a new low-frequency ultrasonic technique with highly sharpened end sensors that behave as point sources was examined. The application of this new ultrasonic technique located in two sensors coupled in the near field was used to explore the relations between the physical properties

^{a)}Electronic mail: gnassar@univ-valenciennes.fr

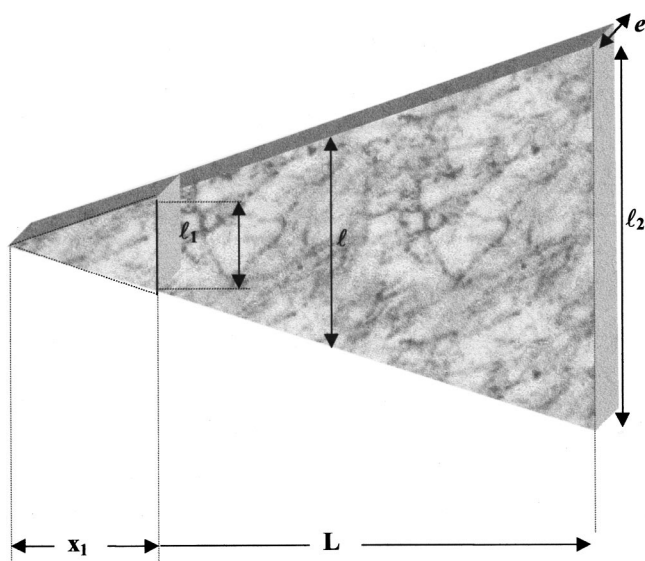


FIG. 1. Physical profile of a ceramic triangular resonator: $L=32$ mm, $l_1=1$ mm, $l_2=7$ mm, and $e=1$ mm. Sound velocity in the material: $c=4300$ m/s.

measured through the evolution of the wave time of flight and structural changes during gel formation.

I. STUDY OF A NEW LOW-FREQUENCY ULTRASONIC DEVICE

The study is aimed at defining and achieving an optimized ultrasonic instrumentation to understand the phenomenon and quantify the viscoelastic properties of changing media.

The usual ultrasonic characterization techniques are generally based on the use of a resonant piezoelectric transducer on a thickness mode. As the resonant frequency of a transducer is inversely proportional to its size, it increases for low frequencies around 100 kHz. Some researchers like Degertekin,^{15–17} Shuyu,^{18,19} Nikolovski,²⁰ and Madigosky²¹ used this physical principle, but associated a tapered volume with the ceramic components to concentrate the mechanical energy.

Our aim was to obtain a low-frequency acoustic point source to generate a spherical wave in the medium. To do this a different procedure from the classic sensor design was used, i.e., a new technique which consisted of setting in resonance all the mechanical structure of a reduced-size unit, with a highly sharpened end in contact with the material to be analyzed. The end size was smaller than the wavelength in the medium to behave like a point acoustic source.

The first part of the work presents a theoretical analysis of low-frequency ultrasonic resonators, beaming a spherical wave in the medium. The choice of a triangular-shaped resonator and its mechanical behavior will be assessed and the study completed by a numerical approach based on the application of the finite-elements method to characterize all the resonator vibration modes and visualize the corresponding distortions when the structure is excited. As the analytical results were in good agreement with the numerical results, they were applied to the whole triangular-shaped sensor to validate the findings experimentally. A correspondence was

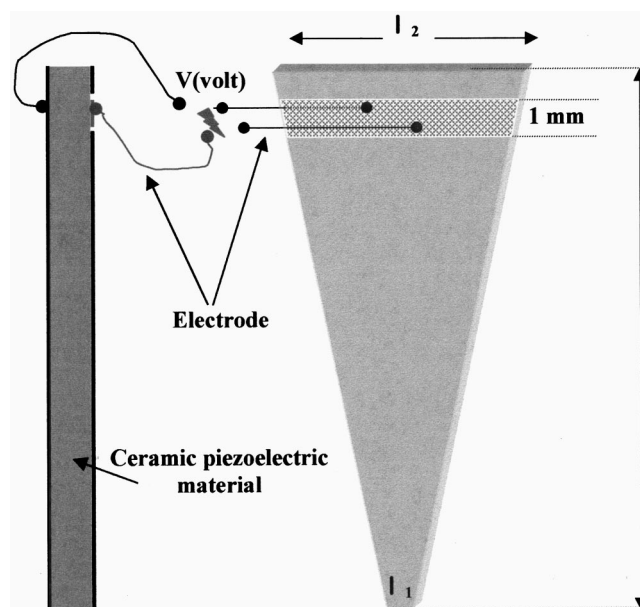


FIG. 2. Architectural structure: source of 1-mm width engraved on the metallic surface (electrode).

then made between the resonance mode frequencies determined by the numerical calculation and the measured electrical impedance.

A. Analytical study

The structure of an ultrasonic sensor based on a 1-mm-thick triangle shape is presented (Fig. 1).

The propagation of longitudinal waves in the triangular part of the sensors was studied to determine the resonance frequency of longitudinal modes and amplification ratio of velocity. Our analysis is based on an extension of Ensminger's²² theory.

According to Fig. 1, the section on the x axis can be written as follows:

$$S = e \cdot l(x). \quad (1)$$

Then

$$S = e \cdot l_1(x_1 + x)/x_1. \quad (2)$$

The differential equation is

$$\frac{\partial^2 v}{\partial x^2} + \frac{1}{(x_1 + x)} \frac{\partial v}{\partial x} + \frac{\omega^2}{c^2} v = 0. \quad (3)$$

The solution of Eq. (3) (see the Appendix) gives the velocity ratio: $v(0)/v(L) = -2.16$ for a resonance frequency $f_1 = 60$ kHz.

B. Numerical study

While an analytical study can only consider a single particular mode of vibration of a triangular part of the sensor, a numerical study based on the finite-elements method²³ can determine all the vibrating modes of these parts as well as those of the whole sensor.

For all the parts of the structure (pavilion, binding rod,...), the displacement differential equations were solved with a continuous regime, taking into account the boundary

TABLE I. The first 30 natural modes of the triangular resonator given by the finite-elements method.

Mode	Frequency (Hz)	Mode	Frequency (Hz)	Mode	Frequency (Hz)
1	752	11	28 474	21	67 241
2	4 112	12	35 437	22	78 013
3	7 742	13	36 204	23	82 952
4	8 926	14	39 583	24	87 627
5	9 965	15	43 704	25	90 308
6	15 405	16	52 136	26	98 037
7	17 262	17	53 021	27	102 000
8	21 887	18	54 589	28	104 000
9	23 190	19	58 142	29	107 000
10	24 734	20	66 012	30	112 000

conditions at the surfaces. The materials were defined by a Young modulus E , Poisson's coefficient ν , and density ρ . The results presented below were applied without loss and they were compared to the characteristics of the longitudinal mode determined elsewhere by analytical calculation. This comparison was also made for the triangular sensor which was studied as a whole.

For our study, the ANSYS 5.2 analysis software was used.

The sensors used were made mainly of piezoelectric material. A source of excitation was engraved in the general structure of the vibrating element (triangle part), giving a mechanical continuity without rupture (Fig. 2). This type of engraving was considered as it has been demonstrated¹⁴ that for the same longitudinal mode, the amplification ratio of the ends is 1.7 times bigger when there is one engraved source providing mechanical continuity with the vibrating element so that one source can impact the structure by gluing. The study of the sensors was carried out with isotropic materials

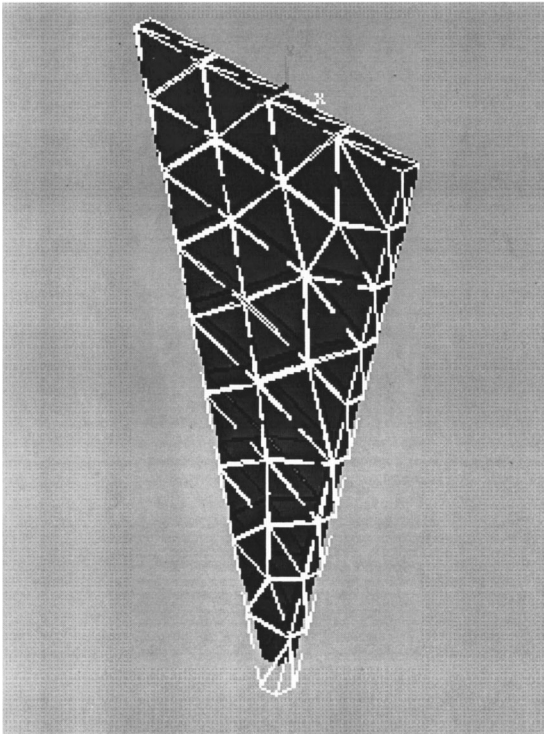


FIG. 3. Identification of the longitudinal mode associated with the frequency $F_L=58$ kHz.

TABLE II. Analytical and numerical illustration of the resonance frequency (longitudinal mode) and the velocity ratio to the ends of the resonator.

	Frequency of longitudinal resonance (kHz)	Ratio of velocity V_1/V_2 between the ends
Analytical approach	60	-2.16
Numerical approach	58	-2.32

and did not take into account the electromechanical coupling function of the piezoelectrical element.

A first modal analysis was carried out with only the triangular part of the sensor.

Table I gives the natural modes of the triangular structure where some simple modes were identified. Void displacements were set on the widest bases of the structure.

Figure 3 shows the distortion of the structure associated with the longitudinal mode.

Table II shows the resonance frequencies and the transformation velocity ratio for the longitudinal mode.

C. Experimental validation

A modal study was carried out for the whole sensor. This sensor included a triangular truncated part whose electrical and mechanical characteristics were identical to those used in the previous part. A binding rod was added (Fig. 4).

Measurement of the electrical impedance was used to check that the modes of calculated resonance were excited by the mechanical distortion produced by the piezoelectrical

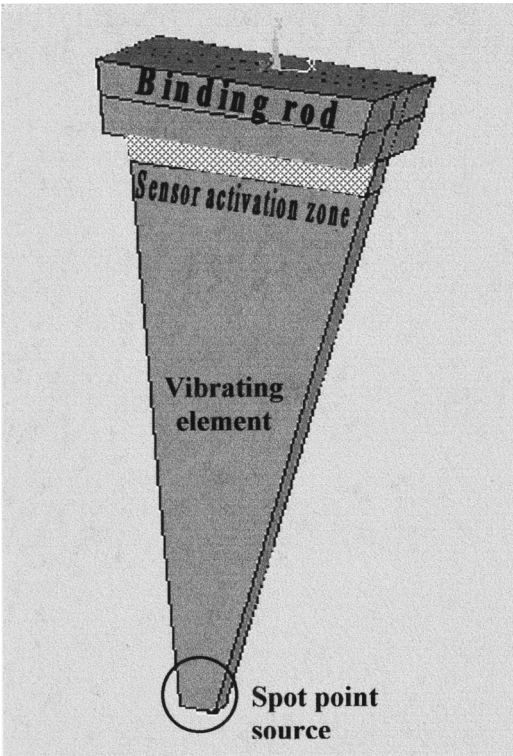


FIG. 4. Ultrasonic sensor designed with the same characteristics (electrical and mechanical) as the triangular resonator, but a brass binding rod was added.

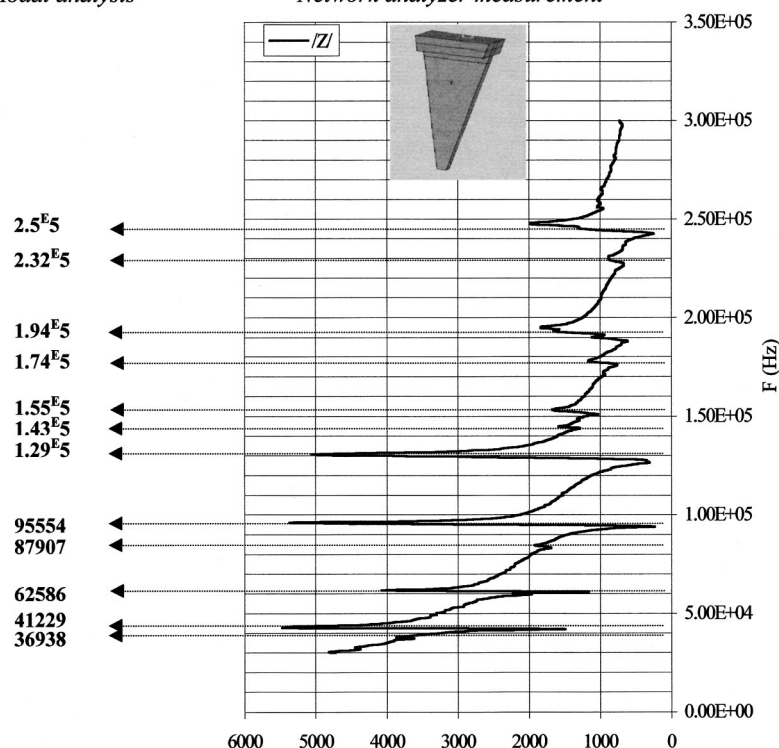


FIG. 5. Schematic illustration of resonance frequencies of the whole sensor: Left: by application of the finite-elements method (modal analysis); Right: by impedance measurement with a network analyzer.

effect. Figure 5 illustrates the changes in the electrical impedance of the sensor measured with a network analyzer. The frequencies of the vibration modes determined by the modal analysis are also labeled to the left of the curve.

There was agreement between the calculated frequencies and those determined by the impedance measurements. This correspondence justifies the approximation that had to be made for the numerical study of the sensor with the finite elements method using ANSYS software, i.e., neglecting the electromechanical coupling.

II. APPLICATION

A. Analysis support: Milk gelation process

The milk gelation process can be considered as an aggregation of different molecules with different size.^{24–26} It was explored as a model for several reasons: available knowledge, experimental conditions known and relatively easy to perform, complex medium with physical properties of liquid and gel states close together.

As the reaction progresses, the average mass of each aggregate increases while the number of molecules in the tested medium decreases. The aggregation process results in a giant macromolecule defining the gel.

B. Materials and method

1. Reconstituted skim milk samples

Milk samples were prepared according to the conditions established by INRA (Poligny, France).

Twelve grams of medium-heat skimmed-milk powder (Ingredia, France) were dissolved in 100 ml of distilled water at ambient temperature. The mixture was stirred for 15 min, 1 ml CaCl_2 1 M was added, and the solution was stirred

again for 15 min. After resting 30 min, the reconstituted skim milk was heated to 30 °C, then 30 μl of rennet were added (520 mg/l of chymosine, Granday, France).

2. Ultrasonic measurements

The milk gelation process was examined using two identical ultrasonic sensors (Fig. 4).

These transducers are coupled to the medium to be characterized, in their near field, and are totally submerged in the medium. The working frequency is 60 kHz.

Figure 6 gives a schematic diagram of the measuring system. The emitter driven by a sharp electrical pulse, of 15 μs duration. These conditions provide a longitudinal vibration mode at the end of the sensor which behaves like a spot point source. This phenomenon generates a divergent ultrasonic wave in the medium, one part of which was measured by a receiver located at a constant distance from the transmitter.

The characteristics of the propagation wave in the medium are more or less a compressional wave, as suggested by the time of flight corresponding to a velocity of 1600 m/s in reconstituted milk samples at 25 °C.

The evolution of the wave time of flight was observed in order to follow the changes in the different steps of the physical change of the product during gelation.

C. Results

1. Measurement stability

Figure 7 give the variation of time of flight dt (ns) of the signal, measured with the ultrasonic apparatus, in distilled

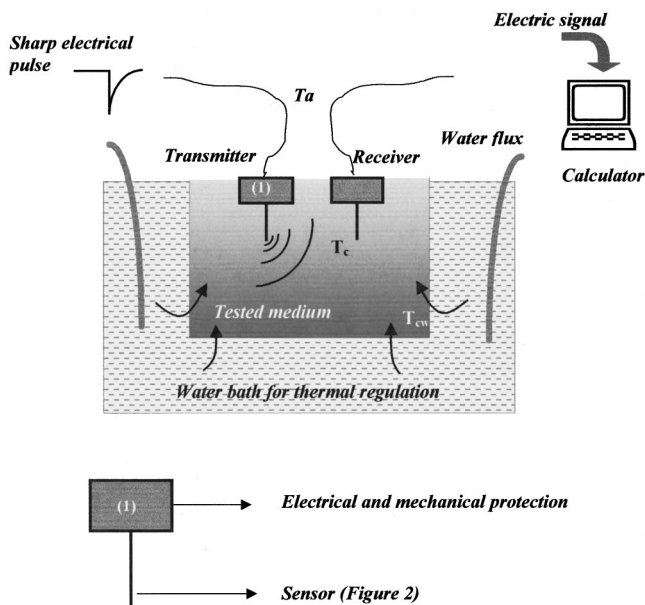


FIG. 6. Diagram of the measuring cell, T_c : temperature of the product at "sensor level;" T_{cw} : temperature of the container walls; T_a : ambient temperature.

temperature-controlled water at $30 \pm 0.1^\circ\text{C}$. The results demonstrated the stability of the ultrasonic measurement in a reference medium.

The precision obtained on the measure of time of flight was 1 ns over a global reply time of $10\ \mu\text{s}$, giving a relative precision of 10^{-4} per measurement at the zero-crossing point (Fig. 8).

2. Measurement reliability

The reliability of a measurement system resides in its reproducibility and its faculty to follow the gelation process through all its stages. In the standard conditions previously indicated and in accordance with the literature,²⁷ transition sol gel clotted between 15 and 16 min. At a 30°C ambient temperature, as shown in Fig. 9, the time of flight of signal decreased, indicating an increase in the mechanical resistance of the product. This variation had not reached a plateau value, showing that the medium was still changing.

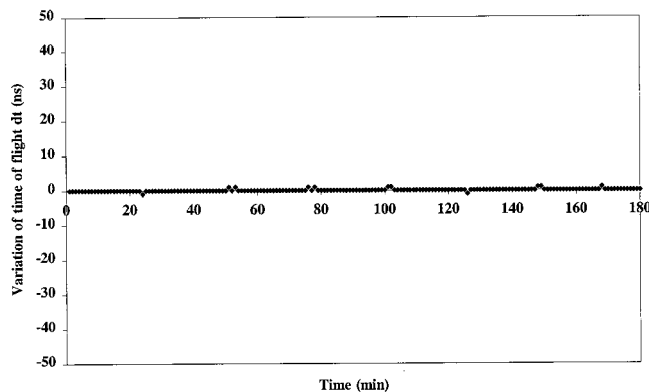


FIG. 7. Illustration of time stability of the ultrasonic signal with the system used. Measurement in temperature-controlled distilled water.

Figure 9 also gives a qualitative estimation of the reproducibility of the ultrasonic measurements. The curves show the progress of the rennet action for two media prepared in the same conditions. The maximum dispersion of the measurements was 5 ns due to the electronic parts and the milk reconstitution process.

3. Gelation process

Figure 10 draws the profile of the variations of time of flight resulting from the variation of the ultrasonic wave velocity during the milk gelation process at different ambient temperature.

The curve observed at an ambient temperature of 30°C , similar to the milk temperature, had a pattern similar to those already known from measurements with other physical methods.²⁸ When the ambient temperature was significantly different from the product temperature, ultrasonic curves had a particular pattern.

III. DISCUSSION

A. Evolution of the network

The gelation process is a transition from an entirely soluble system to a heterogeneous system of two phases: one composed of an insoluble entity (infinite-size macromolecule) and a soluble phase. This transition is accompanied by radical changes in some physical properties of the medium. Below the gelification threshold, the medium viscosity increases as the conversion rate p approaches the critical advancement rate p_c . After the threshold, the medium ceases to flow and develops elasticity.

To illustrate this process schematically, let us consider an initial solution containing units that can link with each other. The network formation occurs if the functionality of the units is greater than 2.¹³ As the reaction progresses, the connectivity status of the system is characterized by the conversion rate p defined as the rate of formed links.^{1,2} At the beginning, the medium behaves like a viscous solution in a sol phase due to the presence of a single type of finite size masses: in this case, p is low [Fig. 11(a)]. When p increases, bigger and bigger masses are formed [Fig. 11(b)]. For a certain critical value of p, p_c , a giant chain appears [continuous connectivity of the space from one side to the other $A \leftrightarrow B$; Fig. 11(c)] defining the gel point. Beyond the threshold p_c , the medium in the gel phase has the macroscopic behavior of a viscoelastic solid. For $p=1$ all the units belong to the giant mass [Fig. 11(e)].

As the properties of a gelling medium are proportional to the degree of progress of the reaction, it is possible to represent the behavior of the viscoelasticity in terms of the rate of progress of the macromolecular connectivity (Fig. 11) according to the following cases:

- For $p < p_c$, the system is a liquid whose viscosity increases as the gel point approaches.
- For $p = p_c$, an elastic behavior appears.
- For $p > p_c$, the medium becomes a solid gel whose elasticity increases with p .

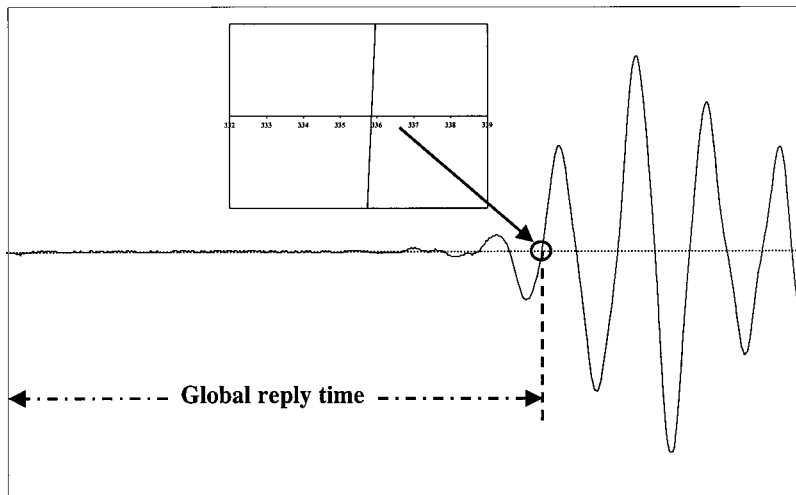


FIG. 8. Signal form in reception indicates the zero-crossing point measurement and the reply global time.

In order to relate the theoretical aspect to the experimental results, the following curve (Fig. 12) divided into five different stages is examined below. The phenomenon describes the gelation process when the ambient temperature was different from the product temperature. It was not detectable when these temperatures were equal.

A quantification of the variation of the different values characterizing the curve of Fig. 12 presenting the milk gelation process can be given by Table III, where T_a is the ambient temperature, ΔT_c is the variation of the common temperature in the medium at the level of the sensors between the renneting time ($t=0$) and the time $t=120$ min, dt_{pw} is the variation of the wave time of flight in pure water, measured at the same temperature conditions as for gelling milk. It is related to ΔT_c ; dt_i is the variation of the time of flight before stage 4; dt_f is the variation of the time of flight after stage 4.

According to Fig. 12, ‘stage 1’ could be interpreted as a proteolytic phase characterized by the appearance of two polypeptides resulting from the effect of the rennet product (chymosin) on casein κ (milk protein).²⁹ The lag time associated with this phase depends on the coagulation conditions. This stage is followed by the association and aggregation of casein micelles characterized by the formation of aggregates of finite size [Fig. 11(b)]. This molecular reorganization

might be related to the change of slope of the curve (stage 2) reflecting a decrease in the time of flight, which means an increase in the viscosity in the sol medium. The temperature inside the medium remained constant ($T_c=30.1^\circ\text{C}$) during this stage and resulted (at that moment) from the propagation of heat by pure and free convection. The formation of more or less voluminous masses in the medium induced the transition from a viscous state to a viscoelastic state, slowing down the free convection. This led to a slight temperature decrease in the medium (stage 3) at the level of the sensors (T_c) to reach a new equilibrium where heat was mainly transmitted through the container walls (regulated temperature: $T_{cw}^\circ\text{C}$), by conduction. The changes in the medium during stage 3 could be interpreted in the following manner.

- Due to a thermal conduction phenomenon, a slight temperature gradient appears in the medium, between the container walls ($T_{cw}=30.1^\circ\text{C}$) and the center of the vat, at the location of the measuring point; $T_c^\circ\text{C}$ (following the ambient temperature). The temperature decrease (Table III) induced an increase of the time of flight.
- The time of flight decreased as the reaction progressed. This decrease can be attributed to the development of an elastic modulus resulting from the formation of

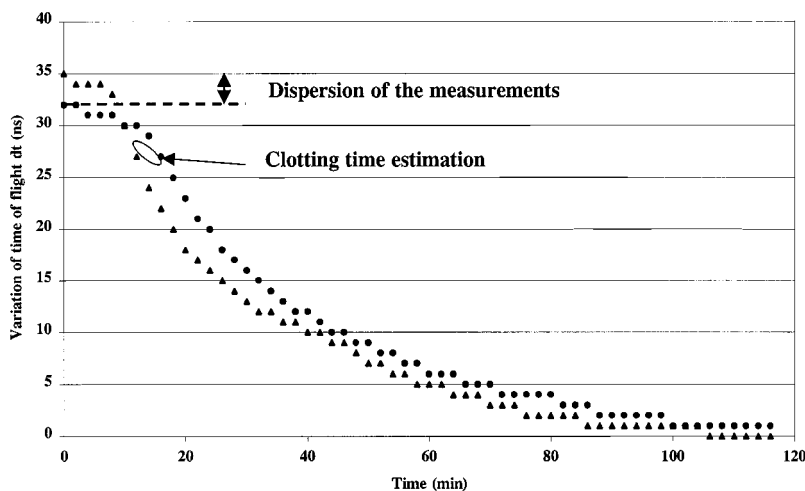


FIG. 9. Typical curves of gelation process at 30.1°C as temperature of reaction viewed by the ultrasonic technique of two media prepared according to the protocol indicated in Sec. II B 1. The difference between the two curves gives a quantitative idea of the global dispersion of the ultrasonic measurement due to the electronics and preparation of sample.

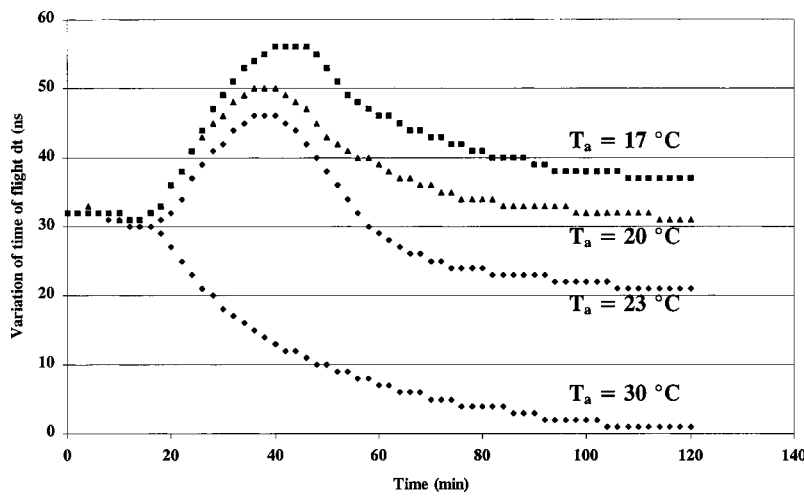


FIG. 10. Ultrasonic sensor responses during milk gelation at 30.1 °C, but in different ambient temperatures, from 17 to 30 °C.

macromolecules, changing the medium from a viscous liquid state to a viscoelastic solid state. The phenomenon was expressed physically by the evolution of the connectivity rate \mathbf{p} towards its critical value \mathbf{p}_c .

These two phenomena make stage 3 a competition between:

- (1) An increase of the time of flight resulting from a decrease in temperature.
- (2) A decrease in the time of flight resulting from the appearance of an elastic component in the changing medium.

This competition can be observed in Table III through the differences between the values expressing the variation of time of flight in water, dt_{pw} (ns) measured in the same

conditions compared to those observed in renneted milk: dt_i (ns), supporting our assumption.

The connectivity rate \mathbf{p} reached a critical value \mathbf{p}_c in stage 4, at the maximum of the curve when the existence of a giant macromolecular chain linking the two extreme sides of the considered space can be assumed [Fig. 11(c)]. During this stage, the mechanical aspect of the medium could dominate the remaining part of the reaction, due to the weak thermal variation (≈ 1 °C) resulting from the difference in temperature between the medium and the environment (ambient temperature).

During stage 5, the gel strengthened. The gel was stronger when the ambient temperature was very close to the temperature medium (Fig. 10, Table III). This phenomenon can be shown experimentally by an increase in Δdt (ns) resulting from a decrease in time of flight, whereas theoretically it was explained by the evolution of the connectivity rate \mathbf{p} towards 1 following the establishment of continuous connections of finite-size masses on the giant molecule, the gel.

IV. CONCLUSION

This work evaluated a low-frequency ultrasonic sensor and a method for studying the gelation process. A strategy was developed to improve the mechanical features of the sensor based on analytical and numerical methods, as well as experimental validation. This mechanical improvement offers the possibility of easily selecting the measuring frequency over a wide range according to the medium to be tested. This sensor could be used in industrial conditions for

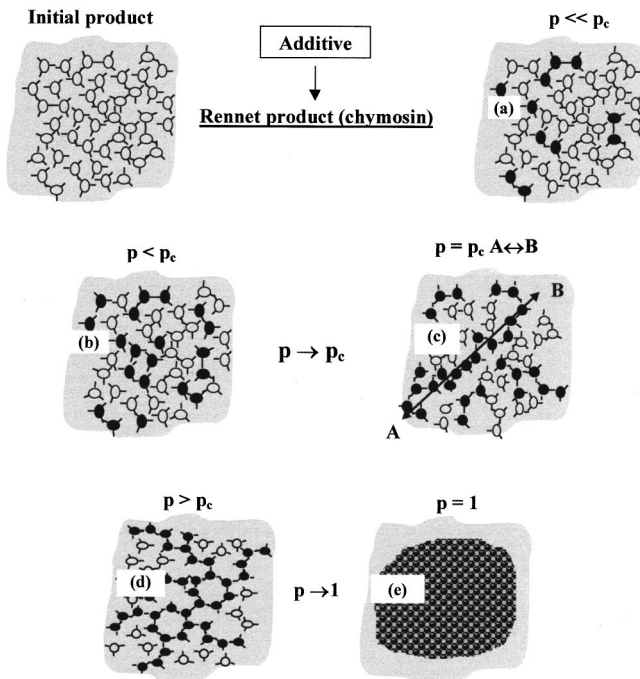


FIG. 11. Schematic representation of the sol-gel transition, initial phase: (a) Suspension of molecules of finite sizes; (b) Agglomeration and formation of macromolecules of large masses; (c) Critical phase connection: $p = p_c$; (d) Network continuity connection; (e) A single giant macromolecule, the gel.

TABLE III. The different values of the physical parameters characterizing the curve of the gelation process when the ambient temperature was different from the product temperature.

Common initial milk temperature	$T_c = 30.1$ °C		
Distance between sensors	$l = 17$ mm		
T_g °C	17	20	23
ΔT_c °C	2	1.6	1.2
dt_{pw} (ns)	36	29	22
dt_i (ns)	24	18	14
dt_f (ns)	19	19	25
$\Delta dt = (dt_f - dt_i)$ (ns)	-5	1	9

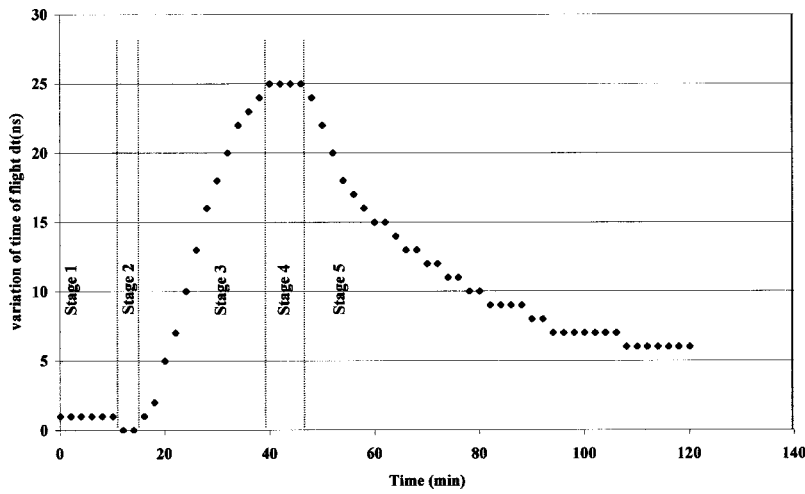


FIG. 12. The specific ultrasonic response during milk gelation when the ambient temperature was different from the milk temperature.

the characterization of viscoelastic media (biological and food industry) and avoids coupling problems with the medium being tested.

By using the longitudinal mode vibration in some very simple conditions, a coherent view of the gelation mechanism and the formation of a macromolecular network were presented. This mechanism was assessed by combined measurements of the physical characteristics resulting from a change in state, expressed by a variation in the elastic and heat conduction properties induced by this change.

The experimental curves showed all the stage of gel formation which, until now, could not be done by other physical or chemical techniques. Compared to other existing techniques, the gel elastic properties can be evaluated with this low-frequency ultrasonic method. A combination of a thermal gradient at the medium surface with the regulated temperature in the tested medium opens new perspectives, particularly in an industrial environment.

Due to the simplicity of the method, it can be applied to the study of physical state change of other viscoelastic complex products found in industry.

APPENDIX

According to Fig. 1, the section on the x axis can be written as follows:

$$S = e \cdot l(x). \quad (\text{A1})$$

Then

$$S = e \cdot l_1(x_1 + x)/x_1. \quad (\text{A2})$$

The differential equation is

$$\frac{\partial^2 v}{\partial x^2} + \frac{1}{(x_1 + x)} \frac{\partial v}{\partial x} + \frac{\omega^2}{c^2} v = 0. \quad (\text{A3})$$

By a changing variable: $z = k(x_1 + x)$ with $k = \omega/c$, the differential equation (A3) can be reduced to the Bessel function of 0 order

$$\frac{\partial^2 v}{\partial z^2} + \frac{1}{z} \frac{\partial v}{\partial z} + v = 0. \quad (\text{A4})$$

A general solution to Eq. (A4) can be written as follows:

$$v(z) = AJ_0(z) + BY_0(z), \quad (\text{A5})$$

$J_0(z)$ and $Y_0(z)$ being, respectively, the first and second kind of zero-order Bessel functions. A and B are two constants determined from the boundary conditions.

Equation (A5) can be expressed as a function of x by

$$v(z) = AJ_0(k(x_1 + x)) + BY_0(k(x_1 + x)), \quad (\text{A6})$$

the first derivative of which is the following:

$$v'(z) = -AkJ_1(k(x_1 + x)) - BkY_1(k(x_1 + x)), \quad (\text{A7})$$

J_1 and Y_1 are, respectively, the first and second kind of first-order Bessel functions.

For $v'(0) = v'(L) = 0$,

$$-AkJ_1(k(x_1)) - BkY_1(k(x_1)) = 0$$

and

$$-AkJ_1(k(x_1 + L)) - BkY_1(k(x_1 + L)) = 0. \quad (\text{A8})$$

$$-AkJ_1(k(x_1 + L)) - BkY_1(k(x_1 + L)) = 0.$$

The first equation of (A8) expresses B as a function of A

$$B = -A \frac{J_1(kx_1)}{Y_1(kx_1)}. \quad (\text{A9})$$

By adding B to the second equation (A8), the following is obtained:

$$-J_1(k(x_1 + L))Y_1(kx_1) + Y_1(k(x_1 + L))J_1(kx_1) = 0. \quad (\text{A10})$$

The resonance frequency is such that the Eq. (A10) in k is respected.

From the geometrical properties of the triangular structure, the following can be deduced:

$$l_2/(L + x_1) = l_1/x_1 \Rightarrow x_1 = 4.6 \text{ mm}.$$

Equation (A10) is true when: $k_1 = 0.11 \text{ mm}^{-1}$ (first solution).

The frequencies of multiple resonance of the longitudinal mode for a given length L can be calculated from the k values.

The velocity transformation ratio is expressed by $v(0)/v(L)$.

From Eqs. (A5) and (A9), the following is obtained:

$$v(0) = AJ_0(kx_1) - AY_0(kx_1) \frac{J_1(kx_1)}{Y_1(kx_1)}. \quad (\text{A11})$$

and

$$v(L) = AJ_0(k(x_1 + L)) - AY_0(k(x_1 + L)) \frac{J_1(k(x_1))}{Y_1(k(x_1))}. \quad (\text{A12})$$

The first mode of the velocity ratio is: $v(0)/v(L) = -2.16$ for a resonance frequency $f_1 = 60$ kHz.

In the case when the triangular part is complete ($x_1 = 0$), the general solution can be written as follows:

$$v(z) = CJ_0(z), \quad (\text{A13})$$

giving the following derivative:

$$v'(z) = -CkJ_1(z), \quad (\text{A14})$$

$v'(0) = 0$ is automatically true. $v'(L) = 0$ gives $z = kL = 3.8$, thus, $v(0) = C$ and $v(L) = -0.4C$.

Finally, the transformation ratio of velocity is: $v(0)/v(L) = J_0(0)/J_0(4) \approx -1/0.4 = -2.5$ for $k_1 = 0.11$; then for a resonance frequency, $f_1 \approx 60$ kHz.

¹P. J. Flory, *Principles of Polymer Chemistry* (Cornell University Press, Ithaca, NY, 1953).

²W. H. Stockmayer, "Theory of molecular size distribution and gel formation in branched-chain polymers," *J. Chem. Phys.* **11**, 45–55 (1943).

³L. C. Case, "Molecular distributions in polycondensations involving unlike reactants. VII. Treatment of reactants involving nonindependent groups," *J. Polym. Sci.* **48**, 27–35 (1960).

⁴S. K. Gupta, A. Kumar, and A. Bhargava, "Molecular weight distribution and moments for condensation polymerization of monomers having reactivity different from their homologues," *Polymer* **20**, 305–310 (1979).

⁵B. E. Eichinger, "Random elastic networks. I. Computer simulation of linked stars," *J. Chem. Phys.* **75**, 1964–1979 (1981).

⁶M. W. Allsopp, "The development and importance of suspension PVC morphology," *Pure Appl. Chem.* **53**, 449–465 (1981).

⁷P. L. San Biagio, D. Bulone, A. Emanuele, F. Madonia, L. Di Stefano, D. Giacomazza, M. Trapanese, M. B. Palma-Vittorelli, and M. U. Palma, "Spinodal demixing, percolation and gelation of biosttural polymers," IUPAC 10th International Symposium on Polymer Networks, 40, Jerusalem, December, 1990.

⁸P. G. De Gennes, *Scaling Concepts in Polymer Physics* (Cornell University Press, Ithaca, NY, 1989).

⁹D. Stauffer, "Can percolation theory be applied to critical phenomena at gel point?," *Pure Appl. Chem.* **53**, 1479–1487 (1983).

¹⁰D. Stauffer, *Introduction to Percolation Theory* (Taylor & Francis Ltd., London, 1985).

¹¹J. P. Clerc, G. Giraud, J. Rousseng, R. Blanc, J. P. Carton, E. Guyon, H. Ottavi, and D. Stauffer, *La Percolation: Modèles, simulation analogiques et numériques*, Annales de Physique Vol. 8 (Masson, Paris, 1983).

¹²D. G. Dalgleish, *Developments in Dairy Chemistry*, edited by P. F. Fox (Applied Science, London, 1982), Vol. 1, Chap. 5.

¹³J. P. Mercier and E. Marechal, *Chimie des Polymères* 1st ed. (Presses Polytechniques et Universitaires Romandes, Lausanne, 1993), Chaps. 1, 3, 8.

¹⁴G. Nassar, "Etude et Optimisation d'un Dispositif Ultrasonore De Suivi en Ligne des propriétés viscoélastiques," Doctoral dissertation, Valenciennes University-France, December 1997.

¹⁵F. L. Degertekin and B. T. Khuri-Yakub, "Hertzian contact transducers for nondestructive evaluation," *J. Acoust. Soc. Am.* **99**, 299–308 (1996).

¹⁶F. L. Degertekin and B. T. Khuri-Yakub, "Single mode lamb wave excitation in thin plates by Hertzian contacts," *Appl. Phys. Lett.* **69**(2), 146–148 (1996).

¹⁷F. L. Degertekin and B. T. Khuri-Yakub, "Lamb wave excitation by Hertzian contacts with applications in NDE," *IEEE Trans. Ultrason. Ferroelectr. Freq. Control* **44**(4), 769–778 (1996).

¹⁸L. Shuyu, "Study on the longitudinal-torsional composite mode exponential ultrasonic horns," *Ultrasonics* **34**, 757–762 (1996).

¹⁹L. Shuyu, "Study on the longitudinal-torsional composite vibration of a sectional exponential horn," *J. Acoust. Soc. Am.* **102**, 1388–1393 (1997).

²⁰J. P. Nikolovski and D. Royer, "Local and selective detection of acoustic waves at the surface of a material," *IEEE Ultrasonics Symposium* (1997), 699–703.

²¹W. M. Madigosky, C. T. Swanson, and E. J. O'Neill, "Dynamic hardness tester and cure meter," in *Non-destructive Testing* (Van Hemelrijck & Anastassopoulos, Balkema, Rotterdam, 1996), pp. 169–173.

²²D. Eensminger, "Solid cone in longitudinal half-wave resonance," *J. Acoust. Soc. Am.* **32**, 194–196 (1960).

²³J. F. Imbert, *Analyse des Structures par Éléments Finis*, 3rd ed. (Cepadues, 1995).

²⁴P. Walstra and V. Vliet, "The physical chemistry of curd making," *Neth. Milk Dairy J.* **40**, 241–259 (1986).

²⁵P. F. Fox, "Proteolysis during cheese manufacture and ripening. A review," *J. Dairy Sci.* **72**, 1379–1385 (1989).

²⁶D. G. Dalgleish, in *Cheese: Chemistry, Physics and Microbiology, General Aspect*, 2nd ed., edited by P. F. Fox (Chapman & Hall, London, 1993), Vol. 1, p. 69.

²⁷Y. Noel, P. Flaud, and D. Quemada, *Traitement Industriel des Fluides Alimentaires Non Newtoniens*, Tome II, Actes du 2^{ème} Colloque la Baule, 11–13 September 1989, pp. 215–224.

²⁸D. J. McMahon and R. J. Brown, "Enzymic coagulation of caseine micelles," *J. Dairy Sci.* **67**, 919–929 (1984).

²⁹S. Visser, "Proteolytic enzymes and their action on milk proteins. A review," *Neth. Dairy J.* **35**, 65–68 (1981).

One-hydrophone method of estimating distance and depth of phonating dolphins in shallow water

Roland Aubauer,^{a)} Marc O. Lammers, and Whitlow W. L. Au
Hawaii Institute of Marine Biology, University of Hawaii, P.O. Box 1106, Kailua, Hawaii 96734

(Received 7 September 1999; accepted for publication 25 January 2000)

Previous attempts at localizing cetaceans have generally used multiple hydrophone arrays and multichannel recording systems. In this paper, a low-budget localization technique using only one hydrophone is described. The time delays of the signals traveling via the surface and bottom reflection paths to the hydrophone, relative to the direct signal, are used to calculate the distance and the depth of a phonating animal. Only two additional measures, the depth of the bottom and hydrophone, have to be taken. The method requires relatively shallow waters and a flat bottom surface. Echolocating and burst pulsing Hawaiian spinner dolphins (*Stenella longirostris*) at the Waianae coast of Oahu, Hawaii, were localized over different bottom substrates. A tracking range of up to 100 m was achieved. The accuracy of the method is estimated by the total error differential technique. The relative distance estimation error is below 35% and the absolute depth error below 0.7 m, so that the location method is sufficiently precise for examining source levels in our study area. Because of its simplicity, the method ideally complements sound recordings and visual sightings of marine mammals and could lead to a better understanding of the nature and use of click trains by dolphins. © 2000 Acoustical Society of America. [S0001-4966(00)00605-6]

PACS numbers: 43.80.Ka, 43.80.Lb [FD]

INTRODUCTION

Compared to most animals, dolphins possess highly developed hearing and sound production capabilities. Dolphins are able to emit tonal and clicklike sounds from a few hundred Hz to about 160 kHz and can perceive sound in the same frequency range. Furthermore, dolphins possess a sophisticated echolocation system that they use for orientation and prey capture. Therefore, it is not surprising that many studies of these marine mammals focus on their acoustics.

In order to obtain better knowledge about the use and function of emitted signals, it is necessary to determine the position of the sound-producing animal relative to the sound recording hydrophone. The range of the calling dolphin is the key to determining the source level of specific signals. The source level of emitted signals is a fundamental parameter necessary to gain insight about dolphin echolocation and social behavior in the wild.

Previous attempts at localizing cetaceans have generally used multiple hydrophone arrays and differences in time-of-arrival measurements. Hydrophones have been towed from a ship as a linear line array, floated independently with an anchored line, or mounted on the sea bottom. At least four hydrophones in an appropriate arrangement are necessary to determine the three-dimensional location of a calling animal by evaluating the travel time differences of the incoming sound. A fifth hydrophone is necessary to resolve the spatial ambiguity of two possible positions (Spiesberger and Frisrup, 1990). All of these methods use multichannel recording and signal evaluation that require expensive and highly sophisticated technical equipment. Recently, Cato (1998) described three relatively simple methods for estimating source levels and distances of marine animal sounds. Two methods

make uses of arrival time and level differences of the signals received by two separated hydrophones. The third method estimates source range from the levels and arrival times of the direct and surface reflected signals.

This paper describes another technique to localize marine animals based on arrival time difference measurements of the direct, the surface, and the bottom reflected signals. The method requires only one hydrophone and determines the distance and depth of the animal. From the distance estimation the source level of emitted signals can also be estimated. Under certain conditions this new method gives significantly better distance, depth, and source level estimates than the methods of Cato.

I. MULTIPATH PROPAGATION IN SHALLOW WATER

Sound in the dolphin frequency range propagates relatively well under water, due to a low absorption loss of less than 0.04 dB/m for frequencies less than 160 kHz and water temperatures less than 25 °C. At the air–water interface and at the water–bottom interface sound is reflected relatively well, because of the large differences in acoustic impedance Z at the interfaces. The reflectivity R is related to the acoustic impedance Z of the substance and of water Z_w ,

$$R = \frac{Z \cdot \sin \varphi_w - Z_w \cdot \sin \varphi}{Z \cdot \sin \varphi_w + Z_w \cdot \sin \varphi}, \quad (1)$$

where φ_w is the incident and reflected angle and φ the transmitted angle, and both angles are related by Snell's law,

$$\frac{\cos \varphi_w}{\cos \varphi} = \frac{c_w}{c}, \quad (2)$$

where c_w is the sound velocity in water and c is the sound velocity in the other medium (see Fig. 1). Typical values of the impedance and reflectivity of various substances in sea water under the assumption of normal sound incident are

^{a)} Author to whom correspondence should be addressed: Mussumer Kirchweg 174, D-46395 Bocholt, Germany.

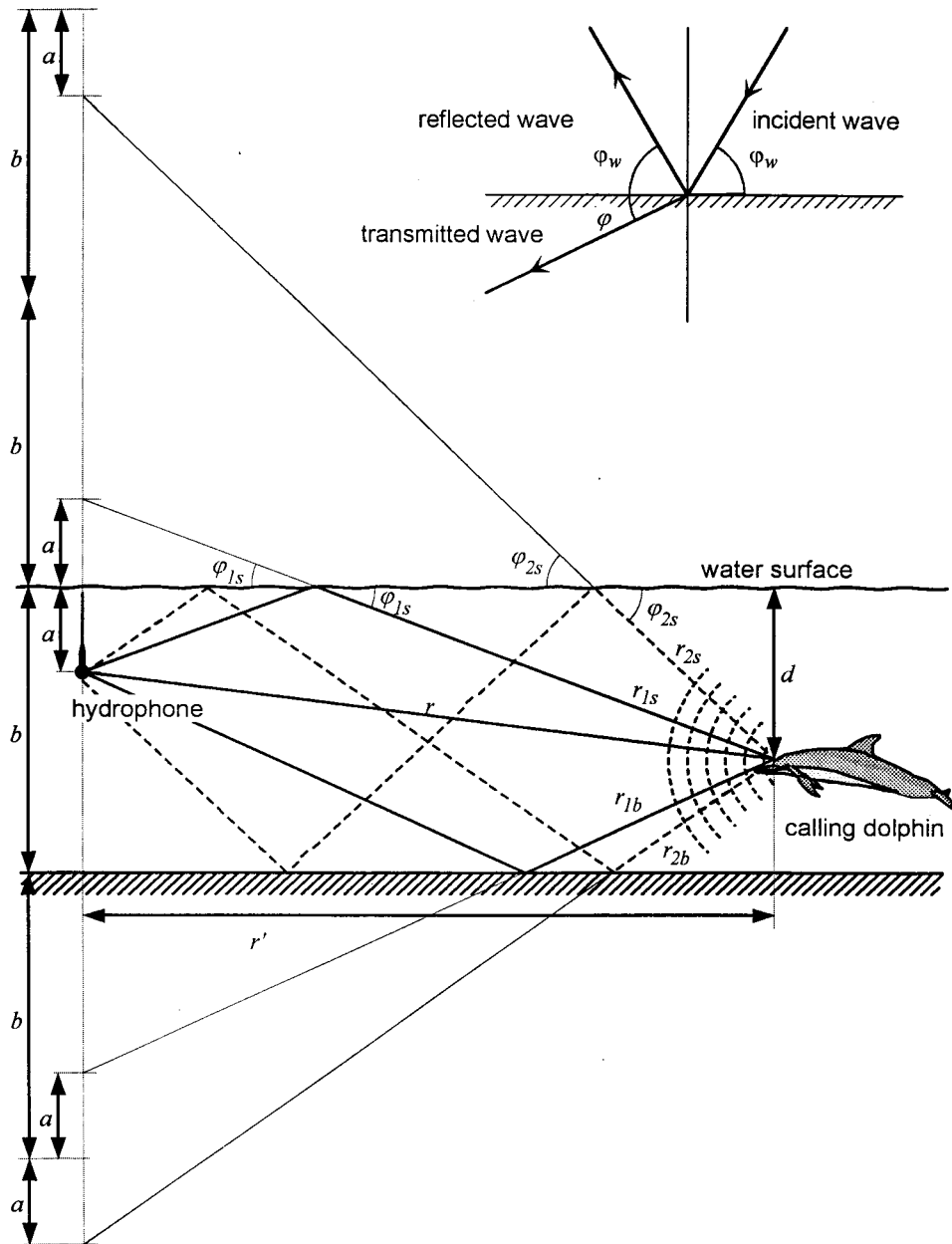


FIG. 1. Multipath propagation in shallow water with a flat ground surface. (—) Direct and first order ($n=1$) propagation paths. (---) Second order ($n=2$) propagation paths. Thin and straight lines starting at the calling dolphin and ending above the water surface and below the sea ground illustrate the construction of Eqs. (4)–(7). The lengths of these lines correspond to the lengths of the shown first and second order propagation r_{1s} , r_{1b} , r_{2s} , and r_{2b} .

given in Table I. Since the angle of refraction φ cannot be less than 0° a critical angle can be defined as

$$\varphi_{wc} = \arccos \frac{c_w}{c}, \quad (3)$$

and the sound is completely reflected for angles $\varphi_w \leq \varphi_{wc}$. For reflection angles φ_w less than the critical angle, the re-

TABLE I. Acoustic impedance and reflectivity of various substances in sea water. The reflectivity values are for normal incidence on an ideal plane interface between water and the material (Stephens, 1970).

Substance	Impedance in $(\text{N/m}^2)/(\text{m/s})$	Reflectivity in %
Air	$0.0004 \cdot 10^6$	-100
Sandstone	$7.6 \cdot 10^6$	66
Granite	$16 \cdot 10^6$	82
Sea water	$1.54 \cdot 10^6$...

flection coefficient R becomes complex. The phase of the bottom reflected signal relative to the source signal is dependent on the reflection angle and moves from 0° ($\varphi_w \gg \varphi_{wc}$) to 180° ($\varphi_w \leq \varphi_{wc}$). For sand (assuming a sound velocity of 1710 m/s) the critical angle is around 28° . No critical angular phenomena occur for the surface reflection since the sound velocity in air is lower than that in water.

In shallow water and for relatively short ranges (water depth $b < 25$ m, range $r < 100$ m) sound propagation is approximately linear, since ray deflection can be neglected because of the relatively homogenous speed of sound in the water column (Clay and Medwin, 1977). In addition, waters have to be calm, since the surface path of the sound can be changed drastically by waves, especially when the calling animal is close to the water surface.

Different orders of multipath propagation must be considered. The order number n indicates the total number of

signal reflections at the air–water and water–bottom interfaces before the signal arrives at the receiver. The propagation path with the initial reflection at the air–water interface is called the surface path, while the path with the first reflection at the water–bottom interface is called the bottom path. Cases with $n > 1$ indicate multiple reflections between the surface and bottom interfaces (see Fig. 1). Equation (4) describes the length r_{ns} of the surface paths while Eq. (5) describes the length r_{nb} of the bottom paths, for orders $n \geq 1$. The length of the direct sound path is r , while the depth of the receiving hydrophone, the water, and the dolphin are a , b , and d , respectively. The horizontal separation between the sound source and the receiver is r' .

$$r_{ns} = \sqrt{r'^2 + h_{ns}^2}, \quad (4)$$

$$r_{nb} = \sqrt{r'^2 + h_{nb}^2}, \quad (5)$$

where

$$r' = \sqrt{r^2 - (d - a)^2}, \quad (6)$$

$$h_{ns} = \begin{cases} b \cdot (n - 1) + a + d, & n \text{ odd} \\ b \cdot n - a + d, & n \text{ even} \end{cases}, \quad (7)$$

$$h_{nb} = \begin{cases} b \cdot (n + 1) - a - d, & n \text{ odd} \\ b \cdot n + a - d, & n \text{ even} \end{cases}. \quad (8)$$

The path lengths r_{ns} and r_{nb} become longer with increasing reflection order, so that signal components traveling on higher-order reflection paths always arrive later at the receiver than lower-order reflections. In addition, the reflection angles φ_{ns} and φ_{nb} increase with the reflection order n [Eqs. (9) and (10)]. Within one reflection path the reflection angles are constant because of Snell's law of reflection,

$$\varphi_{ns} = \arctan \frac{h_{ns}}{r'}, \quad (9)$$

$$\varphi_{nb} = \arctan \frac{h_{nb}}{r'}. \quad (10)$$

The increasing reflection angles with higher-order reflection paths result in large differences in signal amplitudes at the hydrophone because of the transmission beam pattern of dolphin clicks [the half-power angle is less than 10° for an echolocating *Tursiops truncatus* (Au, 1980)]. In addition, the signal is attenuated and spread with every reflection, especially when the reflective surface is rough and of lower reflectivity like the water–bottom interface. Higher-order reflections ($n \geq 2$) are therefore strongly attenuated and can be observed only for high source levels where an animal is directed at the hydrophone and for dolphin to hydrophone distances much larger than the water depth. For these large distances the reflection angle is relatively small, so that the amplitudes of the higher-order reflections are not attenuated much compared to the direct signal.

II. SOUND SOURCE LOCALIZATION

The measurement of arrival time differences between the direct and the indirect signals received with one hydrophone provides an easy way to determine the distance and

depth of calling animals, assuming the propagation conditions are known and constant. At the study site, approximately 200–1000 m offshore along the Waianae coast of Oahu, Hawaii, the sea floor in certain areas is almost evenly flat (less than a 1-m drop in depth per 100-m of distance) and consists mainly of basaltic rock with some sandy patches, almost without any corals or other outcrops. Spinner dolphins (*Stenella longirostris*) use this shallow coastal area (depths below approximately 30 m) mainly during the day for resting and socializing purposes, after returning from nighttime feeding in open waters. During socializing periods, these dolphins are acoustically very active and produce many echolocation click trains, burst pulse sounds, and whistles. Recordings used in this study were only made on calm days (Sea State one or less) from an anchored boat in waters with low current conditions. Depth at the actual study site was determined using a hand-held, personal dive sonar (Under Sea Industries).

Dolphin signals were recorded using a custom-made spherical PZT piezoelectric hydrophone with a flat (± 5 dB) frequency response from 10 to 160 kHz. The hydrophone was connected to a portable broadband data acquisition system, which digitized the incoming signal at a sampling rate of 260 kHz at 12-bit resolution. This recording system is composed of a preamplifier, analog filters, and a data-acquisition PCMCIA card plugged into a notebook computer that allows pretrigger data acquisition (Au *et al.*, 1999).

A burst pulse sequence from a Spinner dolphin is shown in Fig. 2 with the spectrogram representation in Fig. 2(a) and the time-domain representations in Fig. 2(b) and (c). The sequence is fairly typical of most burst pulse sequences recorded at the study site. It is relatively short, contains very broad band clicks with peak frequencies between about 15 kHz to more than 130 kHz and has interclick intervals of 6 to 9 ms. The magnification of click number five in Fig. 2(c) reveals the key property of this sequence and many other click trains recorded at this study site, namely, initial clicks are followed by up to four (three in Fig. 2) more clicks, lower in amplitude and equal or longer in duration. The time delays between the initial and subsequent clicks are consistent over the entire sequence, 404, 869, and 1808 μ s (± 4), respectively.

Assuming that click repetitions are caused by multipath propagation of the emitted dolphin sounds and not by the dolphin itself, the distance and depth of the dolphin can be determined by measuring the time delays, τ_{nb} and τ_{ns} , of the trailing clicks relative to the initial click and by using Eqs. (4) and (5). The distance r for signals with $n = 1$ path is given by

$$r = \frac{\left(\frac{c \cdot \tau_{1b}}{2}\right)^2 + b \cdot (a - b) - \left(\frac{c \cdot \tau_{1s}}{2}\right)^2 \cdot \left(1 - \frac{b}{a}\right)}{\frac{c \cdot \tau_{1s}}{2} \cdot \left(1 - \frac{b}{a}\right) - \frac{c \cdot \tau_{1b}}{2}}. \quad (11)$$

The corresponding depth d is given by

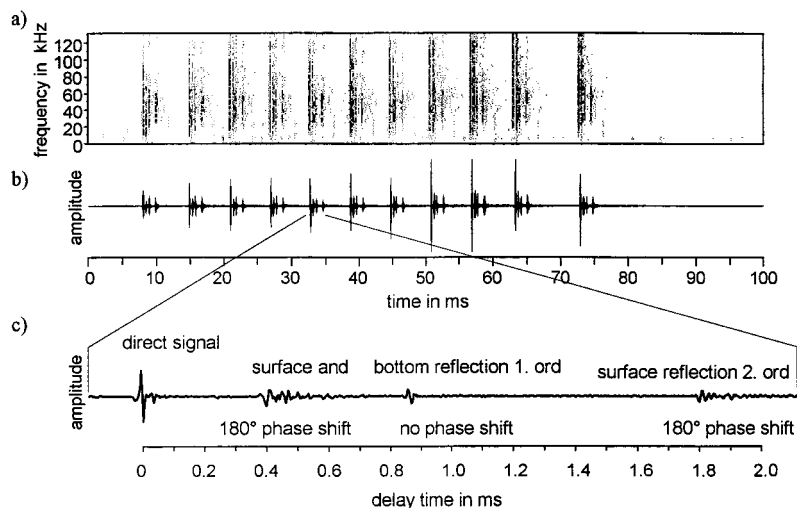


FIG. 2. (a) Spinner dolphin burst pulse signal with multipath propagation recorded in shallow waters along the Waianae Coast of Oahu. The spectrogram of the burst signal sequence shows the broadband nature of these sounds. (b) Complete burst pulse sequence in the time domain, interclick interval 5–10 ms. (c) Enlargement of a single burst pulse with first- and second-order multipath propagation.

$$d = \frac{\left(\frac{c \cdot \tau_{1b}}{2}\right)^2 + \left(\frac{c \cdot \tau_{1b}}{2}\right) \cdot r}{a - b} + b. \quad (12)$$

The uncertainty of assigning the measured time delay to the correct propagation paths can be resolved in many cases considering the phase relationship of the click repetitions. Sound traveling from a higher to a much lower density medium (the water–air interface) is reflected at the medium boundary with a 180° phase shift, whereas sound traveling from a lower to a higher density medium (the water–bottom interface) is reflected with a phase shift between 0 and 180° , depending on the angle of reflection [see Eq. (1) and Fig. 2(c)]. The 180° phase shift can be detected easily at the onsets of the signal repetitions because the first and second cycles are inverted compared to the direct signal (see Fig. 2). The order of the propagation path can be determined by the time of arrival, since higher-order paths are considerably longer so that they always arrive later than the lower ordered signals.

Two reflection signals are necessary for localization. Further reflections, like the third trailing click in the example shown [Fig. 2(c)], are redundant but can be used to verify the

method or to increase the accuracy of the localization through error correction.

III. ERROR ESTIMATION

The accuracy of the localization method can be estimated with the total error differential of the distance and depth functions. The total error differential is the sum of the partial derivatives of all variables multiplied by the error bounds of the individual measurements [see Eqs. (13) and (14)]. The method is used for obtaining an upper bound on the location error considering the measuring errors of the individual variables (Bronstein and Semendjajew, 1983).

In the following analysis, measurement errors of $|\Delta a| = 0.1$ m for the hydrophone depth a , $|\Delta b| = 1$ m for the water depth b , $|\Delta c_w| = 5$ m/s for the sound velocity c_w in water, and $|\Delta \tau| = 8 \mu\text{s}$ for the time delays τ_{1b} and τ_{1s} are assumed. Furthermore, a water depth of $b = 12$ m, typical at the study site, and a hydrophone depth of $a = 3$ m are used. The distance r and the depth d of the dolphin are determined with Eqs. (11) and (12). Only the first-order ($n = 1$) propagation paths are used. The upper bound of the relative distance error $\Delta r/r$ is established using the total error differential Δr ,

$$\frac{\Delta r}{r} = \frac{\left|\frac{\partial r}{\partial a}\right| \cdot |\Delta a| + \left|\frac{\partial r}{\partial b}\right| \cdot |\Delta b| + \left|\frac{\partial r}{\partial c_w}\right| \cdot |\Delta c_w| + \left|\frac{\partial r}{\partial \tau_{1b}}\right| \cdot |\Delta \tau| + \left|\frac{\partial r}{\partial \tau_{1s}}\right| \cdot |\Delta \tau|}{r}. \quad (13)$$

The upper bound of the absolute depth measurement corresponds with the total error differential Δd ,

$$\Delta d = \left|\frac{\partial d}{\partial a}\right| \cdot |\Delta a| + \left|\frac{\partial d}{\partial b}\right| \cdot |\Delta b| + \left|\frac{\partial d}{\partial c_w}\right| \cdot |\Delta c_w| + \left|\frac{\partial d}{\partial \tau_{1b}}\right| \cdot |\Delta \tau| + \left|\frac{\partial d}{\partial \tau_{1s}}\right| \cdot |\Delta \tau|. \quad (14)$$

The error estimation is shown in Fig. 3 for depths of 1, 4, and 8 m of the calling animal. Both upper bounds depend on the position on the dolphin relative to the receiving hydrophone. The upper bound of the relative distance measure-

ment error $\Delta r/r$ decreases with increasing distance and becomes almost constant with distances r larger than 25 m. The absolute depth error Δd has a minimum at distances between 5 and 20 m and then rises constantly with increasing dis-

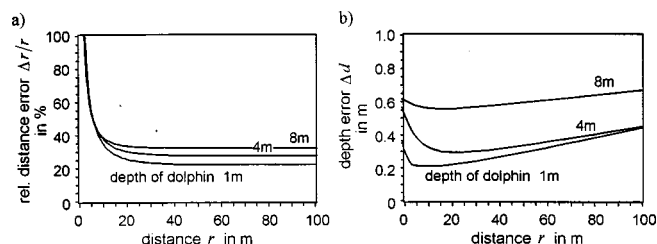


FIG. 3. (a) Relative error in distance measurement $\Delta r/r$ and (b) absolute depth error Δd of the localization.

tance. It is important to point out that these error estimates always give upper bounds. In practice, the actual localization errors are smaller since the measurement errors are typically normally distributed about the zero mean and tend to average out.

IV. RESULTS AND APPLICATIONS

The most interesting application of this localization method is the determination of source levels (SL) of click trains produced by cetaceans in the wild. This parameter is very important for studies of social behavior and foraging of these animals since it determines the range extent for communication and echolocation. Most source level measurements of dolphin echolocation clicks have been done with captive animals and may not necessarily reflect source levels produced by animals in open waters (Au, 1993). The broadband nature of social signals, particularly of the burst pulse signals of certain dolphin species, has been documented only recently, so that little or almost nothing is known about source levels of these signal types (Lammers *et al.*, 1998).

Since dolphins emit click sound directionally and the orientation of an animal cannot be determined with the method presented here, statements about the maximum source level of signals are only hypothetical. Source level is defined as the sound pressure level at 1 m from the source. It is often not measured at that distance so that the transmission loss due to spherical spreading and absorption must be taken into account with the equation

$$SL = SPL_R + 20 \cdot \log r + \alpha \cdot r, \quad (15)$$

where SPL_R is the sound pressure level *re* 1 μPa of the recorded signal, r is the distance between the dolphin and the hydrophone, and α is the absorption coefficient of the water measured in dB/m.

Maximum source levels of echolocation and burst pulse signals were determined for spinner dolphins at our study site. Only click trains of more than 10 clicks and at least 500-ms duration were examined. The maximum source levels were found for click trains where the position of the dolphin could be tracked over some distance and the animal's range continuously declined by 1 m/s or more. It was assumed that the dolphins orientated toward the hydrophone. A maximum peak-to-peak SL of 219 dB *re* 1 μPa was found for echolocation clicks and 214 dB *re* 1 μPa for burst pulse clicks of spinner dolphins, respectively. The upper bound of the distance measurement of $\Delta r/r < 0.35$ for distances $r > 20$ m translates into a source level measurement error ΔSL of

$$|\Delta SL| = 20 \cdot \log \left(1 \pm \frac{\Delta r}{r} \right) \leq 3.7 \text{ dB}. \quad (16)$$

We assumed uncertainty involved with calculating the absorption coefficient does not contribute significantly to the source level measurement [α changes less than 0.02 dB/m for a water temperature change of 10 °C in the frequency range up to 200 kHz (Au, 1993)].

A second application of the presented localization method is the acoustic tracking of dolphins. Figure 4 shows a track of a spinner dolphin recorded along the Waianae coast of Oahu, Hawaii (near Kahe point). Each cross represents a calculated position derived from an echolocation click. The track covers 3 sec in time and a path length of about 6 m. During this time span the dolphin continuously emitted echolocation clicks with durations less than 50 μs and bandwidths exceeding that of our recording system (130 kHz). The interclick interval in this click train varied from 70 to 115 ms. As predicted from the error estimation, the deviation in depth is significantly lower than the deviation in the estimated distance.

About 10% of all the recorded echolocation and burst pulse sounds had two or more distinct signal reflections caused from multipath propagation and high enough signal-to-noise ratios for acoustic localization. Another 25% of the signals had overlapping components caused by propagating along a similar path length. In these cases analysis of the time delay and phase of the signal is uncertain. The range of localized animals was between 20 and 110 m, with signals from remote animals frequently containing three or more signal reflections that allowed the verification of the localization and error correction.

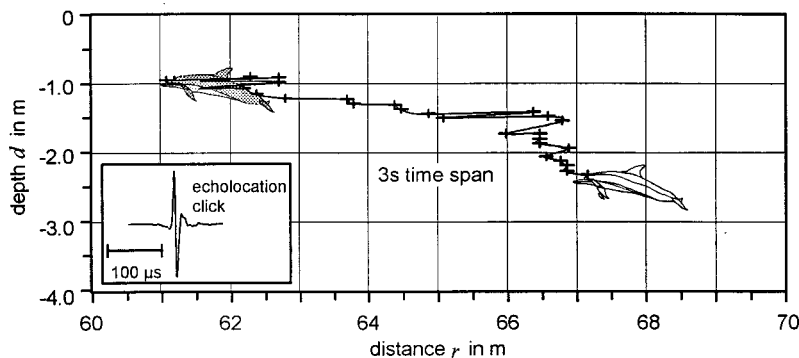


FIG. 4. Acoustic track of a Spinner dolphin at the Waianae coast of Oahu, Hawaii, and echolocation click of the dolphin. The fluctuations of the estimated positions (+) are due to the measurement errors.

V. CONCLUSIONS

The use of multipath propagation of cetacean click sound emissions for acoustic tracking with only one hydrophone provides a simple way to acquire valuable information about sound source levels and animal range and depth for wild dolphins. The technique requires almost no additional costs beyond a single channel recording system. Echolocation and burst pulse click trains allow precise measurements of time-of-arrival differences and the reliable identification of reflections paths by the evaluation of phase relations. The method provides higher two-dimensional (2D) localization (distance, depth) accuracy over large distances than visual range estimates and allows relatively precise estimates of sound source levels. Certain properties of the study site may limit the use of the localization method, since signal propagation over surface and bottom paths is absolutely required. In particular, the sea floor must be relatively even and of relatively high reflectivity (i.e., rock or sand). Signal overlap can complicate the arrival time measurements, so that localization is only possible for certain positions of the animal relative to the receiving hydrophone where surface and bottom reflection paths have distinctly different lengths. In addition, the animal must be oriented somewhat toward the hydrophone since the transmitting beam pattern of the animal is relatively narrow and signals at wide angles from the major axis of the beam can be attenuated significantly. The best results were achieved with animals close to the water surface and ranges considerably larger than the water depth.

In the past, repetitions of dolphin signals with time intervals of less than 1 ms been interpreted as double or multiple clicking by the animal (Norris, 1968). This paper suggests in some cases double clicks may be due to some

surface reflection, that signal repetitions of spinner dolphins are, in fact, multipath propagation of single clicks emitted by the dolphin.

ACKNOWLEDGMENTS

This work was generously supported by research grants provided by the German Research Foundation (DFG) and by the Office of Naval Research (ONR). At the same time, the authors would like to extend their appreciation to Dr. Paul Nachtigall, director of the Marine Mammal Research Program, Hawaii Institute of Marine Biology, for his abundant support and encouragement.

- Au, W. W. L. (1980). "Echolocation signals of the Atlantic bottlenose dolphin (*Tursiops truncatus*) in open water," in *Animal Sonar Systems*, edited by R. G. Busnel and J. F. Fish (Plenum, New York), pp. 251–282.
- Au, W. W. L. (1993). *The Sonar of Dolphins* (Springer, New York).
- Au, W. W. L., Lammers, M. O., and Aubauer, R. (1999). "A portable broadband data acquisition system for field studies in bioacoustics," *Mar. Mammal Sci.* **15**(2), 526–531.
- Bronstein, I. N., and Semendjajew, K. A. (1983). *Taschenbuch der Mathematik* (Nauka, Moskau & Teubner, Leipzig), pp. 100, 276–277.
- Cato, D. H. (1998). "Simple methods of estimating source levels and locations of marine animal sounds," *J. Acoust. Soc. Am.* **104**(3), 1667–1678.
- Clay, S. C., and Medwin, H. (1977). *Acoustical Oceanography: Principles and Applications* (Wiley, New York), pp. 3–11.
- Lammers, M. O., Au, W. W. L., and Aubauer, R. (1998). "The broadband nature of spinner dolphin (*Stenella longirostris*) social signals: Evidence from field studies," Abstracts of the 12th Biannual Conference on the Biology of Marine Mammals in Monaco, January (1998).
- Norris, K. S. (1968). "The echolocation of marine mammals," in *The Biology of Marine Mammals*, edited by T. H. Anderson (Academic, New York), pp. 391–423.
- Spiesberger, J. L., and Fristrup, K. M. (1990). "Passive localization of calling animals and sensing of their acoustic environment using acoustic tomography," *Am. Nat.* **135**, No. 1, 107–153.
- Stephens, R. W. B. (1970). *Underwater Acoustics* (Wiley-Interscience, London), pp. 140–141.

Classification of electronically generated phantom targets by an Atlantic bottlenose dolphin (*Tursiops truncatus*)

Roland Aubauer,^{a)} Whitlow W. L. Au, Paul E. Nachtigall, Deborah A. Pawloski,^{b)} and Caroline M. DeLong

Hawaii Institute of Marine Biology, University of Hawaii, P.O. Box 1106, Kailua, Hawaii 96734

(Received 7 September 1999; accepted for publication 30 January 2000)

Animal behavior experiments require not only stimulus control of the animal's behavior, but also precise control of the stimulus itself. In discrimination experiments with real target presentation, the complex interdependence between the physical dimensions and the backscattering process of an object make it difficult to extract and control relevant echo parameters separately. In other phantom-echo experiments, the echoes were relatively simple and could only simulate certain properties of targets. The echo-simulation method utilized in this paper can be used to transform any animal echolocation sound into phantom echoes of high fidelity and complexity. The developed phantom-echo system is implemented on a digital signal-processing board and gives an experimenter fully programmable control over the echo-generating process and the echo structure itself. In this experiment, the capability of a dolphin to discriminate between acoustically simulated phantom replicas of targets and their real equivalents was tested. Phantom replicas were presented in a probe technique during a materials discrimination experiment. The animal accepted the phantom echoes and classified them in the same manner as it classified real targets. © 2000 Acoustical Society of America. [S0001-4966(00)01205-4]

PACS numbers: 43.80.Ka, 43.80.Lb [FD]

INTRODUCTION

The biological sonar of bats and dolphins is characterized by outstanding discrimination and classification capabilities, even under difficult conditions such as high noise and cluttered environments (Nachtigall and Moore, 1988; Au, 1993). In order to investigate the echolocation system of these mammals, behavioral experiments are valuable to measure perceptual thresholds and to examine the classification of certain target features.

Echo stimuli used in behavioral experiments to study discrimination and classification capabilities of echolocating animals have been from either real targets (mainly in dolphin experiments) or simplified electronically generated echoes (mainly in bat experiments). Both types of stimuli present their own disadvantages. The complex interdependence between the physical dimension and the backscattering characteristics of a target make it difficult to independently control echo parameters such as amplitude, highlight number, duration, and phase relationships. The inability to precisely and independently control echo parameters can severely limit what can be learned from real-target experiments. Past experiments with electronically generated echoes often lacked proximity to reality. The echoes were often "canned" and could not change according to the animal's echolocation signal or consisted only of a few wavefronts and could not simulate reality (Roverud, 1989; Mogdans *et al.*, 1993). This project utilizes an echo-simulation method that combines the advantages of both approaches to echo stimuli presentation (this method was first described in Aubauer and Au, 1998).

I. ACOUSTIC SIMULATION OF UNDERWATER TARGETS

The target impulse response $h(t)$ can be used to describe the acoustic backscattering process of targets (Aubauer and Au, 1998). The backscattered target echo $e(t)$ is the convolution of the incident signal $s(t)$ with the target impulse response and can be expressed in both the time and the frequency domain

$$e(t) = h(t) * s(t) = \int_{-\infty}^{\infty} h(\tau) \cdot s(t - \tau) d\tau, \quad (1)$$

$$E(\omega) = H(\omega) \cdot S(\omega). \quad (2)$$

With these equations, it is possible to determine either the impulse response of a real target or, if the impulse response is known, to simulate an echo of a real target for any incident signal. For the echo simulation, the incident signal has to be measured with a hydrophone and transformed via convolution or Fourier and inverse Fourier transformation into the "phantom echo," which is played back to the echolocating animal with a projector hydrophone. Since a hydrophone placement at the same position where the phantom target appears would cause an unwanted echo overlap of the hydrophone and the phantom target, both receiver and projector hydrophones have to be positioned separate from the phantom target location. Consequently, the signal transmission path, which causes propagation time delay and attenuation, has to be considered in the target impulse response. Moreover, the inverse hydrophone transfer function can be included, to equalize the hydrophone characteristic and achieve a flat frequency response. The phantom-echo generation can be done either in the time or frequency domain and can be executed by digital signal processing. In

^{a)}Present address: Mussumer Kirchweg 174, 46395 Bocholt, Germany;
Electronic mail: aubauer@t-online.de

^{b)}Present address: P.O. Box 940, Waimanalo, HI 96795.

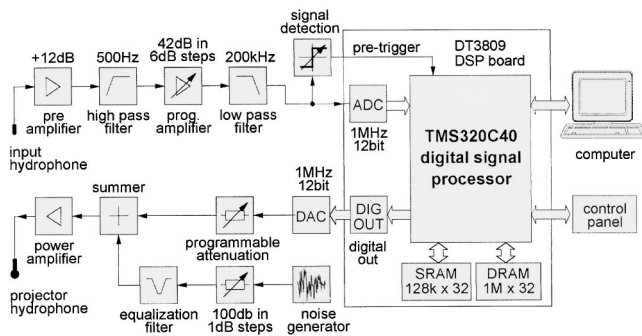


FIG. 1. Block diagram of the phantom-echo generator.

general, the introduced target impulse response represents a target only from one aspect and is dependent on the orientation of the target relative to the sound transmitting and receiving location.

The phantom-echo generator (PEG) was implemented on the DT3809 digital signal-processing (DSP) system from Data Translation plugged into a PC (Aubauer and Au, 1998). The signal input and output and a hand-held control panel was realized externally and connected through diverse analog and digital interfaces to the DSP system. A block diagram of the PEG is shown in Fig. 1.

The input signal received with the Brüel & Kjær 8103 hydrophone was amplified and filtered with a 500-Hz high-pass and a 200-kHz low-pass filter before going to the analog input of the DSP system. A pretrigger signal was generated when the analog input signal exceeded the user-set input threshold. The DSP system was able to collect pre- and post-triggered data before and after the trigger event, so even the beginning of a signal, before it exceeded the trigger level, was acquired. The input signal was digitized with a sampling rate of 1 MHz and 12-bit resolution. The signal transformation based on the fast convolution method was programed on the digital signal processor, which allowed the generation of a phantom echo of 512 digital samples in less than 3 ms. After a certain time delay, which determined the distance of the phantom target from the echolocating animal, the phantom echo was put out with the same sampling rate and resolution as the input signal was acquired. The system was calibrated prior to the experiment so that phantom echoes returning to the animal had the same amplitude as echoes from a real target (Aubauer and Au, 1998). The spherical ITO1042 hydrophone from International Transducer Corporation served as projector hydrophone; its transmit sensitivity curve is shown in Fig. 2. To compensate for the hydrophone frequency characteristics, the phantom echo was digitally pre-equalized on the DSP system for an overall frequency response of 20 to 250 kHz within ± 1 dB. The system was controlled from the PC and a hand-held control panel (Aubauer and Au, 1998).

II. TARGET IMPULSE RESPONSES

The theoretical evaluation of the acoustical impulse response of real targets is difficult, because of the propagation capability of both longitudinal and transversal waves in solids, and is only possible under simplifying assumptions

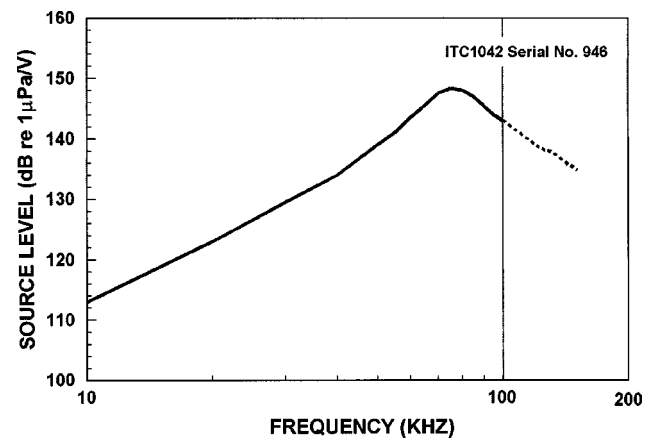


FIG. 2. Transmit calibration curve for the ITC1042. The instrument's curve was plotted out to 100 kHz (solid line). Values up to 150 kHz (dashed line) were obtained from manufacturer's standard curve.

(Neubauer, 1986; Shirley and Diercks, 1970). The impulse responses of certain, especially useful targets were therefore determined experimentally. Spheres of 7.62-cm diameter were selected for the experiment because of their aspect-independent but diverse impulse responses and echoes. The impulse responses were measured in a tank filled with seawater. Two different hydrophones were used for signal projection and echo measurement. The piezoceramic multielement transducer WAU1 (custom-made hydrophone) was used for projection of a delta pulse that was bandlimited between 40 and 260 kHz. The incident signal and the target echo were received with the Brüel & Kjær 8103 measuring hydrophone. The measured echo signals were averaged 2000 times and background clutter was subtracted, in order to improve the signal-to-noise ratio to about 52 dB.

The impulse responses of the hollow water-filled steel (HS), the solid steel (SS), and the solid brass sphere (SB) were determined by recasting Eq. (1) and are shown in Fig. 3 on the left side. The outside diameters of all spheres were 76.2 mm (3 in.). The hollow steel sphere had a wall thickness of 1.9 mm and was filled with seawater. The target

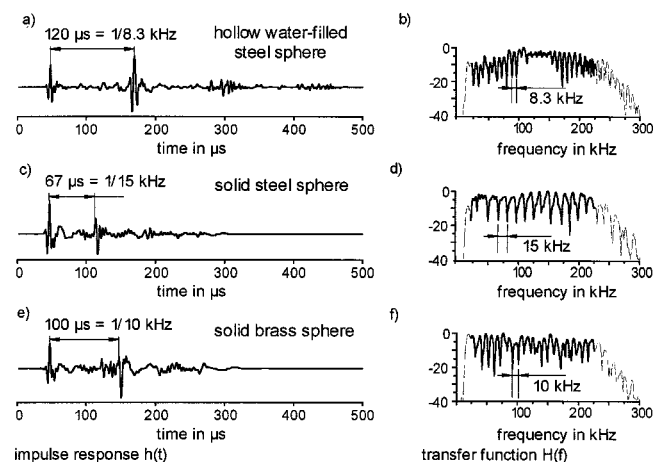


FIG. 3. Target impulse responses of the HS, SS, and SB sphere (a), (c), (e). Target transfer function of the HS, SS, and SB sphere (b), (d), (f). Dashed lines indicate the frequency band limitation of the transfer functions due to the insufficient transmitting characteristic of the measuring system and not of the sphere itself.

response was correct or incorrect. Nonreinforced trials were spaced irregularly within the trial sequence of the session (e.g., trials 4, 7, 17, 23, 32, 36, 40, and 47 out of a 50 trial session were nonreinforced). The SS sphere was the standard target and the SB sphere was the comparison target. Four standard and four comparison trials out of 50 trials were designated as nonreinforced trials prior to a session.

The nonreinforced trials served as a preparation to the introduction of nonreinforced probe trials. In the probe technique, which has been used in other target echo recognition tasks (Hammer and Au, 1980), a small number of probes are presented which share some characteristics with the baseline targets, but systematically differ. These probes can be reinforced or nonreinforced. In this experiment, phantom trials were presented as nonreinforced probe trials. Probes were not reinforced because we didn't want to give the animal any feedback, to avoid the possibility of training her to respond in a certain way to the phantoms.

As soon as the discrimination performance of the animal stabilized on a correct-response level higher than 90%, phantom replicas of the standard (PSS) and the comparison sphere (PSB) replaced two of the nonreinforced standard and two of the nonreinforced comparison trials. A session of the first phase of the phantom target classification experiment thus consisted of 21 reinforced standard (SS), 2 nonreinforced standard (SS), and 2 nonreinforced phantom standard presentations (PSS). The same applied to the comparison presentations (21 reinforced real comparison (SB), 2 nonreinforced real comparison (SB), and 2 nonreinforced phantom comparison (PSB) trials).

In the second phase of the phantom target classification experiment, further comparison targets were introduced and no phantom comparison stimuli were presented. This strategy was chosen in order to test the hypothesis that the dolphin was doing a standard versus nonstandard (*A* versus not-*A*) task, not a standard versus a comparison (*A* versus *B*) classification. A session in phase 2 of the experiment consisted of 21 reinforced standard (SS), two nonreinforced standard (SS), and two nonreinforced phantom standard presentations (PSS). The nonstandard trials consisted of 21 reinforced SB and four nonreinforced nonstandard presentations. These four trials consisted of two solid brass (SB), one solid aluminum (SA), and one solid nylon sphere (SN) presentations (all spheres were 76.2 mm in diameter). The nonreinforced trials were selected randomly prior to a session.

IV. RESULTS

The results of the first six sessions of the phantom classification experiment (phase 1) are shown in Fig. 5. The dolphin's performance at the standard trials (SS) is shown in the upper half and the comparison trials (SB) in the lower half of the diagram. The probe trial results of the phantom standard (PSS) and comparison (PSB) stimuli are shown next to the respective real-target performance. On average, the real standard target was identified 92% correct (8% misses) and the real comparison target at 84% (16% false alarms). In comparison, all phantom replicas of the standard sphere (PSS) were classified as standards, where 11 phan-

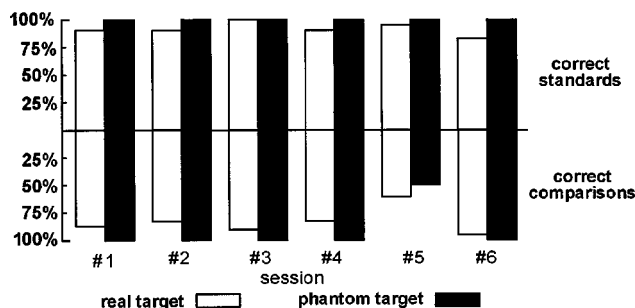


FIG. 5. Dolphin's performance in the first phase of the classification experiment. Standard presentations are shown in the upper half, comparison in the lower half of the diagram. The number *n* denotes the number of trials for each target class per session.

toms of the comparison sphere (PSB) were assigned to the comparison class. Only one phantom comparison target was misclassified as the standard.

The misclassified phantom trial in session #5 was consistent with a relatively poor performance on real comparison stimuli. After the first half of the session we noticed the camera cable hanging in front of the hoop station, which probably served as a distraction to the animal. The cable was removed and the performance of the animal immediately returned to normal in the second half of session #5.

In addition to the response, the latency time for a go response on a real (SS) and phantom standard trial (PSS) was measured (Fig. 6). The latency time is the time interval between lowering the acoustic screen and starting the trial, and the time when the dolphin leaves the hoop for touching the response paddle (rostrum completely out of the hoop). The mean of the go latency in the first phase of the experiment is clearly below the no-go threshold of 6 s and similar for both real (3.0 s) and phantom standard trials (2.8 s). The standard deviation of the SS trials is under 0.6 s.

The results of the second phase of the classification experiment are shown in Fig. 7. Percent-correct responses are drawn for each target class separately. The dolphin got 93% of the standard SS trials (*n* = 299) correct and reported 100% of the phantom PSS probes (*n* = 26) as standards. The brass comparison sphere SB was reported 89% correct, whereas 77% of the aluminum sphere probes and 100% of the nylon sphere probes (both *n* = 13) were classified as comparison stimuli. The mean latency of correct reported SS trials was 3.0 ± 0.7 s and on the PSS trials 2.3 ± 0.5 s.

V. DISCUSSION

Both phases of the experiment clearly show that the dolphin accepted the artificially generated phantom echoes in an

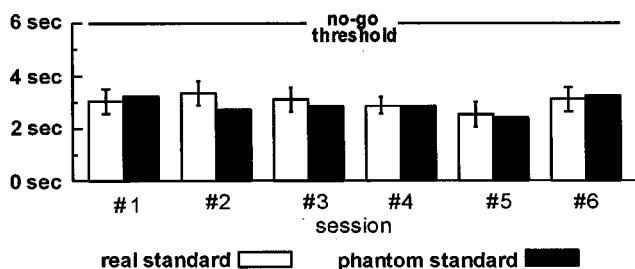


FIG. 6. Go-latency time in the first phase of the classification experiment.

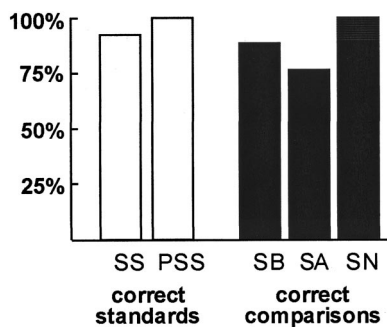


FIG. 7. Correct responses in the second phase of the classification experiment. SS denotes the solid stainless-steel standard sphere, PSS the phantom replica of SS, SB the solid brass sphere, SA the solid aluminum sphere, and SN the solid nylon sphere. The number n denotes the number of trials for each target class.

echolocation task. Furthermore, the phantom replicas of the standard and comparison targets were reported in the corresponding real-target class except for one trial, where the animal may have been disturbed. It can be assumed that the dolphin classifies the phantom echoes in the same manner as the echoes of the real-target models. Since further real-target comparison probes (aluminum and nylon) in the second phase of the experiment are reported with high confidence in the comparison class, it can be assumed the dolphin responded in an A /not- A fashion.

It is not clear why the dolphin reported the phantom standard probes PSS with a slightly higher confidence in the real standard class than the real standard targets SS itself. The dynamics of the backscattering process may give a better understanding of these experimental results. Echolocation measurements with artificial dolphin sonar showed that the phantom and the real-target echo behave slightly differently when the target or the projection hydrophone are moving. Because of the mass of the metal spheres, the targets are able to move independently from the surrounding water, whereas the relatively light and small hydrophones move almost perfectly with the water current. Therefore, waves and currents caused by wind and boat traffic caused the real-target echo to fluctuate in amplitude more than the phantom echo, i.e., the phantom echoes were more stable than the real echoes. However, this does not imply that the phantom echo was not perceived as coming from a real target.

There was a slight difference in procedure between phantom echo and real-target trials. On a phantom-echo trial, a real target was lowered into the water and then raised before the target screen was lowered and the dolphin began echolocating the phantom target. On a real-target trial, the

real target was lowered into the water and remained there for the dolphin to inspect. Therefore, the only difference was the subsequent removal of the real target just prior to a phantom presentation. However, all targets were introduced and extracted from the water very carefully to minimize any extraneous noise.

From our observations of the animal's performance, it seemed as if the dolphin accepted the phantom echo as coming from a real target. Our experiment was not set up to determine whether the dolphin could distinguish between phantom echoes and real echoes; it was designed to determine how the dolphin would classify phantom stimuli. The results of this experiment indicate that the dolphin classified the echoes from the phantom targets in the same manner as the echoes from real targets.

ACKNOWLEDGMENTS

This work was generously supported by several research grants provided by the German Academic Exchange Service (DAAD), by the German Research Foundation (DFG), and by the U.S. Office of Naval Research (ONR) Grant Number N000149510462 from Harold Hawkins. We thank Ronald Schusterman of Long Marine Laboratory, University of California, Santa Cruz, for suggesting the use of partially reinforced trials. This is HIMB contribution number 1101.

- Au, W. L. (1993). *The Sonar of Dolphins* (Springer, New York).
- Aubauer, R., and Au, W. W. L. (1998). "Phantom echo generation: A new technique for investigating dolphin echolocation," *J. Acoust. Soc. Am.* **104**, 1165–1170.
- Gellerman, L. (1933). "Chance orders of alternating stimuli in visual discrimination experiments," *J. Gen. Psychol.* **42**, 206–208.
- Hammer, C. E., and Au, W. W. L. (1980). "Porpoise echo recognition: An analysis of controlling target characteristics," *J. Acoust. Soc. Am.* **68**, 1285–1293.
- Mason, W. P. (1958). *Physical Acoustics and the Properties of Solids* (Van Nostrand, Princeton).
- Mogdans, J., Schnitzler, H. U., and Ostwald, J. (1993). "Discrimination of two-wavefront echoes by the big brown bat, *Eptesicus fuscus*: behavioral experiments and receiver simulations," *J. Comp. Physiol. A* **172**, 309–323.
- Nachtigall, P. E., and Moore, P. W. B. (1988). *Animal Sonar: Processes and Performance* (Plenum, New York).
- Neubauer, W. G. (1986). "Acoustic reflection from surfaces and shapes" (Naval Research Lab, Washington, DC).
- Roverud, R. C. (1989). "Harmonic and frequency structure used for echolocation sound pattern recognition and distance information processing in the rufous horseshoe bat," *J. Comp. Physiol. A* **166**, 251–255.
- Shirley, D. J., and Diercks, K. J. (1970). "Analysis of the frequency response of simple geometric targets," *J. Acoust. Soc. Am.* **48**, 1275–1282.
- Schusterman, R. J. (1980). "Behavioral methodology in echolocation by marine mammals," in *Animal Sonar Systems*, edited by R. G. Busnel and J. F. Fish (Plenum, New York), pp. 11–41.

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.

Kirchhoff evaluation of scattered elastic wavefields in anisotropic media

Martin Spies^{a)}

Fraunhofer-Institut Zerstörungsfreie Prüfverfahren (IZFP), Universität, Geb. 37, 66123 Saarbrücken, Germany

(Received 3 August 1999; accepted for publication 24 January 2000)

Three-dimensional elastic wave scattering is addressed for the case of general anisotropic materials. For traction-free scatterers the elastodynamic Kirchhoff approximation is applied, which is valid at high frequencies and near specular angles. Based on a mathematical formulation including Green's triadic stress tensor function, asymptotic evaluation yields the displacement vector of the scattered wavefield in a form which is convenient for effective numerical computation. The (point source) far-field approximation applied to Green's triadic function is valid in noncaustic regions. © 2000 Acoustical Society of America. [S0001-4966(00)00805-5]

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

INTRODUCTION

Scattering of elastic waves plays a fundamental role in the field of ultrasonic nondestructive testing as well as in the interpretation of seismic data. Numerous investigations have been performed for isotropic media since the 1950's, corresponding work on anisotropic materials started about two decades later. A variety of studies has dealt with transversely isotropic material symmetry for both the specular and the nonspecular scattering regimes.^{1–4} Recently, a number of articles has been published for general anisotropic media, where—dependent on the respective regime—different approximations have been applied in the theoretical treatment.^{5–10}

In this Letter, the three-dimensional scattering problem for anisotropic materials is considered by applying the elastodynamic Kirchhoff approximation to the case of traction-free scatterers. Based on a suitable mathematical formulation, relationships are obtained by asymptotic evaluation which are convenient for numerical computation. The elaborated representation for the displacement vector of the scattered wavefield appears as a point source superposition formulation including the directivities of the point sources located on the scatterer's traction-free surface. Thus, it is in the same form as corresponding representations for transducer-generated wavefields.¹¹

I. DISPLACEMENT VECTOR OF SCATTERED WAVEFIELDS

For volumetric scatterers, a volume integral equation can be used to determine the scattered displacement field.¹² However, since also planar defects are of interest here, the use of a surface integral formulation is more natural. For an elastic wave hitting a traction-free scatterer, the displacement vector of the resulting scattered elastic field can be formulated as¹³

$$\underline{u}_{sc}(\underline{R}, \omega) = \int \int_S dS' \underline{u}(\underline{R}', \omega) \cdot [\hat{\underline{n}} \cdot \underline{\underline{\Sigma}}(\underline{R} - \underline{R}', \omega)]. \quad (1)$$

Here, $\hat{\underline{n}}$ is the unit vector perpendicular to the scatterer's surface S , $\underline{\underline{\Sigma}}$ is Green's triadic function and $\underline{u}(\underline{R}', \omega)$ is the total displacement field on S . In the following, the latter quantities are evaluated by applying a far-field approximation and Kirchhoff-theory, respectively.

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

For a general anisotropic medium, characterized by the elastic stiffness tensor $\underline{\underline{C}}$ and density ϱ , Green's dyadic function $\underline{\underline{G}}$ is defined by the differential equation¹³

$$(\underline{\nabla} \cdot \underline{\underline{C}} \cdot \underline{\nabla}) \cdot \underline{\underline{G}}(\underline{R}, \omega) + \varrho \omega^2 \underline{\underline{G}}(\underline{R}, \omega) = -\underline{\underline{I}} \delta(\underline{R}), \quad (2)$$

where δ is Dirac's delta function, $\underline{\nabla}$ is the gradient vector, and ω denotes the circular frequency, if a time dependence $\sim e^{-j\omega t}$ is assumed. Green's triadic function $\underline{\underline{\Sigma}}$ is defined as

^{a)}Electronic mail: spies@izfp.fhg.de

$$\underline{\underline{\Sigma}}(\underline{\mathbf{R}}, \omega) = \underline{\underline{\mathbf{C}}} : \nabla \underline{\underline{\mathbf{G}}}(\underline{\mathbf{R}}, \omega). \quad (3)$$

The components of these functions represent the displacement and stress field, respectively, at position $\underline{\mathbf{R}}$ generated by three mutually perpendicular (point) forces acting at $\underline{\mathbf{R}}' = \underline{\mathbf{0}}$. The functions are made up by three terms, each one generating the different patterns of wave propagation related to the quasipressure (qP) and quasishear waves ($qSV1,2$) inherent to anisotropic materials.

Using a far-field approximation, each part of Green's dyad can be explicitly expressed according to¹⁴

$$\underline{\underline{\mathbf{G}}}_\alpha^{\text{far}}(\underline{\mathbf{R}}, \omega) = \frac{1}{\varrho \mathcal{K}_\alpha^{1/2}(\hat{\underline{\mathbf{K}}}(\hat{\underline{\mathbf{R}}})) c_\alpha(\hat{\underline{\mathbf{K}}}(\hat{\underline{\mathbf{R}}}))} \times \hat{\underline{\mathbf{u}}}_\alpha(\hat{\underline{\mathbf{K}}}(\hat{\underline{\mathbf{R}}})) \hat{\underline{\mathbf{u}}}_\alpha(\hat{\underline{\mathbf{K}}}(\hat{\underline{\mathbf{R}}})) \frac{e^{j\mathbf{K}_\alpha(\hat{\underline{\mathbf{R}}}) \cdot \underline{\mathbf{R}}}}{4\pi R}, \quad (4)$$

where α designates the wave type. In this equation, \mathcal{K}_α is the Gaussian curvature of the slowness surface, and c_α is the modulus of the group velocity vector. Both quantities—as well as the polarizations $\hat{\underline{\mathbf{u}}}_\alpha$ and the wave vectors $\underline{\mathbf{K}}_\alpha$ —are functions of the wave propagation direction $\hat{\underline{\mathbf{K}}}$ that produces an energy contribution along the spatial direction $\hat{\underline{\mathbf{R}}}$. The $\hat{\underline{\mathbf{K}}}(\hat{\underline{\mathbf{R}}})$ -relationship has in general to be evaluated numerically, only in special cases exact¹⁵ or approximate¹⁶ analytical representations are available. An efficient numerical evaluation scheme is described in Sec. II.

Using the definition of the stress tensor

$$\underline{\underline{\mathbf{T}}} = (\underline{\underline{\mathbf{C}}} : \nabla \underline{\mathbf{u}}), \quad (5)$$

and retaining only terms of order $\mathcal{O}(R^{-1})$, Eqs. (3) and (4) provide Green's triadic function for the far-field according to

$$\underline{\underline{\Sigma}}_\alpha^{\text{far}}(\underline{\mathbf{R}}, \omega) = \mathcal{C}(\hat{\underline{\mathbf{K}}}(\hat{\underline{\mathbf{R}}})) \underline{\underline{\mathbf{T}}}_\alpha(\hat{\underline{\mathbf{K}}}(\hat{\underline{\mathbf{R}}})) \hat{\underline{\mathbf{u}}}_\alpha(\hat{\underline{\mathbf{K}}}(\hat{\underline{\mathbf{R}}})) \frac{e^{j\mathbf{K}_\alpha(\hat{\underline{\mathbf{R}}}) \cdot \underline{\mathbf{R}}}}{4\pi R}, \quad (6)$$

where \mathcal{C} comprises the constant factors and $\underline{\underline{\mathbf{T}}}_\alpha$ designates the part of the stress tensor due to wave mode α . Finally, by applying the simple relationship $\underline{\mathbf{K}}_\alpha(\hat{\underline{\mathbf{R}}}) \cdot \underline{\mathbf{R}} = \omega R / c_\alpha$,¹⁷ it follows that

$$\hat{\underline{\mathbf{n}}} \cdot \underline{\underline{\Sigma}}_\alpha^{\text{far}}(\underline{\mathbf{R}}, \omega) = \underline{\mathbf{t}}_\alpha(\hat{\underline{\mathbf{K}}}(\hat{\underline{\mathbf{R}}})) \hat{\underline{\mathbf{u}}}_\alpha(\hat{\underline{\mathbf{K}}}(\hat{\underline{\mathbf{R}}})) \frac{e^{j\omega R / c_\alpha(\hat{\underline{\mathbf{K}}}(\hat{\underline{\mathbf{R}}}))}}{4\pi R}, \quad (7)$$

where $\underline{\mathbf{t}}_\alpha = \mathcal{C} \hat{\underline{\mathbf{n}}} \cdot \underline{\underline{\mathbf{T}}}_\alpha$ has been introduced.

B. Kirchhoff's approximation

With this result, $\underline{\mathbf{u}}_{sc}$ can be determined from Eq. (1) using Kirchhoff's approximation, which is applicable at high frequencies and near specular angles. It assumes the shadow side of the scatterer to be motionless, while the illuminated side is assumed to move locally as if it were a planar surface illuminated by a plane wave of amplitude equal to that of the local illuminating radiation. Accordingly, the displacement on S is given by

$$\underline{\mathbf{u}}_S(\underline{\mathbf{R}}', \omega) = U^I \frac{\exp[-j\omega|\underline{\mathbf{R}}' - \underline{\mathbf{R}}_0|/c_I(\hat{\underline{\mathbf{K}}}(\underline{\mathbf{R}}' - \underline{\mathbf{R}}_0))]}{4\pi|\underline{\mathbf{R}}' - \underline{\mathbf{R}}_0|} \cdot \left(\hat{\underline{\mathbf{u}}}^I + \sum_\alpha \mathcal{R}_\alpha \hat{\underline{\mathbf{u}}}^{\alpha R} \right), \quad (8)$$

where the wave mode incident onto the scatterer is designated by I , and $\underline{\mathbf{R}}_0$ is the source location of the incident wave. The respective (plane wave) reflection coefficients \mathcal{R}_α can be efficiently evaluated using a matrix algorithm as described in the next section. Thus, from Eq. (1)—using Eq. (7)—the scattered α -wavefield results as

$$\underline{\mathbf{u}}_{sc}^\alpha(\underline{\mathbf{R}}, \omega) \cong \int_S \int_S dS' [\underline{\mathbf{u}}_S(\underline{\mathbf{R}}', \omega) \cdot \underline{\mathbf{t}}_\alpha(\hat{\underline{\mathbf{K}}}(\underline{\mathbf{R}} - \underline{\mathbf{R}}'))] \cdot \hat{\underline{\mathbf{u}}}_\alpha \frac{\exp[-j\omega|\underline{\mathbf{R}} - \underline{\mathbf{R}}'|/c_\alpha(\hat{\underline{\mathbf{K}}}(\underline{\mathbf{R}} - \underline{\mathbf{R}}'))]}{4\pi|\underline{\mathbf{R}} - \underline{\mathbf{R}}'|}. \quad (9)$$

In this representation, the elementary wave term is supplied with the corresponding polarization vector and additionally multiplied with the scalar product of the displacement vector $\underline{\mathbf{u}}_S(\underline{\mathbf{R}}', \omega)$ on the scatterer's surface and the vector $\underline{\mathbf{t}}_\alpha$, which characterizes the traction generated on S by the incident wavefield.

II. COMPUTATIONAL SCHEME

The integration in Eq. (9) has to be performed over the scatterer's illuminated surface. Numerical integration on the basis of an equally spaced rectangular grid can be applied for planar surfaces. For cylindrical or spherical scatterers, a respective projection of a planar grid onto the curved surfaces has to be considered. With $\underline{\mathbf{R}}_m$ designating the position of the m th grid point, the displacement vector describing the scattered α -wavefield under concern follows accordingly as

$$\underline{\mathbf{u}}_{sc}^\alpha(\underline{\mathbf{R}}, \omega) = \sum_m [\underline{\mathbf{u}}_S(\underline{\mathbf{R}}_m, \omega) \cdot \underline{\mathbf{t}}_\alpha(\hat{\underline{\mathbf{K}}}(\underline{\mathbf{R}} - \underline{\mathbf{R}}_m^\Delta))] \hat{\underline{\mathbf{u}}}_\alpha(\hat{\underline{\mathbf{K}}}(\underline{\mathbf{R}} - \underline{\mathbf{R}}_m^\Delta)) \cdot \exp[-j\omega|\underline{\mathbf{R}} - \underline{\mathbf{R}}_m^\Delta|/c_\alpha(\hat{\underline{\mathbf{K}}}(\underline{\mathbf{R}} - \underline{\mathbf{R}}_m^\Delta))] \cdot |\underline{\mathbf{R}} - \underline{\mathbf{R}}_m^\Delta|^{-1}, \quad (10)$$

where the constant integration step size has been omitted. It is $\underline{\mathbf{R}}_m^\Delta = (\underline{\mathbf{R}} - \underline{\mathbf{R}}_m)$ and $\hat{\underline{\mathbf{R}}}_m^\Delta$ is the corresponding unit vector. According to Eq. (10), any scattered (continuous wave) displacement field can be determined using an equidistant distribution of grid points within the scattering surface in accordance with the sampling theorem.

The incident beam field entering into Eq. (8) can be obtained by applying the same computational scheme to the integration over the transducer surface.^{11,18} The only difference exists in the scalar product appearing in Eq. (10): here it can be interpreted as the directivity of a point source located on the scatterer's traction-free surface,¹⁹ while it characterizes the respective pattern for the stress-free material surface in the case of the incident transducer beam field. The results obtained in this way are valid even in the case where the scatterer is within the near-field of the beam.

A. Numerical evaluation algorithms

1. **$\hat{\mathbf{K}}(\hat{\mathbf{R}})$ -relationship.** The evaluation of Eq. (10) requires the determination of the wave vector direction $\hat{\mathbf{K}}$ that produces an energy contribution, i.e., a group velocity direction along the spatial direction $\hat{\mathbf{R}}$. The group velocity vector \mathbf{c} is obtained as a function of the $\hat{\mathbf{K}}$ -direction, so that the inversion of this relationship is required. However, due to the complex mathematical structure of these relationships the analytical inversion succeeds only in special cases. This has been performed for transversely isotropic material symmetry for SH -waves yielding an exact analytical expression,¹⁵ while for qP -waves an approximative formulation has been obtained for a certain class of TI materials like, e.g., unidirectionally grain-structured austenitic steels.¹⁶ In arbitrary anisotropic material the inversion has to be performed numerically, where the shape and possible degeneracy of the slowness surfaces lead to additional difficulties. An efficient $\hat{\mathbf{K}}(\hat{\mathbf{R}})$ -evaluation scheme has been implemented in the following way. The numerical program steps through the wave vector positions, parameterized by an azimuthal angle α and a polar angle κ in steps of, e.g., 0.5° , and determines the corresponding group velocity vectors, which are in turn parameterized accordingly. The angles α and κ characterizing $\hat{\mathbf{K}}$ are stored as an array, where the azimuthal and polar angles determined for the \mathbf{c} -vectors are used as array indices. In evaluating Eq. (10), the polar and azimuthal angles for a given direction $\hat{\mathbf{R}} (= \hat{\mathbf{c}})$ are then used to look up the corresponding $\hat{\mathbf{K}}$ -direction from the stored array. Additional care has to be taken in directions where group velocity is multivalued—here multiple look-up arrays are generated and applied, which also allows us to take care of the different wave decay properties in these directions ($\mathcal{O}(R^{-1/2})$ and $\mathcal{O}(R^{-5/6})$ instead of $\mathcal{O}(R^{-1})$ ⁽¹⁴⁾).

2. **Reflection coefficients.** The same program is used to determine the reflection coefficients for plane waves traveling in arbitrary directions relative to a traction-free scattering surface, which is assumed to lie within the x - y -plane. With the incident and the reflected waves given by

$$\underline{\mathbf{u}}^{R\alpha}(\underline{\mathbf{R}}, \omega) = U^{I,R\alpha} \underline{\mathbf{u}}^{I,R\alpha} e^{j\omega \underline{\mathbf{s}}^{I,R\alpha} \cdot \underline{\mathbf{R}}}, \quad (11)$$

the amplitudes $U^{R\alpha}$ and the slownesses $\underline{\mathbf{s}}^{R\alpha}$ are to be determined. These quantities are obtained by considering the continuity of the slownesses ($\underline{\mathbf{s}} \cdot \underline{\mathbf{e}}_i$ continuous, $i=x,y$) and the condition of a stress-free boundary according to

$$\underline{\mathbf{e}}_z \cdot \underline{\mathbf{T}}^I + \sum_{\alpha} \underline{\mathbf{e}}_z \cdot \underline{\mathbf{T}}^{R\alpha} = \mathbf{0}. \quad (12)$$

Considering wave propagation in the x - z -plane, Snell's Law requires all projections of the $\underline{\mathbf{s}}$ -vectors onto the interface to be identical ($s_x^I = s_x^{R\alpha} \equiv s_x$) and their y -components to vanish ($s_y^I = s_y^{R\alpha} = 0$). With these conditions, the $s_z^{R\alpha}$ -components are left to be determined. Solving the dispersion relations for the slowness- z -components yields a 6th order polynomial according to

$$\mathcal{P}(s_z^{R\alpha}) = \sum_{i=0}^6 C_i (s_z^{R\alpha})^{6-i} = 0. \quad (13)$$

Evaluation of the $s_z^{R\alpha}(s_x)$, which in some cases can have a nonvanishing imaginary part, is performed numerically using Bairstow's method²⁰ in a straightforward manner. From the six solutions (branches) of Eq. (13), characterizing three upgoing and three downgoing waves, the three reflected ones are selected by the condition that the energy flow is directed away from the traction-free (scattering) surface, i.e., back into the medium. Additionally, the $s_z^{RqS1,2}$ - and s_z^{RqP} -values have to be correctly identified, since their determination yields six unordered solutions. Their proper assignment to upgoing and downgoing waves is achieved in the following way: At the origin, the six q -branches are identified by directly comparing modulus and sign of the $s_z^{R\alpha}(s_x=0)$ -values. The numerical program then steps through the values of s_x and uses the results from the previous step to identify the current $s_z^{R\alpha}$ -values. Special care in this procedure, which has originally been proposed by van der Hijden,²¹ has to be taken at the branch points. The imaginary parts are selected from the condition, that the evanescent wave components have to decay exponentially with increasing distance from the traction-free surface. Examples illustrating the complex s_z -behavior as a function of s_x can be found in Ref. 18 for the transversely isotropic case.

With the slowness vectors thus obtained, the reflection coefficients are determined from Eq. (12). With $\underline{\mathbf{V}} \rightarrow j\omega \underline{\mathbf{s}}$, it follows from Eq. (5) that

$$\underline{\mathbf{e}}_z \cdot \underline{\mathbf{T}} = \underline{\mathbf{e}}_z \cdot (\underline{\mathbf{C}} \cdot \underline{\mathbf{V}} \underline{\mathbf{u}}) = j\omega U \underline{\mathbf{e}}_z \cdot (\underline{\mathbf{C}} \cdot \underline{\mathbf{s}} \underline{\mathbf{u}}) \equiv j\omega U \underline{\mathbf{t}}_z. \quad (14)$$

Using this relationship and defining the matrix

$$\underline{\mathbf{T}}_z^R \equiv (\underline{\mathbf{t}}_z^{RqS1} \quad \underline{\mathbf{t}}_z^{RqS2} \quad \underline{\mathbf{t}}_z^{RqP}), \quad (15)$$

as well as the vector

$$\underline{\mathbf{U}}^R \equiv (U^{RqS1} \quad U^{RqS2} \quad U^{RqP}), \quad (16)$$

allows to write Eq. (12) as

$$U^I \underline{\mathbf{t}}_z^I + \underline{\mathbf{U}}^R \cdot \underline{\mathbf{T}}_z^R = \mathbf{0}. \quad (17)$$

Solving this equation for the amplitudes of the reflected plane waves finally yields

$$\underline{\mathbf{U}}^R = -U^I \underline{\mathbf{t}}_z^I \cdot (\underline{\mathbf{T}}_z^R)^{-1}. \quad (18)$$

Equation (18) allows the numerical evaluation of the (amplitude) reflection coefficients $\mathcal{R}^\alpha = U^{R\alpha}/U^I$ for any direction of the incident wave, properly accounting for the effect of mode conversion inherent to the reflection process.

III. NUMERICAL EXAMPLE

The computational scheme has been applied to ultrasonic inspection simulation for transversely isotropic media in Ref. 19: results obtained for circular and rectangular scatterers correspond well with respective experimental data. The principal procedure applied in such a pulse-echo C-scan experiment is illustrated in Fig. 1. An ultrasonic transducer is moved across the material's surface, transmitting ultrasonic signals and receiving those reflected either at the specimen's backwall or at a defect. The C-scanning results are usually documented in form of echo dynamic curves. For a piezo-

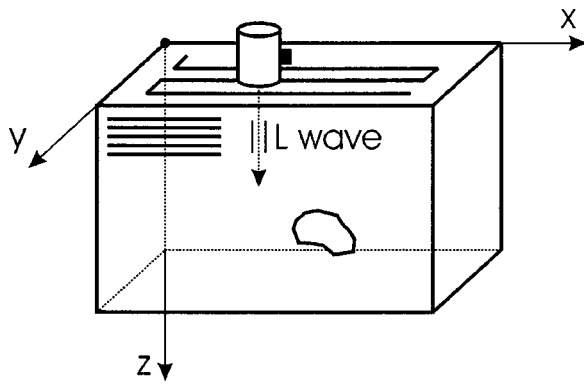


FIG. 1. Principal setup of a pulse-echo C-scan experiment in ultrasonic nondestructive testing.

electric normal-transducer (frequency 2.25 MHz, diameter 6.3 mm), generating (quasi-) longitudinal waves, simulated echo dynamic curves are shown in Figs. 2 and 3 for an orthotropic, layered $[0_3/90]$ -composite material. The defect is assumed to be of circular shape with a diameter of 5 and 10 mm, respectively, located at a depth of 25 mm. The curves are obtained for different orientations of the material, where the layering is oriented at 0, 10, and 20° to the specimen surface. The figures show the received amplitudes versus probe position relative to the defect (zero position), where only the quasilongitudinal (qP) wave contributions are taken into account. It can be seen that both the maximum amplitude and its position change with the material's orientation, the largest amplitude occurring in the case of 20°-orientation. For the 10 mm defect, the transducer is in the near-field of the scatterer, leading to correspondingly distorted signals (Fig. 3). The continuous-wave simulation of the detected signals has been performed on a standard PC (Pentium II, 300 MHz) within a calculation time of about 1 min for each scan path.

IV. SUMMARY

The theoretical evaluation of elastic wave scattering in general anisotropic media has been addressed. Using Kirchhoff's approximation and a far-field approximation for Green's tensor functions, an analytical representation for the scattered displacement field has been obtained which is well-suited for numerical evaluation. The formulation derived for

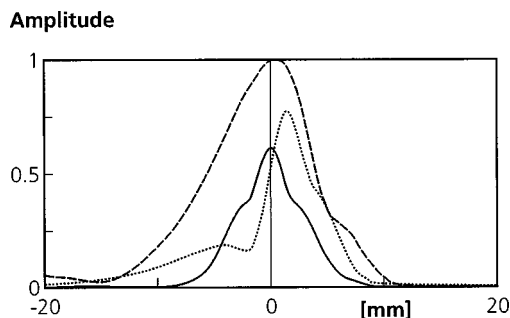


FIG. 2. Simulated echo dynamic curves: received amplitude versus probe position for a circular defect (5 diameter, 25 mm depth) in an orthotropic composite material with its layering aligned at 0 (solid), 10 (dotted), and 20° (dashed) to the surface.

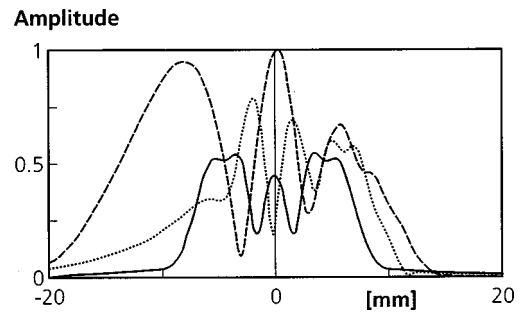


FIG. 3. Received amplitude versus probe position for a circular defect (10 diameter, 25 mm depth) in an orthotropic composite material with its layering aligned at 0 (solid), 10 (dotted), and 20° (dashed) to the surface.

the scattered displacement field is also valid in the case of isotropy. By explicit consideration of the corresponding isotropic relationships, Eq. (9) can be brought into the same form as the one given, e.g., in Ref. 22. Since the application regime of Kirchhoff's approximation is confined to the region of specular reflection, other methods such as Geometrical Theory of Diffraction⁷ have to be applied to account for the nonspecular regime.

ACKNOWLEDGMENTS

This work was sponsored by the German Science Foundation *Deutsche Forschungsgemeinschaft DFG*, which is gratefully acknowledged. The author would like to thank the anonymous reviewers for their valuable comments and suggestions.

- ¹ A. N. Norris and J. D. Achenbach, "Crack-tip diffraction in a transversely isotropic solid," in *Review of Progress in QNDE*, edited by D. O. Thompson and D. E. Chimenti (Plenum, New York, 1984), Vol. 3, pp. 133–141.
- ² A. N. Norris and J. D. Achenbach, "Elastic wave diffraction by a semi-infinite crack in a transversely isotropic solid," *Q. J. Mech. Appl. Math.* **37**, 565–580 (1984).
- ³ W. Lin and L. M. Keer, "Three-dimensional analysis of cracks in layered transversely isotropic media," *Proc. R. Soc. London, Ser. A* **424**, 307–322 (1989).
- ⁴ T. Kundu and A. Boström, "Elastic wave scattering by a circular crack in a transversely isotropic solid," *Wave Motion* **15**, 285–300 (1992).
- ⁵ C. H. Chapman and R. T. Coates, "Generalized born scattering in anisotropic media," *Wave Motion* **19**, 309–341 (1994).
- ⁶ R. K. N. D. Rajapakse and D. Gross, "Transient response of an orthotropic elastic medium with a cavity," *Wave Motion* **21**, 231–252 (1995).
- ⁷ P. A. Lewis, J. A. G. Temple, and G. R. Wickham, "Elastic wave diffraction at cracks in anisotropic materials," in *Review of Progress in QNDE*, edited by D. O. Thompson and D. E. Chimenti (Plenum, New York, 1996), Vol. 15, pp. 41–48.
- ⁸ J. Mattson, A. J. Niklasson, and A. Eriksson, "Three-dimensional ultrasonic crack detection in anisotropic materials," *Res. Nondestruct. Eval.* **9**, 59–79 (1997).
- ⁹ B. Ursin and M. Tygel, "Reciprocal volume and surface scattering integrals for anisotropic elastic media," *Wave Motion* **26**, 31–42 (1997).
- ¹⁰ M. Tygel and B. Ursin, "Weak-contrast and vertex diffractions in anisotropic elastic media," *Wave Motion* **29**, 363–373 (1999).
- ¹¹ M. Spies, "A general formulation for elastic wavefield modeling applied to arbitrarily oriented orthotropic media," *J. Acoust. Soc. Am.* (to be submitted).
- ¹² J. E. Gubernatis, E. Domany, and J. A. Krumhansl, "Formal aspects of the scattering of ultrasound by flaws in elastic materials," *J. Appl. Phys.* **48**, 2804–2807 (1977).
- ¹³ Y. H. Pao and V. Varatharajulu, "Huygens' principle, radiation conditions and integral formulas for the scattering of elastic waves," *J. Acoust. Soc. Am.* **59**, 1361–1371 (1976).

- ¹⁴V. T. Buchwald, "Elastic waves in anisotropic media," Proc. R. Soc. London, Ser. A **253**, 563–580 (1959).
- ¹⁵M. Spies and M. Kröning, "A computationally efficient modeling code for SH-wave propagation in austenitic welds using an explicit space-time Green's function," in *Review of Progress in QNDE*, edited by D. O. Thompson and D. E. Chimenti (Plenum, New York, 1996), Vol. 15, pp. 145–152.
- ¹⁶M. Spies and M. Kröning, "Accelerated space-time modeling of quasi-longitudinal waves in austenitic weld structures," in *Review of Progress in QNDE*, edited by D. O. Thompson and D. E. Chimenti (Plenum, New York, 1997), Vol. 16, pp. 1175–1182.
- ¹⁷Since $\underline{\mathbf{K}}_\alpha \cdot \underline{\mathbf{c}}_\alpha = \omega$ and $\underline{\mathbf{R}} = \hat{\underline{\mathbf{c}}}$, it is $\underline{\mathbf{K}}_\alpha \cdot \underline{\mathbf{R}} = \underline{\mathbf{K}}_\alpha \cdot \underline{\mathbf{c}}_\alpha R/c_\alpha = \omega R/c_\alpha$.
- ¹⁸M. Spies, "Transducer-modeling in general transversely isotropic media via point-source-synthesis. Theory," J. Nondestruct. Eval. **13**, 85–99 (1994).
- ¹⁹M. Spies and M. Kröning, "Theoretical evaluation of ultrasonic scattering by flaws in anisotropic media," in *Review of Progress in QNDE*, edited by D. O. Thompson and D. E. Chimenti (Plenum, New York, 1998), Vol. 17, pp. 57–63.
- ²⁰L. V. Atkinson, P. J. Harley, and J. D. Hudson, *Numerical Methods with Fortran 77: A Practical Introduction* (Addison-Wesley, Wokingham, 1989).
- ²¹J. H. M. T. van der Hijden, *Propagation of Transient Elastic Waves in Stratified Anisotropic Media* (North-Holland, Amsterdam, 1987).
- ²²J. A. Ogilvy and I. D. Culverwell, "Elastic model for simulating ultrasonic inspection of smooth and rough defects," Ultrasonics **29**, 490–496 (1991).

The noise emitted from vehicles at roundabouts

W. M. To

Department of Mechanical Engineering, Hong Kong University of Science and Technology,
Clear Water Bay, Kowloon, Hong Kong

T. M. Chan

Daniel Chan & Associates Ltd., Wanchai, Hong Kong

(Received 21 October 1999; accepted for publication 15 February 2000)

In Asian cities, many noise-sensitive receivers are located in the vicinity of highways and roundabouts because of the limited supply of usable land and the demand for good transportation. Although roundabouts can carry smooth traffic flow by minimizing the start-stop operations of drivers, the noise emitted from moving vehicles would definitely affect those people who live, work, or study in the vicinity. This article presents an analytical solution of the noise emitted from vehicles at roundabouts by assuming vehicles to be incoherent sources lying continuously on the roundabouts. A simple equation is introduced by which it is possible to relate the noise level at any distance from the center of the roundabout to the noise level measured at the center of the roundabout. © 2000 Acoustical Society of America. [S0001-4966(00)04105-9]

PACS numbers: 43.50.Lj, 43.50.Sr [MRS]

INTRODUCTION

The accurate prediction of noise levels due to road traffic is of great importance during a planning process. In the United States, the Federal Highway Administration developed a sophisticated computer model for the prediction and control of highway noise.^{1,2} In the United Kingdom, the Department of Environment publication "Calculation of Road Traffic Noise"³ describes a similar approach, derived from both field data and computer-simulation techniques, which enables L_{10} noise levels to be predicted. This method is also adopted in Hong Kong to predict L_{10} noise levels at sensitive receivers due to road traffic. A large amount of data on the noise from unrestricted, relative straight, road traffic gathered in the UK and Hong Kong^{4,5} indicated this prediction method to be generally reliable.

In Asian cities, a large number of residential buildings have been built right next to complicated road networks over the past decade. To ease traffic congestion, roundabouts have become one of the most popular choices adopted by town planners for carrying smooth traffic flow as they can minimize the start-stop operation of vehicles. As a result, many residential units and other sensitive receivers such as schools and playgrounds are located in the vicinity of highways as well as roundabouts. The noise levels at those receivers can be obtained by summing the contributions from highways and roundabouts. Lewis and James⁶ suggested in their paper that the roundabout, a circular source, could be divided into a series of contiguous arcs which subtended an angle of ϕ radians at center. The total noise level (L_{10} or L_{eq}) at the reception point could be obtained by summing the noise levels due to individual segment. It should be noted that several companies^{7,8} have developed commercial software in which a roundabout is modeled as many small in-line elements. With a sufficiently large number of such elements, this modeling will give an accurate representation. The required number of elements, for a specified accuracy, can be obtained by comparison to an analytical solution for the noise emission.

This article presents an analytical solution of the noise emitted from vehicles at roundabouts. The results show that the segmentation of a roundabout suggested by Lewis and James will introduce a small discretization error when a receiver is located in the vicinity.

I. MATHEMATICAL DERIVATION

When there are many sound sources in a row, such as vehicles on a motorway, they can be considered as an infinite line source on the ground. It is also assumed that these vehicles are represented by an infinite set of incoherent point sources lying continuously on the line in random phases, resulting in wave behavior such as constructive and destructive interference to be ignored. In a semi-free-field environment, the sound waves spread in a cylindrical form around the line source on the ground that is the axis of the cylinder. When the sound power per unit length of the line source is W , the sound intensity I at distance d is given by

$$I = \frac{W}{\pi d}. \quad (1)$$

Similarly, traffic on a roundabout can be represented as a circular line source on the ground. It is a valid assumption when the size of the roundabout is huge and most vehicles will emit more or less the same noise after entering the circular path. In this sense, the roundabout is considered as a statistically stationary and ergodic system. As the point sources can be considered as lying continuously from 0 to 2π , as shown in Fig. 1, the intensity at the reception point P at a distance r from the source dE becomes

$$dI = \frac{WR d\theta}{2\pi r^2}. \quad (2)$$

The overall intensity at the reception point P due to the circular source is written as

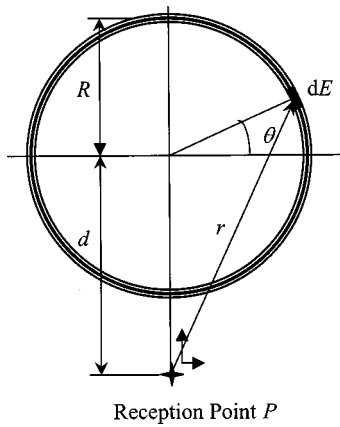


FIG. 1. A roundabout.

$$I = 2 \int_{\pi/2}^{\pi/2} dI = \int_{\pi/2}^{\pi/2} \frac{WR}{\pi r^2} d\theta. \quad (3)$$

The separation r can be expressed as a function of the radius of roundabout R , the distance d between the center of the roundabout and the reception point, and the angle θ . From the geometry, one can show that

$$r = (x^2 + y^2)^{1/2} = (R^2 + d^2 + 2dR \sin \theta)^{1/2}, \quad (4)$$

where $x = R \cos \theta$ and $y = d + R \sin \theta$.

Substituting Eq. (4) into Eq. (3), one can obtain

$$I = \frac{WR}{\pi} \int_{\pi/2}^{\pi/2} \frac{d\theta}{(R^2 + d^2 + 2dR \sin \theta)}. \quad (5)$$

When d is greater than R , as shown in Fig. 1, the integration on the right-hand side gives

$$I = \frac{2WR}{\pi(d+R)(d-R)} \left[\tan^{-1} \frac{d+R}{d-R} - \tan^{-1} \frac{-(d-R)}{d+R} \right]. \quad (6)$$

If the reception position is at the center of the roundabout, Eq. (3) can be rewritten as

$$I_{\text{center}} = \int_0^{2\pi} \frac{WR}{2\pi R^2} d\theta = \frac{W}{R}. \quad (7)$$

Now, one can determine the difference in sound-pressure level ΔSPL between the noise level at an arbitrary reception point and the noise level at the center of the roundabout. It is defined by

$$\Delta\text{SPL} = 10 \log \frac{I}{I_{\text{center}}} = 10 \log \left\{ \frac{2R^2}{\pi(d+R)(d-R)} \times \left[\tan^{-1} \frac{d+R}{d-R} - \tan^{-1} \frac{-(d-R)}{d+R} \right] \right\}. \quad (8)$$

If the distance d is normalized to the radius of roundabout R , Eq. (8) can be expressed as

$$\Delta\text{SPL} = 10 \log \left\{ \frac{2}{\pi \left(\frac{d}{R} + 1 \right) \left(\frac{d}{R} - 1 \right)} \times \left[\tan^{-1} \frac{\frac{d}{R} + 1}{\frac{d}{R} - 1} - \tan^{-1} \frac{-(\frac{d}{R} - 1)}{\frac{d}{R} + 1} \right] \right\}. \quad (9)$$

It should be noted that in practice the reception point may not be at the same level of the roundabout. In this situation, the general expression for the separation r should be $(x^2 + y^2 + z^2)^{1/2}$, where z is the height of reception point above the roundabout. As a result, the generalized equation for the difference in sound-pressure level is

$$\Delta\text{SPL} = 10 \log \left\{ \frac{2}{\pi A} \left[\tan^{-1} \frac{\left(\frac{d}{R} + 1 \right)^2 + \frac{z^2}{R^2}}{A} - \tan^{-1} \frac{-(\frac{d}{R} - 1)^2 - \frac{z^2}{R^2}}{A} \right] \right\}, \quad (10)$$

where $A = [(d^2/R^2 - 1)^2 + 2(d^2/R^2 + 1)(z^2/R^2) + z^4/R^4]^{1/2}$.

II. DISCRETE MODELS

In the past, researchers^{6,9} suggested that the roundabout could be divided into a series of contiguous arcs. The total noise level at the reception point can be calculated by summing the noise levels due to individual segment. The difference in sound pressure level ΔSPL between the noise level at an arbitrary reception point and the noise level at the center of the roundabout is now calculated from

$$\Delta\text{SPL} = \text{SPL}_{\text{at } P} - \text{SPL}_{\text{center}} = 10 \log \frac{10^{\text{SPL}_{\text{at } P}/10}}{10^{\text{SPL}_{\text{center}}/10}}, \quad (11)$$

where

$$\text{SPL}_{\text{at } P} \quad \text{or} \quad \text{SPL}_{\text{center}} = 10 \log \sum_{k=1}^{2\pi/\phi} 10^{L_k/10}.$$

It should be noted that L_k is constant when the point of interest is at the center of the roundabout. L_k is a function of location for P depending on the radius of roundabout and the separation between the center of the roundabout and the reception point.

If the circular path is equally divided into eight sections, as shown in Fig. 2, the sound intensity at P will be approximated by

$$I = \sum_{k=1}^8 \frac{\frac{W(2\pi R)}{8}}{2\pi \left[R^2 + d^2 + 2dR \sin \frac{\pi(2k-1)}{8} \right]}. \quad (12)$$

As the sound intensity at the center of the roundabout is W/R , the difference in sound-pressure level is calculated from the intensity values

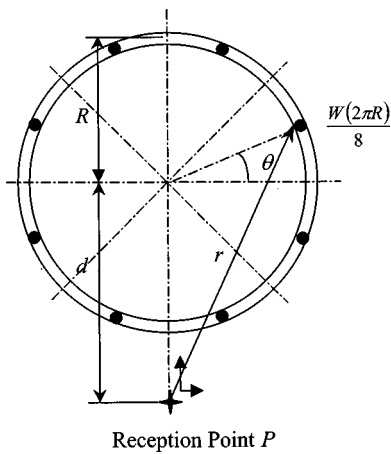


FIG. 2. An 8-segment discrete model for the noise emitted from a roundabout.

$$\begin{aligned}\Delta\text{SPL} &= 10 \log \frac{I}{I_{\text{center}}} \\ &= 10 \log \sum_{k=1}^8 \frac{R^2}{8 \left[R^2 + d^2 + 2dR \sin \frac{\pi(2k-1)}{8} \right]}\end{aligned}\quad (13)$$

This equation can be generalized as

$$\begin{aligned}\Delta\text{SPL} &= 10 \log \frac{I}{I_{\text{center}}} \\ &= 10 \log \sum_{k=1}^n \frac{R^2}{n \left[R^2 + d^2 + 2dR \sin \frac{\pi(2k-1)}{n} \right]}\end{aligned}\quad (14)$$

where $n=4, 6, 8, 10, 12, 16, 18, 20$, etc. Four is the minimum number because a circle shall be divided by at least four quadrants and n is an even number to make sure that 2π is easily divisible.

When the distance d is normalized to the radius of roundabout R , Eq. (14) is rewritten as

$$\Delta\text{SPL} = 10 \log \sum_{k=1}^n \frac{1}{n \left[1 + \left(\frac{d}{R} \right)^2 + 2 \frac{d}{R} \sin \frac{\pi(2k-1)}{n} \right]}\quad (15)$$

III. NUMERICAL SIMULATIONS

The difference in sound-pressure level between the noise level at some point of interest outside of the roundabout and the noise level at the center of the roundabout for the analytical model and discrete models using 4, 8, and 16 segments are given in Fig. 3 with ratios d/R of from 1.1 to 10. One can observe that at long ranges (d/R equal to or greater than 3), the discrete models will give a very accurate estimation when compared to the difference in sound-pressure level predicted using the analytical model. By inspecting the numerical values of those curves shown in Fig. 3, it was found

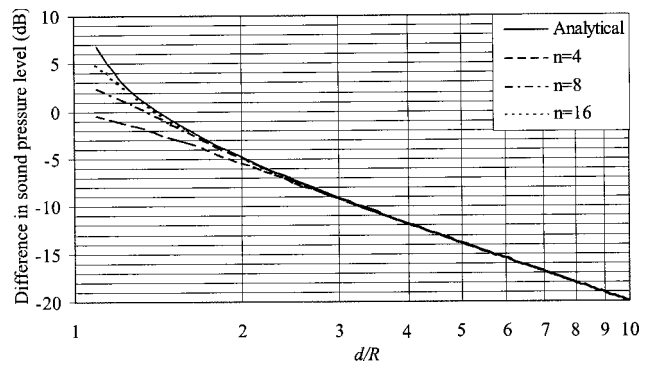


FIG. 3. The difference in sound-pressure level between the noise level at some point of interest outside of the roundabout and the noise level at the center of the roundabout.

that when d/R equals 1.8, the 16-segment discrete model gives the same prediction as the one obtained using the analytical model. The eight-segment discrete model produces an accurate prediction when d/R is equal to or greater than 2.4. These curves also confirm a well-known fact that at long ranges the decrease of sound-pressure level is of about 6 dB per doubling of distance from the roundabout.

When some noise-sensitive receivers are located in the vicinity of roundabouts, change in noise level as a function of the horizontal distance from the center of the roundabout and height above the roundabout should be considered. Figure 4 shows the difference in sound-pressure level between the noise level at the reception point and the noise level at the center of the roundabout for propagation over hard intervening ground. The difference in sound-pressure level is obtained using Eq. (10) and takes into account distance from the roundabout center and height above roundabout of the reception point.

Now, we can assume that a noise-monitoring station is positioned at the center of a roundabout with a radius of 40 m and the measured sound-pressure level is 72 dB. The corresponding root-mean-square pressure is 0.0796 Pa. As the intensity is equal to the square of the root-mean-square pressure divided by the characteristic acoustic impedance of air (about 410 Rayls), the intensity at the center of the roundabout is $1.546 \times 10^{-5} \text{ W/m}^2$. Using Eq. (7), one can deter-

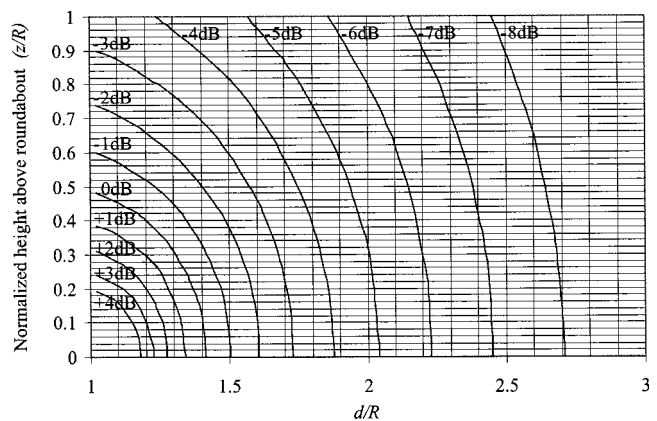


FIG. 4. Change in sound-pressure level as a function of the horizontal distance from the center of the roundabout and height above the roundabout (with respect to the noise level at the center of the roundabout and $z=0$).

mine that the sound power of moving vehicles is 6.185×10^{-4} W/m. If a noise-sensitive receiver is located 16 m from the roundabout (or 56 m from the center of this roundabout), one can determine that the noise level is (72 dB + 0.18 dB =) 72.18 dB. The noise level is (72 dB - 0.41 dB =) 71.59 dB when the eight-segment discrete model is employed.

IV. CONCLUSION

Roundabouts have become one of the most popular choices adopted by town planners for carrying smooth traffic flow as they can minimize the start-stop operation of drivers. However, many people are adversely affected by the noise emitted from vehicles at roundabouts as they live, study, or work in the vicinity.

In this article, we derived a simple equation by which it is possible to relate the noise level at any distance from the center of the roundabout to the noise level measured at the center of the roundabout (or the sound power of moving vehicles estimated from the measured sound-pressure level). This equation is applicable to places such as Hong Kong, in which noise-monitoring stations have been operated for years and some prediction models for planning purposes are based on existing measured parameters including noise lev-

els. Numerical simulations showed that the eight-segment discrete model introduced an error of only 0.6 dB when d/R was equal to 1.4. This discretization error is quite small when compared to the uncertainties of several dB caused by meteorological effects and scattering from buildings in outdoor sound propagation. One can show that the same degree of accuracy can be achieved when the number of small inline elements is equal to eight.

¹A. J. Angiola, V. M. Lee, and R. A. Michalove, "Field evaluation of the FHWA highway traffic noise prediction model," *Proceedings of Inter-Noise 80*, pp. 1137-1140 (1980).

²W. Bowlby, "Integration of a traffic noise model and AutoCAD," *Proceedings of Microcomputer in Transportation*, pp. 774-785 (1993).

³The UK Department of the Environment, *Calculation of Road Traffic Noise* (Her Majesty's Stationery Office, London, 1975).

⁴J. W. Sargent, "Data from free flow traffic noise measurements," *J. Sound Vib.* **57**(2), 307 (1978).

⁵I. W. K. Ng, S. W. H. Wong, and C. C. Chiu, "Prediction and field validation of road traffic noise," *Proceedings of the IOA Autumn Conference—Environmental Noise* (1993).

⁶P. T. Lewis and A. James, "Noise levels in the vicinity of traffic roundabouts," *J. Sound Vib.* **72**(1), 51-69 (1980).

⁷Wolfel, IMMI 5.0 Noise Prediction and Noise Mapping Software.

⁸DataKustik GmbH, CADNA/A® Computer Aided Noise Abatement Program for the prediction and assessment of outdoor noise levels.

⁹P. M. Nelson, *Transportation Noise Reference Book* (Butterworth, London, 1987), Chap. 10.

Effect of membrane motor on the axial stiffness of the cochlear outer hair cell

K. H. Iwasa^{a)}

Biophysics Section, Laboratory of Cellular Biology, National Institute on Deafness and Other Communication Disorders, National Institutes of Health, 9 Center Drive, Bethesda, Maryland 20892-0922

(Received 9 November 1999; accepted for publication 17 February 2000)

Voltage-dependent motility of the cell body of the cochlear outer hair cell has been successfully explained by an “area motor” model, which is based on tight coupling between charge transfer across the membrane and changes in the membrane area. The model might appear to contradict a recent report that the axial stiffness of the outer hair cell is dependent on the membrane potential. Here, it is shown that the area motor model indeed predicts a considerable voltage dependence of the axial stiffness of the outer hair cell. The predicted changes resemble those observed experimentally, although the predicted changes are smaller at the two extremes of the membrane potential. [S0001-4966(00)04805-0]

PACS numbers: 43.64.Ld, 43.64.Kc, 43.64.Bt [LHC]

INTRODUCTION

The outer hair cell has a cylindrical cell body, which undergoes length changes in response to changes in the membrane potential (Brownell *et al.*, 1985). This motility is thought to be associated with the biological function of the cell in increasing frequency tuning and in modulating the gain of the ear. The amplitude of the change reaches about 5% of its length (Ashmore, 1987). Force produced by the cell associated with this motility (Hallworth, 1995; Iwasa and Adachi, 1997) is up to 0.1 nN/mV. This motility can be understood by a membrane motor model in which charge transfer across the membrane is coupled with mechanical displacements in the membrane (Iwasa, 1993; Dallos *et al.*, 1993; Santos-Sacchi, 1993; Iwasa, 1994; Iwasa and Adachi, 1997). The validity of such a motile mechanism, specifically that the mechanical displacements are area changes, referred to as the area motor model (Iwasa, 1994), was further confirmed recently by the observation that charge transfer is inhibited by constraining the membrane area (Adachi and Iwasa, 1999).

Intriguingly, it has recently been reported that this motility is accompanied by changes in the axial stiffness of the cell body (He and Dallos, 1999). This finding raises a possible importance of stiffness changes in the biological role of the cell. The observation also raises two interesting questions regarding the underlying cellular mechanism. How can the axial stiffness depend on the membrane potential? How are the changes in the stiffness related to the motile mechanism?

In this paper, I show that the area motor model predicts that the axial stiffness decreases with depolarization. The predicted magnitude may exceed 25%, which is appreciable but smaller than the observed changes by a factor of 2 or 3.

I. AREA MOTOR MODEL

A. Assumptions

The area motor model consists of the following three main assumptions (Iwasa, 1994; Iwasa and Adachi, 1997):

- (1) The total mechanical displacement of the cylindrical cell consists of elastic displacements and motor displacements. This assumption can be represented by the total axial strain ϵ_z and the total circumferential strain ϵ_c

$$\epsilon_z = \epsilon'_z + na_z P_l, \quad (1)$$

$$\epsilon_c = \epsilon'_c + na_c P_l, \quad (2)$$

where ϵ' 's are elastic strains, n the number density of the motor, and P_l the fraction of the motor in the expanded state. The quantities a_z and a_c are displacement per motor unit in the axial direction and the one in the circumferential direction, respectively. It is assumed that the stiffness of the motor is the same in the two states and is independent of the membrane potential.

- (2) Tension σ_z and σ_c of the membrane in the two directions due to elasticity must be balanced with tension due to pressure difference P across the membrane and external force F

$$\sigma_z = d_1 \epsilon'_z + c \epsilon'_c = (1/2)RP + f, \quad (3)$$

$$\sigma_c = c \epsilon'_z + d_2 \epsilon'_c = RP. \quad (4)$$

Here, d_1 , d_2 , and c are elastic moduli of the membrane. Pressure inside the cell is higher than the outside by P , R is the radius of the cell, and tension f is due to an externally applied axial force [$f = F/(2\pi R)$]. The validity of these equations is limited to small strains.

- (3) The fraction of the motor P_l in the extended conformation follows Boltzmann distribution,

$$P_l = \frac{1}{1 + \exp[\beta(F_0 + qV + \sigma_z a_z + \sigma_c a_c)]}. \quad (5)$$

Here, F_0 is a constant and $\beta = 1/(k_B T)$, with k_B being the Boltzmann constant and T the temperature. The quantity q is the difference in charge in the two states and V is the membrane potential. The motor area differ-

^{a)}Electronic mail: iwasa@nih.gov

ence a_z , a_c , and tension σ in the two directions are described earlier. The energy difference between the two motor states thus include both electrical energy and mechanical energy.

B. Axial strain

If we can impose a condition that the cell volume is kept constant,

$$\epsilon_z + 2\epsilon_c = \epsilon_{v0}, \quad (6)$$

these equations reduce to a single nonlinear equation with respect to a single variable and can be solved numerically for a given value of the membrane potential V and external tension f . For example, the axial strain ϵ_z can be expressed as

$$\epsilon_z = \frac{4f - 2\epsilon_{v0}(2c - d_2) - 2(2f_z + f_c)P_l}{4d_1 - 4c + d_2}. \quad (7)$$

Here, f_z and f_c are constants that involve motor parameters (n , a_z , and a_c), and the elastic moduli of the membrane. They determine the effectiveness of the motor in generating tension. An analogous expression can be obtained for ϵ_c , the circumferential strain.

The terms for the mechanical energy of a motor unit, $\sigma_z a_z + \sigma_c a_c$, which appear in Eq. (5), depend on pressure and axial force

$$\sigma_z a_z + \sigma_c a_c = RP(a_z/2 + a_c) + a_z f. \quad (8)$$

With Eqs. (1), (2), and (4), cell pressure P is represented by

$$RP = c\epsilon_z + d_2\epsilon_c - (ca_z + d_2a_c)nP_l. \quad (9)$$

Thus, the mechanical energy $\sigma_z a_z + \sigma_c a_c$ can be expressed by the strains ϵ_z and ϵ_c , which in turn are determined by P_l and f as shown by Eq. (7). Substitution of the resulting expression for $\sigma_z a_z + \sigma_c a_c$ in Eq. (5) yields a nonlinear equation with respect to P_l . This equation shows that the motor state P_l is determined by a Boltzmann function of the motor state P_l as well as of V and F . This is because the motor is affected by membrane tension that it generates. The equation can be solved numerically to obtain P_l . Then the axial strain ϵ_z for given values of V and F is obtained with Eq. (7).

II. AXIAL STIFFNESS OF THE CELL BODY

The stiffness of the cell body can be determined by evaluating the increase of the length strain $\Delta\epsilon_z$ in response to an increase in the axial force $2\pi R\Delta f$. Conformational changes in the motor can make the axial stiffness of the cell dependent on the membrane potential.

The importance of conformational changes on the axial stiffness of the cell can be examined numerically by evaluating the membrane potential dependence and load dependence of the axial stiffness using stiffness moduli and motor parameters previously determined. Two parameter sets are chosen for the examination. The first set (A) is the one given in Iwasa (1994). It is based on mechanical isotropy of the lateral membrane and the elastic moduli are determined by the stress-strain relationship in an experiment in which cell pressure was changed (Iwasa and Chadwick, 1992). That is a special case when $d_1 = d_2$. The second set (B) is the one

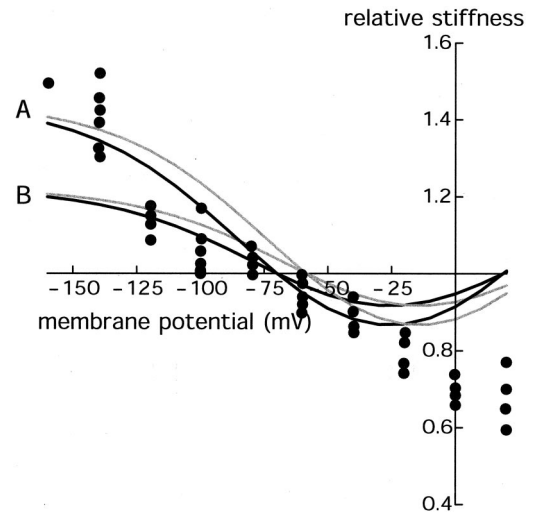


FIG. 1. Voltage dependence of the axial stiffness. Relative changes of the axial stiffness is plotted against the membrane potential. The filled circles are taken from He and Dallos, 1999. The solid lines represent theoretical values based on the area motor model. The motor parameters are $q = 0.85 e$, the density of the motor n is $3.7 \times 10^3 \mu\text{m}^2$, the area displacement of a motor unit is $a_z = 8.0 \text{ nm}^2$ in the axial direction and $a_c = -1.5 \text{ nm}^2$ in the circumferential direction. Two sets of values for the elastic moduli were used. Set A: $d_1 = 0.077$, $d_2 = 0.077$, and $c = 0.063 \text{ (N/m)}$. The axial stiffness is normalized to $0.73 \mu\text{N}$ per unit strain. It assumes isotropy and is consistent with a pressure experiment (Iwasa and Chadwick, 1992). Set B: $d_1 = 0.046$, $d_2 = 0.068$, and $c = 0.046 \text{ (N/m)}$. The axial stiffness is normalized to $0.42 \mu\text{N}$. This set is consistent with the pressure experiment (Iwasa and Chadwick, 1992) and an axial stiffness experiment (Iwasa and Adachi, 1997). The axial loading is near zero (dark lines) and 5 nN (light lines).

given in Iwasa and Adachi (1997). It assumes mechanical orthotropy of the lateral membrane and the elastic moduli are determined by stress-strain relationship using pressure and the axial stiffness at -75 mV , which is about 500 nN per unit strain (Iwasa and Adachi, 1997). In both sets, the motor parameters are determined to satisfy the voltage- and pressure dependence of the membrane capacitance and the amplitude of voltage-dependent length changes (Iwasa, 1994; Iwasa and Adachi, 1997). The plot for each parameter set is normalized to the stiffness at -70 mV for small load. In both cases the predicted stiffness is higher when the cell is hyperpolarized and decreases as the cell depolarizes (Fig. 1). The magnitude of the change depends on the parameter sets used: It is about 50% for the set (A) and 25% for the set (B). The effect is greater for the set (A) primarily because the axial stiffness is twice as large.

The effect of the membrane potential on the axial stiffness is the consequence of tension dependence of the motor. An increase in axial tension due to externally applied axial-stretching force favors the extended state of the motor and this effect facilitates elongation, resulting in the apparent reduction in the axial stiffness of the cell. The sensitivity of the motor to tension is highest when motor state P_l is near 0.5, namely at the membrane potential at which the extended and shrunk conformations are equal. Thus, the stiffness has a minimum, just as the membrane capacitance has a maximum, between -20 and 0 mV (depending on the load), and a further depolarization results in an increase in the axial stiffness.

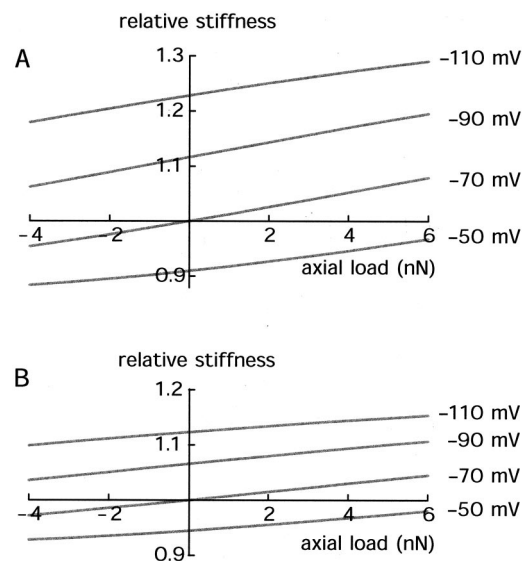


FIG. 2. Dependence of the axial stiffness on load (negative load is compressive). The two sets, A (top) and B (bottom), of parameter values are the same as in Fig. 1. The axial stiffness is normalized in the same manner as in Fig. 1. The values of membrane potential are indicated on the right.

Another outcome of the area motor model is that the apparent axial stiffness is expected to depend on the load. If a steady force is applied to the cell, the resulting tension changes the motor state. This effect also changes the axial stiffness (Fig. 1). Such an effect may not be as pronounced as that of the membrane potential (Fig. 2) because of the range of the force that can be applied to the cell would be limited.

III. DISCUSSION

Experimentally observed changes in the axial stiffness resemble the prediction (Fig. 1). The stiffness decreases as the membrane potential goes up from the resting level between -70 and -80 mV. It increases as the membrane potential goes down (hyperpolarizes) from the resting level.

The experimental data show larger changes in the stiffness than the prediction, particularly in the hyperpolarizing direction. The difference in the depolarized range is also appreciable. While the predicted stiffness change shows a minimum at about -20 mV, such a minimum is not clear for the experimental data.

There is a factor that could complicate experimental determination of the axial stiffness. It has been observed that cell volume increases when the cell is hyperpolarized and decreases when depolarized (Shehata *et al.*, 1991) if ion channels are not blocked. Such a volume effect, which ac-

companies length changes, is expected to be time dependent and more pronounced while the membrane potential is held at voltage levels farther away from its resting level of about -75 mV. When the membrane potential is brought back to the resting level, volume recovery takes some time. Some of the traces [in Fig. 3(a)] in the report by He and Dallos (1999) appear to be consistent with such an interpretation. Volume increases during hyperpolarizing voltage steps are expected to reduce the length of the cell (Iwasa and Chadwick, 1992). It thus reduces the compressive load in their recording configuration. Because a reduction in compressive load (negative load in Fig. 2) increases the axial stiffness, the effect is an increase in the stiffness during hyperpolarizing pulses. The opposite is expected during depolarizing pulses. In both cases the effect reduces the difference between the data and the prediction. However, this effect is not significant if the voltage pulses are short. That appears to be the case because the shifts of the probe position are not large.

In conclusion, the mechanism based on the area motor model does not predict stiffness changes as large as the observed ones. Nonetheless, it would be fair to state that tension-dependent conformational transitions of the area motor must have a significant contribution to the observed stiffness changes.

- Adachi, M. and Iwasa, K. H. (1999). "Electrically driven motor in the outer hair cell: Effect of a mechanical constraint," *Proc. Natl. Acad. Sci. USA* **96**, 7244–7249.
- Ashmore, J. F. (1987). "A fast motile response in guinea-pig outer hair cells: The molecular basis of the cochlear amplifier," *J. Physiol. (London)* **388**, 323–347.
- Brownell, W., Bader, C., Bertrand, D., and Ribaupierre, Y. (1985). "Evoked mechanical responses of isolated outer hair cells," *Science* **227**, 194–196.
- Dallos, P., Hallworth, R., and Evans, B. N. (1993). "Theory of electrically driven shape changes of cochlear outer hair cells," *J. Neurophysiol.* **70**, 299–323.
- Hallworth, R. (1995). "Passive compliance and active force generation in the guinea pig outer hair cell," *J. Neurophysiol.* **74**, 2319–2328.
- He, D. Z. Z., and Dallos, P. (1999). "Somatic stiffness of cochlear outer hair cells is voltage-dependent," *Proc. Natl. Acad. Sci. USA* **96**, 8223–8228.
- Iwasa, K. H. (1993). "Effect of stress on the membrane capacitance of the auditory outer hair cell," *Biophys. J.* **65**, 492–498.
- Iwasa, K. H. (1994). "A membrane model for the fast motility of the outer hair cell," *J. Acoust. Soc. Am.* **96**, 2216–2224.
- Iwasa, K. H., and Adachi, M. (1997). "Force generation in the outer hair cell of the cochlea," *Biophys. J.* **73**, 546–555.
- Iwasa, K. H., and Chadwick, R. S. (1992). "Elasticity and active force generation of cochlear outer hair cells," *J. Acoust. Soc. Am.* **92**, 3169–3173.
- Santos-Sacchi, J. (1993). "Harmonics of outer hair cell motility," *Biophys. J.* **65**, 2217–2227.
- Shehata, W. E., Brownell, W. E., and Dieler, R. (1991). "Effects of salicylate on shape, electromotility and membrane characteristics of isolated hair cells from guinea pig cochlea," *Acta Oto-Laryngol.* **111**, 707–718.

PROGRAM OF

The 139th Meeting of the Acoustical Society of America
Westin Peachtree Plaza • Atlanta, Georgia • 30 May–3 June 2000

1a TUE. AM

NOTE: All Journal articles and Letters to the Editor are peer reviewed before publication. Program abstracts, however, are not reviewed before publication, since we are prohibited by time and schedule.

TUESDAY MORNING, 30 MAY 2000

ENGLISH ROOM, 9:00 A.M. TO 12:40 P.M.

Session 1aAO

Acoustical Oceanography and Underwater Acoustics: Acoustical Surveys of Ocean Bottom Geology I

Christian de Moustier, Chair

Marine Physical Laboratory, Scripps Institution of Oceanography, 9500 Gilman Drive, La Jolla, California 92093-0205

Chair's Introduction—9:00

Invited Papers

9:05

1aAO1. Geological analysis of acoustic subbottom profiles as a tool for prediction of acoustic property distribution. Louis R. Bartek (Dept. of Geological Sci., Univ. of North Carolina, Chapel Hill, CB #3315, Mitchell Hall, Chapel Hill, NC 27599, bartek@email.unc.edu)

Geoscientists have developed analytical techniques for extraction of geologic and acoustic data from acoustic profiles as well as prediction of how these properties may vary in areas where data are lacking and therefore these may be of interest to acousticians. Analysis of reflection termination patterns in subbottom profile data are used to identify sediment deposited contemporaneously by a set of related depositional environments. A variation in the interval that separates reflections is usually related to a change in bed thickness. Closely spaced reflections are associated with environments of deposition where the mode of deposition changes frequently and those with widely spaced reflections are associated with environments where the mode of deposition rarely changes. Continuity of reflections is related to the lateral continuity of beds. Environments that have relatively short lateral continuity, such as terrestrial environments like river valleys that have been flooded, and many marginal marine environments are associated with reflections that have relatively short lateral continuity. These variations in separation and lateral continuity of beds and termination and orientation of beds that are all linked to variations in reflection character are produced by depositional processes that are associated with depositional systems and therefore distributed in a relatively predictable manner.

10:05

1aAO2. Seafloor mapping and 3-D visualization: An integration of technologies providing new insights into seafloor processes. Larry A. Mayer (Ctr. for Coastal and Ocean Mapping, Univ. of New Hampshire, Durham, NH 03824, lmayer@cisunix.unh.edu) and Christian de Moustier (Scripps Inst. of Oceanogr., La Jolla, CA 92093-0205)

The past few years have seen rapid advancements in sonar development, in vessel orientation and positioning technology, and in our ability to manage and visualize the massive amounts of data that modern seafloor sonars are capable of collecting. The confluence of these technologies has resulted in revolutionary changes in our ability to map and visualize the seafloor. Until recently, maps of the seafloor were based on the interpolation (typically through the construction of contours) of sparsely spaced discrete soundings or lines of soundings. With the advent of multibeam or swath-sonar technology, we can now produce near-100% coverage of large areas of the seafloor with high resolution (both vertical and horizontal) in a timely and cost-effective manner. More recently multibeam sonars have also begun to collect coregistered acoustic backscatter information (and interferometric sidescan sonars have begun to collect coincident bathymetric information), potentially providing information on the composition (texture or material type) as well as the morphology of the seafloor. When combined with modern interactive 3-D visualization tools, these new data sets result in a radically different representation of the seafloor than that provided by traditional echo sounders that can lead to unprecedented levels of understanding of seafloor processes.

Contributed Papers

10:55

1aAO3. Multibeam bathymetry and Deep Tow sidescan sonar surveys SW of San Clemente Island, CA. Christian de Moustier (Marine Physical Lab., Scripps Inst. of Oceanogr., 9500 Gilman Dr., La Jolla, CA 92093-0205, cpm@mpl.ucsd.edu)

In March 1999, about 170 square nm of bathymetry were mapped from the 40- to 700-m contour lines using a 50-kHz Elac BottomChart multi-beam echo-sounder temporarily installed aboard Scripps' R/V NEW HORIZON. This survey provided the base map for a near-bottom 110-kHz sidescan sonar and 4-kHz subbottom profiler survey with the Marine Physical Lab. Deep Tow package. Notable seafloor geological features include a broad terrace, left over from the last glacial low sea level stand and sloping gently from the island's shore to the 120-m depth contour, canyons and sediment aprons below, steep rocky extensions of China Point (island's southern-most point), and the sheer wall of the San Clemente Escarpment to the east. The sidescan sonar map has poor contrast over the sandy bottom along the island's western shore, whereas the rock spurs off China Point and a number of erosional channels carving the slopes between 200- and 800-m depth correspond to areas of high acoustic backscatter. Deep Tow subbottom profiles along the southern slope provide evidence of erosional channels cutting through sediment bedforms that are presumably compressional side effects of the strike-slip faulting along the San Clemente Escarpment.

11:10

1aAO4. Performance of one-dimensional chirp sonar inversion in complex environments. Altan Turgut (Naval Res. Lab., Acoust. Div., Washington, DC 20375, turgut@wave.nrl.navy.mil)

Chirp sonar surveys can provide subbottom images with cm-scale vertical resolution and tens of meters of penetration into the sediment. Recently, chirp sonar reflection amplitude data have been used for the inversion of subbottom attenuation and impedance profiles [e.g., Panda *et al.*, J. Acoust. Soc. Am. **96**, 3022–3035 (1994)]. Furthermore, by making use of physics-based sediment models and correlation between certain parameters, real-time estimation of sediment properties such as density, porosity, and sound-speed profiles has become possible [Turgut and Wolf, J. Acoust. Soc. Am. **98**, 2972(A) (1995)]. In this work, effects of fine layering and lateral inhomogeneities on the performance of chirp sonar inversion have been studied. Simulated reflection amplitude data were obtained by using a 2-D pseudospectral method for a poro-viscoacoustic sediment model based on the Biot theory. Limits of one-dimensional inversion were discussed in this framework and several recommendations were made to improve the data collection methods in more complex environments. [Work supported by ONR.]

11:25

1aAO5. Inversion of reverberation data for rapid environmental assessment. John R. Preston (ARL, Penn State Univ., P.O. Box 30, State College, PA 16804, preston@ciao.arl.psu.edu)

From 1996 through 1998 SACLANTCEN participated in three rapid environmental assessment (REA) trials, including experiments for bottom reverberation. Eight sites were visited: near Sicily, in the Ionian Sea, and near the Straits of Gibraltar. SUS charges were used as sources. The receivers were horizontal arrays (four apertures) spanning a wide frequency

range. Sets of nearly monostatic recordings were analyzed in frequency bands from 80 to 4000 Hz. Previous results by the author [Berlin ASA meeting (1999)] showed polar plots of the beam time series superimposed on bathymetric charts, revealing a number of scattering features not on the charts and that directional reverberation measurements are a useful remote-sensing tool. Another objective of the reverberation experiments was to quickly invert for estimates of scattering and geo-acoustic parameters that can be used in a wider variety of conditions. A manual inversion procedure was originally used to obtain at-sea results. New work uses a simulated annealing algorithm giving more formal inversion results. The horizontal array data are compared with the generic sonar model (GSM) predictions for both the manual and automated inverse schemes. A summary of the estimated geo-acoustic parameters using both methods is presented. [Work supported by ONR and DREA.]

11:40

1aAO6. Sediment roughness estimation with ROV mounted lasers. Nicholas P. Chotiros, James N. Piper, Brian S. Ward, D. Eric Smith (Appl. Res. Labs., Univ. of Texas, P.O. Box 8029, Austin, TX 78713-8029), David C. Gage, Roger A. Banks, and David L. Smith (Univ. of Texas, Austin, TX 78713-8029, chotiros@arl.utexas.edu)

A simple method of measuring underwater sediment surface roughness was tested over a sandy sediment off the west coast of Florida, as part of the Sediment Acoustics Experiment (SAX99). The process was integrated into an acoustic data acquisition system, and allowed roughness measurements of the seafloor without interference with the acoustic data acquisition process. The method employed lasers mounted on a remotely operated vehicle (ROV). Each laser was a red He–Ne class IIIa with output power less than 5 mW. A cylindrical lens was used to spread the laser beam into fan, which was directed downward to produce a stripe on the sediment. Viewing from an oblique angle, the profile of the sediment directly under the stripe is obtained. The information was recorded on videotape, digitized, and processed to obtain roughness statistics. [Work supported by ONR, Ocean Acoustics.]

11:55

1aAO7. Continuous P-SV coupling and strongly dispersive surface waves in soft marine sediments. Oleg A. Godin (NOAA/ETL, R/E/ET1, 325 Broadway, Boulder, CO 80303, ogodin@etl.noaa.gov) and David M. F. Chapman (Defence Res. Establishment Atlantic, Dartmouth, NS B2Y 3Z7, Canada)

In the upper tens to few hundred meters of clay and silt marine sediments with high porosity, the shear wave velocity is much smaller than the compressional wave velocity. The shear velocity has very large gradients close to the ocean floor leading to strong P-SV coupling in such "soft" sediments. A theory of elastic wave propagation in continuously stratified soft sediments is developed that fully accounts for the coupling. Elastic waves in soft sediments consist of "fast" waves propagating with velocities close to the compressional velocity and "slow" waves propagating with velocities on the order of the shear velocity. For the slow waves, the theory predicts existence of surface waves at the ocean/sediment boundary. An explicit, exact solution is obtained for the surface waves in the case of linear increase of shear rigidity with depth. Asymptotic and perturbation techniques are used to extend the result to more general environ-

ments. Velocity of vertically polarized interface waves is sensitive to sediment density and shear rigidity. Theoretical dispersion relations agree well with available experimental data and are shown to lead to a simple and robust inversion of interface waves travel times for shear-velocity profiles in the sediment.

12:10

1aAO8. Shear wave velocity measurements of layered sediment models using radiation impedance. Masao Kimura (Dept. of Ocean Eng., Tokai Univ., 3-20-1 Orido, Shimizu, Shizuoka 424-8610, Japan)

Shear wave velocity is important in characterizing surface marine sediments. It was demonstrated that the shear wave velocity in viscoelastic media can be obtained using the value of the frequency at which the radiation reactance becomes zero, and the shear wave velocity of sands can be measured. Moreover, the effect of the size of the radius of the circular vibrating plate on the measured values of shear wave velocities in sands is shown. In this study, the shear wave velocities in layered sediment model are measured. The sediment model is a two-layered model, which is composed of clay as the upper medium and sand or urethane rubber as the lower medium. The effect of the thickness of the upper medium on the measured values of shear wave velocities is also considered.

12:25

1aAO9. Geoacoustic properties of sediments determined from measurements of head waves. Thomas K. Berger and Michael J. Buckingham (Scripps Inst. of Oceanogr., Univ. of California—San Diego, 9500 Gilman Dr., La Jolla, CA 92093-0238, tberger@ucsd.edu)

As part of the SAX 99 experiments, head wave signals have been analyzed to determine geoacoustic properties for the sandy sediment at the shallow water site near Destin, FL. A small-aperture (2.33 m), vertical array of four hydrophones was deployed at mid-water depth over the sea floor, where the water column is 20 m deep. Transient signals were generated by an airgun, at a fixed location, resting on the ocean bottom. At a range of approximately 1 km from the source, the array was allowed to drift to vary its position; the location was measured with DGPS equipment. Two estimates of the compressional wave speed in the sandy sediment were obtained from the incoming signals: a spatially averaged value over the track between the source and receiver, determined at a single sensor from the difference in arrival times between the head wave and the water-borne wave; and a localized value, computed from the critical angle derived from the vertical directionality of the received head wave signals. A comparison of the independent local and spatially averaged estimates provides a useful consistency check, and establishes the spatial variation of the sediment more completely than by either purely averaged or local measurements alone.

TUESDAY AFTERNOON, 30 MAY 2000

GEORGIAN ROOM, 2:00 TO 3:45 P.M.

Session 1pAB

Animal Bioacoustics: Dolphins, Whales and Bats

Peter H. Rogers, Chair

School of Mechanical Engineering, Georgia Institute of Technology, Atlanta, Georgia 30332

Contributed Papers

2:00

1pAB1. Simple vocal identification method of porpoise sonar signals using an acoustic data logger and a time depth recorder. Tomonari Akamatsu (Natl. Res. Inst. of Fish. Eng., Ebikai, Hasaki, Kashima, Ibaraki 314-0421, Japan, akamatsu@nri.affrc.go.jp), Ding Wang, Kexiong Wang (Inst. of Hydrobiol., Wuhan 430072, PROC), and Yasuhiko Naito (Natl. Inst. of Polar Res., Tokyo 173-8515, Japan)

Individual identification is needed for the standard behavioral observation. However, the vocal identification of cetacean sound is difficult even using an acoustic data logger attached to the animal, because it is hard to exclude the sounds from a nearby conspecific. This paper shows that simultaneous measurement of echolocation signals and the dive depth of a free-ranging finless porpoise (*Neophocaena phocaenoides*) enables vocal identification. Echolocation signals of a juvenile finless porpoise were observed with an acoustic data logger attached to the upper posterior side of the pectoral fin by a suction cup. The acoustic data logger had 4.8-kHz sampling frequency and recorded each event of pulse, processed by an amplitude peak hold circuit. Dive depth and swim speed of the porpoise were recorded by another data logger every 1 s. Surface reflection of echolocation signals was clearly detected during shallow dives less than 2 m in depth. The delay time of the reflected signal corresponded with the two-way transmission time from the animal dive depth to the surface. This shows that the signals were from the animal wearing the data loggers.

2:15

1pAB2. Object matching across vision and echolocation by dolphins and humans. Heidi Harley (New College of the Univ. of South Florida and Epcot's Living Seas, Div. of Social Sci., Sarasota, FL 34243, harley@sar.usf.edu)

Bottlenose dolphins can recognize objects across modalities, i.e., given a sample object visually or echoically, a dolphin can identify that object from among a group of alternatives in the second modality (echoic or visual, respectively). Dolphins may be able to perform this task, because (1) they can learn to associate their visual and echoic experiences of an object; (2) they can recognize an equivalence between their visual and echoic experiences through the detection of object features that are available to both modalities; (3) they can create a spatial-analog representation (an "echo-image") of an object based on its echoes. Data from a dolphin performing a three-alternative cross-modal matching task in which identical (object A to object A) or conditional (object B to object C) matches were reinforced suggest that the dolphin can perform this task based on a recognition of equivalence between his visual and echoic experiences; the dolphin consistently matched based on object identity regardless of reinforcement for conditional matches. Data from a cross-modal matching task with humans listening to slowed-down dolphin echoes and choosing from among photographs of objects suggest that it is possible to perform this task without echo imaging.

ments. Velocity of vertically polarized interface waves is sensitive to sediment density and shear rigidity. Theoretical dispersion relations agree well with available experimental data and are shown to lead to a simple and robust inversion of interface waves travel times for shear-velocity profiles in the sediment.

12:10

1aAO8. Shear wave velocity measurements of layered sediment models using radiation impedance. Masao Kimura (Dept. of Ocean Eng., Tokai Univ., 3-20-1 Orido, Shimizu, Shizuoka 424-8610, Japan)

Shear wave velocity is important in characterizing surface marine sediments. It was demonstrated that the shear wave velocity in viscoelastic media can be obtained using the value of the frequency at which the radiation reactance becomes zero, and the shear wave velocity of sands can be measured. Moreover, the effect of the size of the radius of the circular vibrating plate on the measured values of shear wave velocities in sands is shown. In this study, the shear wave velocities in layered sediment model are measured. The sediment model is a two-layered model, which is composed of clay as the upper medium and sand or urethane rubber as the lower medium. The effect of the thickness of the upper medium on the measured values of shear wave velocities is also considered.

12:25

1aAO9. Geoacoustic properties of sediments determined from measurements of head waves. Thomas K. Berger and Michael J. Buckingham (Scripps Inst. of Oceanogr., Univ. of California—San Diego, 9500 Gilman Dr., La Jolla, CA 92093-0238, tberger@ucsd.edu)

As part of the SAX 99 experiments, head wave signals have been analyzed to determine geoacoustic properties for the sandy sediment at the shallow water site near Destin, FL. A small-aperture (2.33 m), vertical array of four hydrophones was deployed at mid-water depth over the sea floor, where the water column is 20 m deep. Transient signals were generated by an airgun, at a fixed location, resting on the ocean bottom. At a range of approximately 1 km from the source, the array was allowed to drift to vary its position; the location was measured with DGPS equipment. Two estimates of the compressional wave speed in the sandy sediment were obtained from the incoming signals: a spatially averaged value over the track between the source and receiver, determined at a single sensor from the difference in arrival times between the head wave and the water-borne wave; and a localized value, computed from the critical angle derived from the vertical directionality of the received head wave signals. A comparison of the independent local and spatially averaged estimates provides a useful consistency check, and establishes the spatial variation of the sediment more completely than by either purely averaged or local measurements alone.

1p TUE. PM

TUESDAY AFTERNOON, 30 MAY 2000

GEORGIAN ROOM, 2:00 TO 3:45 P.M.

Session 1pAB

Animal Bioacoustics: Dolphins, Whales and Bats

Peter H. Rogers, Chair

School of Mechanical Engineering, Georgia Institute of Technology, Atlanta, Georgia 30332

Contributed Papers

2:00

1pAB1. Simple vocal identification method of porpoise sonar signals using an acoustic data logger and a time depth recorder. Tomonari Akamatsu (Natl. Res. Inst. of Fish. Eng., Ebikai, Hasaki, Kashima, Ibaraki 314-0421, Japan, akamatsu@nri.affrc.go.jp), Ding Wang, Kexiong Wang (Inst. of Hydrobiol., Wuhan 430072, PROC), and Yasuhiko Naito (Natl. Inst. of Polar Res., Tokyo 173-8515, Japan)

Individual identification is needed for the standard behavioral observation. However, the vocal identification of cetacean sound is difficult even using an acoustic data logger attached to the animal, because it is hard to exclude the sounds from a nearby conspecific. This paper shows that simultaneous measurement of echolocation signals and the dive depth of a free-ranging finless porpoise (*Neophocaena phocaenoides*) enables vocal identification. Echolocation signals of a juvenile finless porpoise were observed with an acoustic data logger attached to the upper posterior side of the pectoral fin by a suction cup. The acoustic data logger had 4.8-kHz sampling frequency and recorded each event of pulse, processed by an amplitude peak hold circuit. Dive depth and swim speed of the porpoise were recorded by another data logger every 1 s. Surface reflection of echolocation signals was clearly detected during shallow dives less than 2 m in depth. The delay time of the reflected signal corresponded with the two-way transmission time from the animal dive depth to the surface. This shows that the signals were from the animal wearing the data loggers.

2:15

1pAB2. Object matching across vision and echolocation by dolphins and humans. Heidi Harley (New College of the Univ. of South Florida and Epcot's Living Seas, Div. of Social Sci., Sarasota, FL 34243, harley@sar.usf.edu)

Bottlenose dolphins can recognize objects across modalities, i.e., given a sample object visually or echoically, a dolphin can identify that object from among a group of alternatives in the second modality (echoic or visual, respectively). Dolphins may be able to perform this task, because (1) they can learn to associate their visual and echoic experiences of an object; (2) they can recognize an equivalence between their visual and echoic experiences through the detection of object features that are available to both modalities; (3) they can create a spatial-analog representation (an "echo-image") of an object based on its echoes. Data from a dolphin performing a three-alternative cross-modal matching task in which identical (object A to object A) or conditional (object B to object C) matches were reinforced suggest that the dolphin can perform this task based on a recognition of equivalence between his visual and echoic experiences; the dolphin consistently matched based on object identity regardless of reinforcement for conditional matches. Data from a cross-modal matching task with humans listening to slowed-down dolphin echoes and choosing from among photographs of objects suggest that it is possible to perform this task without echo imaging.

1pAB3. Introduced song replaces existing song in east Australian humpback whale population. Michael J. Noad (Dept. of Veterinary Anatomy and Pathol., Univ. of Sydney, NSW 2006, Australia, mnoad@mail.usyd.edu.au), Douglas H. Cato (Defence Sci. and Technol. Organisation, Pyrmont, NSW, Australia), and M. M. Bryden (Univ. of Sydney, NSW 2006, Australia)

Observations of humpback whale song in various parts of the world have shown that the song is a complex sequence of sounds that at any time is stereotyped within a geographical stock but differs between stocks. The song changes with time at a variable rate, usually in a gradual and evolutionary fashion. This paper reports observations of song change that differ from anything previously observed. A completely new song appeared spontaneously in a small number of singers in the east Australian stock, increasing in prevalence to completely replace the original song within two years. The two songs coexisted in the population for this period but individual singers generally produced one song or the other, with few occurrences of an intermediate variety. The new song was found to be the same as that observed off Western Australia. Previous comparisons have shown the east and west coast songs to be unrelated despite a small amount of interchange between the two stocks. The rapid rise to dominance of an introduced song indicates that song change is not necessarily evolutionary, gradual, or internally driven.

1pAB4. Playback experiments with long-finned pilot whales using a new broadband transmitter. Dennis F. Jones (Defence Res. Establishment Atlantic, P.O. Box 1012, Dartmouth, NS B2Y 3Z7, Canada) and Luke E. Rendell (Dalhousie Univ., Halifax, NS B3H 4J1, Canada)

Recently, a portable broadband acoustic transmission system (BATS) was developed as a joint venture between the Defence Research Establishment Atlantic and Sensor Technology Limited in Collingwood, Ontario. This transmitter is battery operated, easily deployed from small research vessels, and can produce source levels of 165 dB/1 μ Pa-m throughout its 1–14 kHz operating frequency band. The BATS was used by researchers from the Whitehead Laboratory at Dalhousie University to investigate vocal mimicry in long-finned pilot whales (*Globicephala melas*) off the coast of Cape Breton Island, Nova Scotia during the summer of 1999. A series of frequency-modulated tones in the 2–10 kHz band and using 33 cycles was played back to the pilot whales from the BATS and the acoustic responses were recorded using a hydrophone and tape recorder. Evidence that the whales heard these stimuli was extracted from the recordings and estimates of received levels were made. Although there were no instances of clear mimicry of the stimuli, there were a few tantalizing whale vocalizations that have highlighted the need for further experimentation. [Work supported in part by NSERC, Canada.]

1pAB5. Response of gray whales to low-frequency sounds. John R. Buck (ECE Dept. and CMaST, UMass Dartmouth, 285 Old Westport Rd., North Dartmouth, MA 02747, jbuck@umassd.edu) and Peter L. Tyack (Woods Hole Oceanogr. Inst., Woods Hole, MA 02543)

One transducer from the U.S. Navy SURTASS-LFA source was used during January 1998 to expose migrating gray whales off the California coast to low-frequency sound. These transmissions were 42 s in duration, repeating every 6 min, and 160–330 Hz. The whales were observed from shore stations north and south of the source location. Each pod of whales' closest point of approach (CPA) to the sound source was estimated from the shore observations. The received level (RL) at the CPA was estimated

using measured transmission loss curves and the known source level. The probability of observing whale pods at each RL is computed under both playback and control conditions. The difference between these probabilities as a function of RL provides a measure of avoidance for the animals. Previous work [Malme *et al.* (1984)] found a 50% change in these probabilities at a RL of 120 dB *re*: 1 μ Pa. For an inshore source location, the current work found 50% avoidance occurred at a RL of 141 dB, with 95% confidence bounds of ± 3 dB derived by bootstrap statistical methods. For the offshore source location, 50% avoidance was not achieved for any observed RL. [Work supported by ONR.]

1pAB6. Reconstruction of flight and acoustic behavior of echolocating bats. Kyler M. Eastman, James A. Simmons (Dept. of Neurosci., Brown Univ., Providence, RI 02192), and Michael J. O'Farrell (O'Farrell Biological Consulting, Las Vegas, NV)

A method for localizing sounds was applied to simultaneously reconstruct the emission pattern and flight path of echolocating bats. An infrared camera and an array of three microphones were used to record simultaneous audio signals and video images of groups of bats while foraging. The bats studied (*Eptesicus*, *Myotis*, *Pipistrellus*) use frequency sweeping "chirps" in the 25–45 kHz to 100–120 kHz band for echolocation. Cross correlation of the bats' chirps was used to find delays in chirp arrival time between microphones in the array. These delays were then used to locate the chirps' direction. A time-series of these coordinates was found, then calibrated to pixel points of the corresponding infrared video recording to reconstruct the flight path and chirp pattern for each of the bats over 15-s intervals. By providing simultaneous emission and position information throughout time, this method addresses questions about dynamic adaptation of the emission patterns of bats to the acoustic environment. Simultaneous adaptation between bats flying in groups is particularly interesting. An accompanying abstract discusses the implications of these findings for the current paradigm of information processing in the FM bat's auditory pathway.

1pAB7. Deployment of biosonar in diverse tasks by FM echolocating bats. James A. Simmons, Kyler M. Eastman (Dept. of Neurosci., Brown Univ., Providence, RI 02192), and Michael J. O'Farrell (O'Farrell Biological Consulting, Las Vegas, NV)

Using an infrared body-heat-imaging video camera and a microphone array, we made field recordings of FM echolocating bats (*Eptesicus*, *Myotis*, *Pipistrellus*) while they deployed their sonar alone or in company with other bats. We observed cruising flights of bats in transit from roosts to feeding areas, bats flying low over water for drinking or hunting, and aerial interception maneuvers in open areas (larger insects were visible, too). However, we also recorded frequent captures of insects on or near the ground, flights into vegetation to capture beetles, and groups of bats flying in close proximity while drinking or making aerial captures of insects. Bat-to-bat dogfights were common, some involving two or more bats diving and violently maneuvering close to the ground or near vegetation. For single bats or several bats not very close to each other, we recorded typical interception maneuvers with pursuit sequences consisting of search-, approach- (tracking-), and terminal-stage signals. By reconstruction of flight and emission patterns, we found that bats flying in close proximity alternated their broadcasts, and large aggregations of bats feeding close together perpetually emitted sounds characteristic of the mid-approach stage, with short durations, high repetition rates, and no discernible terminal stages.

Session 1pAO

Acoustical Oceanography and Underwater Acoustics: Acoustical Surveys of Ocean Bottom Geology II

Luiz A. L. Souza, Chair

MIT/WHOI Joint Program in Oceanography and Oceanographic Engineering, Woods Hole Oceanographic Institution,
Woods Hole, Massachusetts 02543

Chair's Introduction—2:00

Invited Paper

2:05

1pAO1. Acoustical techniques for observing the processes that control the formation of ocean bottom geology. Peter Traykovski (Woods Hole Oceanogr. Inst., Dept. of Appl. Ocean Phys. and Eng., MS#11, Woods Hole, MA 02543, ptraykovski@whoi.edu)

While acoustical methods provide useful techniques for surveying geological properties of the seafloor, acoustical techniques are also useful for examining the processes which create the observable geological properties. The dynamic processes which control sediment deposition and erosion are complex, and understanding these processes requires knowledge of the interactions of the seafloor topography with the hydrodynamics just above the seafloor and the subsequent sediment erosion, transport, and deposition. Observing these interactions requires multi-dimensional systems with high resolution in both space and time. This mini-tutorial talk will review some of the acoustic techniques for observing sediment transport processes which are currently being used. These techniques include acoustic backscatter measurements of suspended sediment concentration, acoustic imaging techniques such as rotary sidescan sonars and rotary pencil beam sonars, and the latest advances in high-resolution Doppler current profiling. Examples of data from these techniques will also be presented to give an overview of the transport processes including sediment transport due to the combined action of both waves and mean currents, sediment transport due to gravitational flows, and the relation between bedform migration and bedload transport.

Contributed Papers

3:05

1pAO2. Analysis of sound waves in a sandy sediment by focused array. Nicholas P. Chotiros, D. Eric Smith, James N. Piper, Brett K. McCurley (Appl. Res. Labs., Univ. of Texas, P.O. Box 8029, Austin, TX 78713-8029), David C. Gage, and Harvey Ma (Univ. of Texas, Austin, TX 78713-8029)

The objective was to determine the underlying physical processes in the penetration of sound into sandy ocean sediments, particularly at shallow grazing angles. Signals were collected from a buried acoustic receiving array in a sandy sediment off the west coast of Florida, as part of the Sediment Acoustics Experiment (SAX99). The array was insonified by a wideband sound source carried on a remotely operated vehicle (ROV). The discrimination between refraction and scattering processes was of particular interest. To this end, an array-processing method was used that employed focusing to distinguish between refracted signals from a distant source and scattered signals from scatterers in the near field of the array. Sound waves entering the sediment at steep angles were clearly refracted. In the case of shallow grazing angles, the results were much more complicated. [Work supported by ONR, Ocean Acoustics.]

3:20

1pAO3. Geoacoustic inversion results from the modal mapping experiment. Kyle M. Becker (MIT/WHOI Joint Prog. in Oceanogr. and Oceanogr. Eng., Woods Hole Oceanogr. Inst., Woods Hole, MA 02543), George V. Frisk (Woods Hole Oceanogr. Inst., Woods Hole, MA 02543), and Subramaniam D. Rajan (Sci. Solutions, Inc., Hollis, NH 03049)

The modal mapping experiment (MOMAX) was conducted in March 1997 off the New Jersey coast in about 70 m of water [G. V. Frisk *et al.*, J. Acoust. Soc. Am. **103**, 3028(A) (1998)]. A cw source emitting several discrete tones in the frequency range 50–175 Hz was used to excite the acoustic pressure field, which was measured using an array of freely drift-

ing hydrophone buoys equipped with GPS navigation and radio telemetry. The wave-number content of the spatially varying pressure field was extracted using a Hankel-transform-based method. The discrete portion of the resulting spectrum was used as input data to a perturbative inversion algorithm for inferring the geoacoustic properties of the seabed. Results are shown as depth-dependent sound-velocity profiles in the sediment. Comparisons are made among profiles obtained in different locations, highlighting the effect of variations in wave-number content on the inferred geoacoustic properties. [Work supported by ONR.]

3:35

1pAO4. Practical approaches for utilizing the plane-wave reflection coefficient as input data for inferring geoacoustic properties of the seabed. Luiz L. Souza, Kyle M. Becker (MIT/WHOI Joint Prog. in Oceanogr. and Oceanogr. Eng., Woods Hole Oceanogr. Inst., Woods Hole, MA 02543), and George V. Frisk (Woods Hole Oceanogr. Inst., Woods Hole, MA 02543)

Exact inverse techniques for inferring seabed geoacoustic properties typically require measurements of the plane-wave reflection coefficient as input data. Specifically, assuming the same constant density throughout the water column and the seabed, the reflection coefficient as a function of incident angle at a fixed frequency can be used to obtain a unique estimate of the sound-velocity profile in the bottom. In this paper, practical issues associated with implementing techniques such as the Gelfand–Levitan method in the deep and shallow ocean are discussed. These include the estimation of the reflection coefficient from point-source measurements using the Hankel transform, and the incorporation of a known density discontinuity at the water–bottom interface into the inversion scheme. Examples based on realistic seabed environments are presented. [Work supported by ONR.]

3:50

1pAO5. The predictability of broadband sound propagation in shallow water and geoaoustic inversion. David P. Knobles, Lewis A. Thompson, Robert A. Koch, and Karl C. Focke (Appl. Res. Labs., Univ. of Texas, Austin, TX 78713-8029)

An experiment was performed in a shallow-water region in the Gulf of Mexico to collect data for testing the performance of geoaoustic inversion algorithms using acoustic propagation measurements. Previous geophysical and acoustic measurements, along with inversions for seabed parameters in the general area, provide an excellent standard for validating several acoustic data inversion techniques. Continuous wave acoustic signals at multiple frequencies from a towed source, broadband impulsive source signals, and broadband signals generated by a moving surface ship were recorded on an advanced system that included a horizontal line array. Simulated annealing is used in conjunction with a normal-mode description of the propagation and various types of broadband matched-field cost functions to invert for a geoaoustic representation of the seabed. Consistent inversion results are obtained using all three types of acoustic data and are in general agreement with the reported physical measurements. The nature of the soft seabed results in a unique acoustic signature in the pressure time series from impulsive sources and the interference patterns in the transmission loss as a function of source-receiver range. The results indicate that the use of surface ships of opportunity can be successfully employed for geoaoustic inversion.

4:05–4:25 Break

4:25

1pAO6. Inversion for geometric and geoaoustic parameters of Haro Strait data using a *posteriori* density functions. John Viechnicki and Ross Chapman (Ocean Acoust. Group, School of Earth and Ocean Sci., Univ. of Victoria, Victoria, BC V8W 3P6, Canada, jayvee@ducks.seos.uvic.ca)

Inversions for geometric and geoaoustic parameters of a multitone, multishot experiment are examined using a *posteriori* probability distributions based on simulated annealing importance sampling. Data were collected as part of the matched-field tomography component of the Haro Strait PRIMER Experiment of June/July 1996 which can be categorized as low frequency, shallow water, and bathymetrically complex. The inherently unquantifiable sources of mismatch between data and model in environmentally complex systems favor a statistical approach over more conventional global optimization techniques. Large uniqueness and parameterization error complicate an accurate interpretation of the data covariance matrix necessary for appropriate choice of Gibbs sampling temperature. Trends in density functions are therefore examined as the relative sampling temperature is lowered. The mean and half-width value of the density distribution for each parameter are evaluated as a function of sampling temperature. [Work supported by ONR.]

4:40

1pAO7. Inversion of sediment parameters in a shallow-water range-dependent environment. Gopu R. Potty and James H. Miller (Dept. of Ocean Eng., Univ. of Rhode Island, Narragansett, RI 02882)

The propagation of acoustic energy in a range dependent environment will be analyzed. The broadband acoustic signals from explosive sources collected during the Shelf Break PRIMER Experiment will be used for this study. Time-frequency behavior of these signals will be analyzed

using wavelets. Upslope propagation along varying bathymetry (from 800 m at source location to 90 m at the receiver) and sound speed is considered. A simple model for the sediment compressional speeds will be extracted using a hybrid inversion scheme. This hybrid scheme uses a genetic algorithm and Levenberg–Marquardt nonlinear least-squares method. Inversion is based on group speed dispersion. Theoretical group speeds will be calculated based on range-dependent propagation. [Work supported by ONR.]

4:55

1pAO8. Determining a geoaoustic model from shallow-water transmission loss data using parameter linkage and a hybrid inversion algorithm. Marshall V. Hall (Maritime Operations Div. Defence Sci. & Technol. Organisation, P.O. Box 44, Pyrmont, NSW 2009, Australia)

In order to characterize the propagation conditions along a shallow-water sound range at low frequencies, measurements have been made of both cw transmission loss versus distance, and travel times of airgun-generated head waves. The head wave data yield the sound speed and time intercept of a reflecting interface, and these results are used as known parameters when the cw data are inverted to obtain a complete geoaoustic model. The inversion algorithm was a hybrid of the simplex and simulated annealing methods, similar to versions developed recently at the University of Victoria, Canada. The geoaoustic model was assumed to consist of two uniform solid layers overlying a solid uniform basement. The sound speed of the upper layer was estimated from the measured seafloor grain size, in accordance with empirical data. To further reduce the number of parameters, regression equations were devised to relate the less critical parameters (density, shear speed, and shear absorption) to those that are usually found to be more important (sound speed, sound absorption, and layer thickness), although the basement shear speed was also found to be an important parameter. With only six unknown parameters, the inversion algorithm generally found a satisfactory geoaoustic model after only several hundred runs.

5:10

1pAO9. A genetic algorithm for sediment sound-speed profile estimation from source and receiver position and travel time data. Brett McCurley, Nicholas P. Chotiros, James Piper, and Eric Smith (Appl. Res. Labs., Univ. of Texas, P.O. Box 8029, Austin, TX 78713)

Data were taken from a source projecting sound (10 to 100 kHz) at low grazing angles to a buried array (at a 30-deg decline) of nine receivers. These data were collected in October 1999 as part of the Sediment Acoustics Experiment (SAX99) off the coast of Fort Walton Beach, Florida. From the source and receiver positions and accurate measurements of travel time, the sound-speed profile can be estimated from the water–sediment boundary to the deepest receiver. The relationship between sound speed and depth can be determined. Also, refraction layers within the sediment can be ascertained. This estimation can be accomplished with the use of a genetic algorithm that searches a given sound-speed range to estimate the speed profile at specific frequencies. The algorithm equates the known travel time to the time associated with the refracted sound to each receiver position (distance from source). [Work supported by ONR, Ocean Acoustics.]

5:25–6:00

Panel Discussion

Session 1pBB**Biomedical Ultrasound/Bioresponse to Vibration and Physical Acoustics: New Techniques in Biomedical Imaging**

Thomas L. Szabo, Cochair

Agilent Technologies, MS-095, 3000 Minuteman Road, Andover, Massachusetts 01810

Ibrahim M. Hallaj, Cochair

*Brigham and Womens Hospital, Department of Radiology/MRI, 54 School Street, Lexington, Massachusetts 02421***Invited Papers****1:00****1pBB1. Clinical applications of acoustical microscopy.** Joie P. Jones (Dept. of Radiological Sci., Univ. of California—Irvine, Irvine, CA 92697-5000)

Very high frequency ultrasonic imaging offers an important tool for biomedical research as well as for clinical practice. Using ultrasound systems operating between 20 and 200 MHz, morphological and functional images of the skin may be obtained. For example, ultrasound may be used to assess wound healing, determine the extent of skin tumors, measure skin thickness, and visualize skin vascularization. At higher frequencies (600 to 1000 MHz), ultrasonic imaging can be used for histological studies with far greater sensitivity than optical microscopy and with a spatial resolution equivalent to optical methods. For example, ultrasound may be used to evaluate biopsy specimens more quickly and more accurately than optical microscopy. Here we report our experiences with very high frequency ultrasound as a clinical tool and, using this technology, the development of new standards by which tissue changes and disease processes may be defined.

1:20**1pBB2. Harmonic imaging in tissue.** Helen Routh (ATL Ultrasound, P.O. Box 3003, Bothell, WA 98041-3003)

In the last few years it has become clear that harmonic imaging in tissue (or tissue harmonic imaging) will improve image quality in many “difficult to image” patients. This talk will focus on the practical and clinical aspects of this imaging method which makes use of the nonlinear propagation of ultrasound in tissue. By filtering the received signal around the second harmonic, much of the haze and clutter in the image is reduced. Although there has been much research on the propagation of nonlinear waves in tissue, the exact mechanisms of this image quality improvement are still unclear. One hypothesis is based on the fact that the intensity of the second harmonic signal increases as the signal pass deeper into tissue. As the majority of artifacts arise from structures close to the transducer, the artifacts in the second harmonic signal are significantly lower than those in the fundamental. By imaging the second harmonic signal, much of the haze and clutter in the image can then be reduced. This talk will give a basic description of tissue harmonic imaging illustrated with clinical examples and a hypothesis for the improvements in image quality.

1:40**1pBB3. Dynamic sonoelastography—Making the impossible, possible.** Stephen F. Levinson (Departments of Physical Medicine and Rehabilitation, Univ. of Rochester, 601 Elmwood Ave., Box 664, Rochester, NY 14642), Dongshan Fu, Kevin J. Parker, and Sheryl M. Gracewski (Univ. of Rochester, Rochester, NY 14642)

Sonoelastography refers to the use of ultrasound to image the viscoelastic properties of soft tissues. Although a number of techniques have emerged, most are qualitative and static in nature. The availability of dynamic quantitative methods could open a new field of functional imaging for the study of biomechanics. We have developed an approach using 2D speckle tracking and finite element-based elastic modulus reconstruction. The speckle tracking method uses a deformable mesh with nodes assigned to points with a high feature content. The tissue is harmonically excited at a frequency of less than 100 Hz and vibration vectors are obtained at the nodes of the mesh. Young's modulus is then obtained within a series of small regions of interest using an iterative forward estimation approach. In each region, a finite element algorithm is used to predict the internal displacements from the measured motion on the boundary. A sum squared difference operator is used to estimate the associated error, from which an improved Young's modulus is obtained for use in the next iteration. Results from simulations and from measurements in soft tissue phantoms and in human muscle will be presented. Limitations of the technique will be discussed along with potential solutions.

2:00

1pBB4. Characterization of bone using longitudinal and shear wave imaging techniques. Junru Wu (Dept. of Phys., Univ. of Vermont, Burlington, VT 05405), Shigong Ye (Nanjing Univ., Nanjing, PROC), and John Peach (Univ. of Vermont, Burlington, VT 05405)

Recently a computer-controlled scanning ultrasound imaging system was developed in our laboratory. It includes a pair of broadband 48-MHz focusing copolymer transducers. The apertures of the transducers and their f -numbers were identically equal to 2 mm and 2.25, respectively. A specimen can be moved in a 10- μ m increment in its plane and its normal direction can be rotated along the ultrasound propagation direction. It can be used to produce transmission-mode shear-wave as well as longitudinal-wave images of solid specimens. Shear waves in solids were generated by mode conversion. The results of the longitudinal and shear-wave ultrasound images for a piece of compact bovine bone obtained using this system are presented. Shear-wave images combined with longitudinal images can provide a complete evaluation of elastic moduli of bone.

2:15

1pBB5. Fundamental frequency signal cancellation for tissue harmonic imaging. Fatma N. Ucar (GE Yokogawa Medical Systems, Ltd., Hino-shi, Tokyo 191, Japan) and Yoshiki Yamakoshi (Gunma Univ., Gunma 376, Japan)

An alternative method for tissue harmonic imaging is proposed. An adaptive filter (AF) is applied for each receiving RF signal acquired by the ultrasonic (US) transducer array. The technique uses single transmit pulse but cancels out the fundamental frequency (FF) of the received signal by appropriately changing the delay terms of the AF's. The resultant signal is slightly improved at second harmonic frequency. The M -tap-delay AF processing can be illustrated for N -element transducer array as below:

$$\text{Min}\{S_k[n-D] + S_{k+1}[n], S_k[n-D-\Delta D] + S_{k+1}[n], \dots, S_k[n-D$$

$$-M^*\Delta D] + S_{k+1}[n]\} = 3DS_k^1[n] \text{ for } 1 \leq k < N, (D+M) \leq n$$

$$< (\text{Sample No} - D - M),$$

where $S_k[n]$ is the received US echo signal from the k th transducer element at n th sample and $S_k^1[n]$ is the output of AF of the same index. The delay terms D and ΔD apply π -phase shift on first input of AF to minimize the sum. The autoadaptive filtering is applied for the further suppression of the signal at FF. The experiments to test the performance of the technique were carried out in a water tank for ten different random scatterer objects. A GE-LQ 500, 64-element S220-type probe was operated at a central transmitting frequency of 2.0 MHz. The echo signals were acquired by using a membrane hydrophone and sampled at 200 MHz. The average total improvement factor was found as 24.5 dB with standard deviation of 0.8.

2:30

1pBB6. Microbubbles and medical ultrasound: Optimization of harmonic imaging. Matthew Wyczalkowski and Andrew J. Szeri (Dept. of Mech. Eng., Univ. of California, Berkeley, CA 94720-1740)

Due to their nonlinear nature, bubbles driven at a particular frequency will radiate a considerable portion of their total acoustic emission at a frequency twice that of the driving—at the second harmonic. Tissues in the body, however, radiate predominantly at the driving frequency only. In

ultrasound applications where bubbles serve as contrast agents, this effect leads to reduced noise at the second harmonic and forms the basis of harmonic imaging. This paper explores techniques by which the amount of energy radiated by a bubble at a frequency of twice the driving may be increased. A modulating signal of a frequency distinct from the fundamental is added to the forcing. The amplitude and phase of the modulating signal are chosen so as to be optimal; that is, so as to maximize a particular measure of the acoustic response. Results are presented for several different schemes of multimode forcing. The total acoustic emission, as well as emission at the first and second harmonics, is obtained for bubbles of a spectrum of sizes and a range of amplitudes of the driving. These results are compared with the response of bubbles at the single-frequency forcing.

2:45

1pBB7. In vitro study of mechanisms of transient response imaging. Suor Kim and E. Carr Everbach (Dept. of Eng., Swarthmore College, Swarthmore, PA 19081)

Transient response imaging (TRI), or intermittent imaging, has been shown to produce increased myocardial contrast when fluorocarbon-filled microbubble contrast agents are used to visualize the heart. By lowering the frame rate from around 40 Hz to below 1 Hz, it has been hypothesized that less microbubble destruction occurs via inertial cavitation [E. C. Everbach, Shouping Li, and Thomas R. Porter, J. Acoust. Soc. Am. **102**, 3154 (1997)]. To test this hypothesis *in vitro*, a flow-through system of diluted PESDA contrast agent in distilled water was insonified at 1 MHz using 10, 100, and 1000 microsecond pulses at PRFs of 0.5, 1.0, 10, and 30 Hz. Inertial cavitation was quantified by a narrow-band 20-MHz passive transducer placed confocally with the 1-MHz insonifying transducer. Results and relevance to *in vivo* experiments will be discussed.

3:00–3:15 Break

3:15

1pBB8. New and old approaches for nonlinear-contrast agents imaging. Alexander M. Sutin (Stevens Inst. of Technol., 711 Hudson St., Hoboken, NJ 07030)

Microbubble, nonlinear-contrast agents are used in nonlinear ultrasound imaging to increase the contrast between these agents and the surrounding tissue. Similar approaches were developed in Russia in the 1980s for bubble measurements at sea. Two methods can enhance nonlinear imaging. (1) Sum-frequency nonlinear imaging has several advantages over second harmonic imaging due to the absence of nonlinear distortion in electronic circuits and emitters. This method allows one to use significantly less ultrasound amplitude than in the method of second harmonic imaging. This method was used for bubble detection in sea sediments. (2) The nonlinear difference frequency Doppler effect is based on moving bubbles that are insonified by two acoustic beams of different frequencies directed from different directions. Difference-frequency scattering provides significant enhancement of the Doppler frequency shift. This method was tested in laboratory conditions. Nonlinear scattering by microbubbles could also lead to remote, direct blood-pressure measurements. This method could be based on measurements of microbubble resonance frequencies by using nonlinear scattering. The resonance frequency depends on the ambient pressure; nonlinear scattering provides high-resolution measurements of the resonance frequencies and avoids the influence of scattered signals reflected by different inhomogeneities in the human body. [Work was supported by Stevens.]

1pBB9. Radiation impedance of a finite circular piston vibrating normal to an isotropic viscoelastic half-space with application to medical diagnosis. Xiangling Zhang, Thomas J. Royston (Univ. of Illinois at Chicago, Chicago, IL 60607), Hussein A. Mansy, and Richard H. Sandler (Rush Medical College, Chicago, IL 60612)

In a recent study [J. Acoust. Soc. Am. **106**, 3678–3686 (1999)], a new analytical solution was developed and validated experimentally for the problem of surface wave generation on a linear viscoelastic half-space by a rigid circular disk located on the surface and oscillating normal to it. The results of that study suggested that, for the low audible frequency range, some previously reported values of shear viscosity for soft biological tissues may be inaccurate. Those values were determined by matching radiation impedance measurements with theoretical calculations reported previously [J. Acoust. Soc. Am. **23**, 707–714 (1951)]. In the current study, the new derivation is extended to the calculation of radiation impedance. Comparisons are made with prior theory and experiments to determine which theory is more accurate and what range of values for shear viscosity is more accurate. Measurement of skin surface radiation impedance has been studied by a few researchers for rapid, noninvasive diagnosis of a variety of specific medical ailments. It is hoped that the developments reported here will advance these techniques and also provide insight into related diagnostic methods, such as sonoelastic imaging and other methodologies that utilize disease-related variations in soft tissue viscoelastic properties. [Work supported by a grant from the Whitaker Foundation.]

3:45

1pBB10. Bessel beams of finite amplitude in absorbing media. Kevin B. Cunningham and Mark F. Hamilton (Dept. of Mech. Eng., Univ. of Texas, Austin, TX 78712-1063)

The profile of a Bessel beam is $J_0(\rho)$, where J_0 is the zeroth-order Bessel function and ρ is a dimensionless distance from the axis. The beam propagates without diffraction, a property that has stimulated interest in connection with medical ultrasound imaging. Previous analyses of second-harmonic generation in Bessel beams are limited to lossless media. Du, Zhang, and Zhu [Proc. 14th Intl. Symp. Nonlin. Acoust., edited by R. J. Wei (Nanjing U.P., Nanjing, 1996), pp. 189–194] showed that the beam profile of the second harmonic in the near field is given approximately by $J_0^2(\rho)$. Ding and Lu [Appl. Phys. Lett. **68**, 608 (1996)] obtained $J_0(2\rho)$ for the far field. We provide a more general analysis of second-harmonic generation, first by investigating solutions for lossless propagation in greater detail, and second by including absorption. It is shown that the far-field beam profile is $J_0(2\rho)$ only for very small values of a characteristic absorption parameter. As the absorption parameter increases, the beam profile evolves toward a distribution given approximately (i.e., away from minima) by $J_0^2(\rho)$. Numerical results are presented for higher harmonics and for waveform distortion in a Bessel beam that forms a shock. [Work supported by ONR.]

4:00

1pBB11. Adaptive transmit focusing by numerical backpropagation of ultrasonic waveforms. James C. Lacefield and Robert C. Waag (Dept. of Elec. and Comput. Eng., Univ. of Rochester, Rochester, NY 14627, lacefiel@ece.rochester.edu)

Physical aspects of adaptive transmit-focusing algorithms that model an inhomogeneous medium as a time-shift screen distant from the aperture are described. Exact restoration of an ideal focus by numerical backpropagation of a time-shifted field to the aperture followed by physical propagation of transmitted waveforms requires an aperture that is sampled at

greater than twice the highest spatial frequency in the field and occupies an area larger than the nonzero region of the field. Time-shift screens that are estimated from backpropagated pulse–echo data include oscillations at outer sample positions. The oscillations are caused by edge waves diffracted from the aperture boundary. Simulations performed using point-reflector data acquired through a distributed aberration phantom with a two-dimensional array system demonstrate that nonidealities from spatially windowed data increase the peripheral energy of the compensated focus compared to the ideal case at levels below –20 dB. Omission of the outer time-shift estimates and apodization of the time-shifted field reduce the significance of these nonidealities. For biomedical applications in which pulse distortions more severe than simple time shifts are observed, compensated foci produced using the backpropagation method nevertheless exhibit lower peripheral energy than foci produced after time-shift compensation in the aperture.

4:15

1pBB12. The application of a material parameter extraction algorithm to MRI-based displacement measurements. Anthony J. Romano, Joseph A. Bucaro (Naval Res. Lab., Washington, DC 20375-5350), and Richard L. Ehman (Mayo Clinic and Foundation)

A novel algorithm based on the variational formulation is presented for the determination of the material parameters $(\lambda + 2\mu)/\rho$ and μ/ρ throughout inhomogeneous media given volumetric displacement information. The performance of the inversion algorithm is subsequently investigated when applied to measured displacement data throughout an inhomogeneous test phantom. The vector displacement components throughout the test phantom subject to monochromatic shear excitation measured in time using MRI (Magnetic Resonance Imaging) were temporally Fourier transformed to extract the components of monochromatic excitation, and the data were delivered to the inversion algorithm. A sequence of studies is presented on the basis of subsequent wavenumber filtering, polarization selection, and variation in the size of the volume elements of resolution. The resulting performance is assessed and recommendations for future efforts are discussed. [Work supported by ONR and NIH.]

4:30

1pBB13. Increasing field of view of high frame rate ultrasonic imaging. Jian-yu Lu and Shiping He (Ultrasound Lab., Dept. of Bioeng., The Univ. of Toledo, Toledo, OH 43606, jilu@eng.utoledo.edu)

A high frame rate imaging method has been developed recently based on the theory of limited diffraction beams for medical diagnosis. This method constructs images in a direction normal to the transducer surface. In this report, the high frame rate method is extended so that it will not only construct images in the normal direction, but also in any oblique directions. This increases the image construction area because images from several oblique angles can be combined. A new equation has been derived for this extension. A computer simulation was performed. Results show that the field of view of images constructed with the extended method is increased from about 15.74 deg to about 96.57 deg with a linear array of a width (foot print) of about 18.288 mm after combining images from only seven angles (the depth of the images is assumed to be 66.12 mm). This work is significant because a large field of view with a small transducer footprint is desirable for high frame rate heart imaging. [Work supported in part by Grant No. HL60301 from NIH.]

Session 1pEA

Engineering Acoustics: Transducer Materials

R. Lowell Smith, Chair

1749 Linbrook Drive, San Diego, California 92111-7112

Contributed Papers

1:00

1pEA1. Developing a simple heuristic viscoelastic theory of piezoelectricity. R. Lowell Smith (1749 Linbrook Dr., San Diego, CA 92111)

In its simplest form, the linear theory of piezoelectricity takes no account of losses or dissipation phenomena in electroactive materials. This deficiency is often addressed by allowing the dielectric, mechanical, and piezoelectric coefficients to be represented as complex numbers. This paper discusses a related exercise in which the electroactive material is considered to be a three-parameter viscoelastic solid. The piezoelectric equations are redeveloped via the incorporation of a generalized version of Hooke's law representing a single relaxation process. Through examination of the responses of the system in creep, stress relaxation, and dynamic loading scenarios, the time and frequency dependences of the associated relaxation and dynamic compliances are inferred. The formalism is applied to a simple longitudinal vibrator geometry, and it is demonstrated that primitive viscoelastic behavior can be advantageously described using standard equivalent circuit analogies. The method can readily be generalized to accommodate more complete viscoelastic models.

1:15

1pEA2. Decoupling electrical and mechanical loss measurements in piezoelectric ceramic. R. Lowell Smith (1749 Linbrook Dr., San Diego, CA 92111)

Specimens of commercial piezoelectric ceramic material are commonly screened before use by observing their two-terminal impedance or admittance behavior as a function of frequency near the most prominent electromechanical structure resonance. Often a companion admittance measurement is made at 1 kHz, and interpreted in terms of the free capacitance and its associated dielectric loss tangent. Due to the strong electromechanical coupling of this class of materials, the low-frequency dissipation should be regarded as due to combined electrical and mechanical losses. This paper discusses an inductive tuning measurement strategy that permits substantially decoupled observations of the electrical and mechanical losses to be made. Motivation for the method is developed using standard equivalent circuit analogies, which also provide a useful means for interpreting two-terminal electrical measurements on the materials of interest. Such an approach is beneficial because linear piezoelectric materials such as lead zirconate titanate are not very lossy, and a low-frequency dissipation measurement typically does not yield a very precise loss tangent determination.

1:30

1pEA3. Comparison of minor to major loop properties in PMN-PT ceramics. Harold C. Robinson and Elizabeth A. McLaughlin (NAVSEA Undersea Warfare Ctr. Div. Newport, Newport, RI 02841, robinsonhc@npt.nuwc.navy.mil)

The electromechanical properties of lead magnesium niobate-lead titanate (PMN-PT) ceramics are of great interest to transducer designers. Typically, the strain versus electric field and polarization versus electric field curves used to obtain those properties are measured over a major

loop, i.e., one with a large dc bias and large ac signal. However, many devices are designed to operate at lower fields that fall entirely within the major loop, i.e., in a minor loop. Since the electromechanical properties of PMN-PT are strong functions of the bias and drive levels, it is crucial to understand how well properties derived from the major loop measurement correspond to the actual minor loop properties. In this study, electromechanical measurements on PMN-PT corresponding to minor loops at varying dc bias, ac drive, mechanical prestress, and temperature levels were made. The equivalent piezoelectric constant, dielectric constant, and dielectric loss factor for these minor loops will be compared to the corresponding major loop parameters. It will be shown that, in many cases, the parameters derived from the major loop agree quite well with the directly measured minor loops. An attempt to establish the lower limit for extracting minor loop properties from major loop data will be made.

1:45

1pEA4. An energy-based method for the experimental determination of electrostrictive coefficients on polyurethane films. Francois M. Guillot and Jacek Jarzynski (School of Mech. Eng., Georgia Inst. of Technol., Atlanta, GA 30332-0405, gt7565a@prism.gatech.edu)

A new method for measuring electrostrictive coefficients on thin plastic films is presented. This method is applied to polyurethane samples, whose large electrostriction makes them interesting new transduction materials. The experimental part of the technique involves a small sample to which both an ac voltage and a large dc bias voltage are applied in the thickness direction. The sample is encapsulated in a soft silicone rubber and its electric field-induced deformations are optically measured in three perpendicular directions by a laser Doppler vibrometer. These data are input into a Rayleigh-Ritz energy minimization program implemented symbolically in MATHCAD. The program outputs the second-order electromechanical coupling coefficients M3311, M3322, and M3333 which represent the variations of the dielectric constant of the material with respect to stress. These are true electrostrictive coefficients, in the sense that they are material properties of the polyurethanes, independent of the mechanical boundary conditions imposed to the sample and of the second-order effect resulting from Maxwell stress. A first-order version of this method applied to piezoelectric polymers was validated numerically and experimentally. Experimental values of the electrostrictive coefficients are presented here for several types of structurally different polyurethane samples.

2:00-2:15 Break

1pEA5. Application of Lamb waves to characterize adhesive bond properties. Robert Seifried, Laurence Jacobs (School of Civil and Environ. Eng., Georgia Tech, Atlanta, GA 30332-0355), and Jianmin Qu (Georgia Tech, Atlanta, GA 30332-0405)

This research develops an analytical model that quantifies the effect of an adhesive bond's properties, on the propagation of guided ultrasonic waves in a three-layered, adhesively bonded component. A separate experimental study has developed the relationship between a bonded specimen's dispersion spectrum (the modes in its frequency-wave number relationship), and the condition of the adhesive bond; these experiments show that bond degradation causes some of the specimen's higher modes to damp-out and vanish. The goal of the current research is to develop a theoretical understanding of this experimentally observed behavior; a fundamental understanding of the mechanics of this phenomenon is needed to develop a quantitative nondestructive evaluation procedure, capable of measuring the stiffness of a bonded component. A continuum mechanics based model for the steady-state propagation of guided waves in a three-layer component is developed. This model enables a parameter study—by varying the adhesive bond's material properties—to quantify the effect of the adhesive's properties on the (overall) bonded component's dispersion curves, thus characterizing the effect of aging. In addition, the transient behavior of the bonded component is established with a finite element model.

2:30

1pEA6. Application of nonlinear wave processes for the nondestructive characterization of adhesives. Thomas Meurer, Laurence Jacobs (Civil & Environ. Eng., Georgia Tech, Atlanta, GA 30332-0355), and Jianmin Qu (Georgia Tech, Atlanta, GA 30332-0405)

This study investigates the mechanics and mathematical fundamentals of the propagation of nonlinear ultrasonic waves in adhesives. The objective of this research is to develop an understanding of the underlying mechanics of nonlinear waves, in order to develop quantitative techniques capable of interrogating damaged adhesive bonds. This study develops a mathematical model that represents a nonlinear stress-strain relationship, hysteresis, and then applies it to one-dimensional wave propagation. The numerical procedure uses a finite-difference scheme to study the propagation of pulses in this nonlinear media. This research includes an experimental verification of these numerical results, as well as a physical interpretation.

1pEA7. A new analytic model of piezoelectric stack actuators including ferroelectric nonlinearities. Yann Pasco and Alain Berry (G.A.U.S., Mech. Eng. Dept., Université de Sherbrooke, Sherbrooke, PQ J1K 2R1, Canada, Yann.Pasco@gme.usherb.ca)

The proposed numerical modeling of piezoelectric phenomena combines a mathematical model of hysteresis and a macroscopic theory of ferroelectricity including the nonlinear effect of polarized ceramics. Most of the time, modeling is sought between applied electric field versus displacement. Such an analytic model does not take into account ferroelectric phenomena and as a consequence never gives information about ferroelectric characteristics of ceramics, but gives a simple model of the ceramics' behavior. Another advantage of the new model is that it can still be accurate even if the ceramic is blocked in force. In order to simplify equations, the quasistationarity of electric field is assumed and three-dimensional equations will be reduced to a one-dimensional stack actuator. Such a model could be used in active control of vibrations.

3:00

1pEA8. Vibration of thin and thick annular plates. Yan Xiang, David A. Brown, Lawrence Reinhart, and Boris Aronov (Acoust. Res. Lab., ECE Dept., Univ. of Massachusetts Dartmouth, 285 Old Westport Rd., North Dartmouth, MA 02747)

This work is on the determination of the resonant frequency and mode shapes of thin and thick annular plates. We summarize an extensive experimental and theoretical investigation on the vibration of annular plates driven symmetrically with a center post. Experimental data for the thin plates agree well with theory. We also present the energy analysis method to describe thick plates and obtain an equivalent equation for mode shape as predicted by Mindlin's theory. Several sets of polycarbonate plates with inner radius to outer radius ratio range from 0.1–0.5 and thickness to outer radius ratio range from 0.1–0.5 were fabricated and tested using a non-contact photonic sensor. The resonance frequency, mode shape, and surface stress distribution were investigated. These results are useful in the design of acoustical accelerometers and pressure gradient hydrophone designs. [This work was supported in part by the University of Massachusetts Dartmouth and ONR, Code 321.]

TUESDAY AFTERNOON, 30 MAY 2000

SPANISH ROOM, 2:00 TO 4:35 P.M.

Session 1pPA

Physical Acoustics and Signal Processing in Acoustics: Signal Processing for Low Diffraction Beams

D. Kent Lewis, Chair

Lawrence Livermore National Laboratory, Advanced Applications, 7000 East Avenue, Livermore, California 94551

Chair's Introduction—2:00

Invited Papers

2:05

1pPA1. Experimental results of localized wave experiments. D. Kent Lewis (Lawrence Livermore Natl. Lab., 7000 East Ave., L-372, Livermore, CA 94550, kent-lewis@llnl.gov)

The theory for several low diffraction beams has been upheld by experimental verification in several tests to date. Both the modified pulse spectrum and the superposed Gaussian beams have been verified by Livermore projects in the laboratory and in open water tests. Test methods, instrumentation, data analysis techniques, and results will be discussed in this presentation. [This work was performed under the auspices of the Department of Energy by Lawrence Livermore National Laboratory under Contract W-7405-ENG-48.]

1pPA2. Generation and propagation of acoustic Bessel bullets. Peter Stepanishen (Dept. of Ocean Eng., Univ. of Rhode Island, Narragansett, RI 02882-1197, stepanishen@oce.uri.edu)

Acoustic Bessel bullets are a new class of localized wave solutions to the linearized wave equation with controllable spectral properties. Ideally such waves maintain their spatially dependent shape and time history as they propagate in free space. A brief review of the history of acoustic Bessel bullets (BBs) will be first presented. Both Fourier transform and recently developed convolution and impulse response methods of investigating the space-time properties of acoustic BBs will be presented. The generation of acoustic BBs from an array of planar annular rings with appropriate space-time source excitations will be addressed via the use of the space-time Kirchhoff integral solution. Various properties of acoustic BBs generated from an infinite planar aperture will be presented and contrasted to the space-time characteristics of the acoustic BB's generated by a finite planar aperture—the latter exhibits the characteristics of the former over a limited space-time region. The supersonic nature of the on-axis acoustic BB fields will be simply explained using space-time characteristics of the source excitations. Numerical results will be presented to illustrate the tradeoffs between the radial extent and the near- to far-field transition of acoustic BBs, and several interesting characteristics of the far field will also be discussed.

1pPA3. Helical focus wave modes and helical nondiffracting beams. Pamela L. Overfelt (Phys. and Computational Sci. Div. Res. Dept., Naval Air Warfare Ctr., Weapons Div., China Lake, CA 93555, overfeltpl@navair.navy.mil)

A number of years ago Durnin and colleagues considered the possibility of nondiffracting beams, beams defined to be nonsingular free space exact solutions of the scalar wave equation having the same intensity distribution in every plane normal to the axis of propagation [J. Durnin, J. Opt. Soc. Am. A **4**, 651–655 (1987)]. Assuming harmonic time dependence and cylindrical coordinates, i.e., a Fourier decomposition of the time and space coordinates, the Bessel function mode solutions of the scalar Helmholtz equation were pointed out as a class of diffraction-free solutions. Similarly, using a nonorthogonal helical coordinate system, helical mode solutions of the scalar Helmholtz equation that are nondiffracting in the same sense as the Bessel beams can be determined [P. L. Overfelt, Phys. Rev. A **46**, 3516–3522 (1992)]. In this paper, when a bidirectional plane wave decomposition [I. M. Besieris *et al.*, J. Math. Phys. **30**, 1254–1269 (1989)] is assumed, this same helical coordinate system can be used to obtain helical focus wave mode (FWM) solutions to the scalar wave equation. The fundamental Gaussian FWM can be decomposed into two oppositely polarized helical FWMs. Many relationships between the helical and cylindrical FWMs and their beam solution counterparts will be shown.

1pPA4. Performance prediction of localized waves using a simulator. David Chambers (Lawrence Livermore Natl. Lab., P.O. Box 808 L-154, Livermore, CA 94551, chambers2@llnl.gov)

Numerical simulations of the performance of localized waves have proved useful in planning experiments and doing parameter studies. In this talk predictions of axial energy decay and far-field spot patterns for the modified pulse spectrum and superposed Gaussian pulses are investigated as a function of their input parameters. These predictions are confirmed by recent experimental data and comparisons with conventional tone bursts are shown. Comparisons are also made with analytical predictions for near-field transition distance and far-field radiation patterns.

1pPA5. Beams with large depth of field. Paola Gori, Gabriella Cincotti, and Massimo Pappalardo (Dept. of Electron., Univ. of Roma TRE, Via della Vasca Navale, 84-00146 Roma, Italy, p.gori@uniroma3.it)

Acoustic bullets, also known as limited diffraction beams, are solutions of the scalar wave equation with the property of maintaining their shape upon propagation, without spreading in space and in time. A necessary and sufficient condition for the generation of acoustic bullets is presented. In particular, it is proved that a necessary condition for acoustic bullets is that all the wave vectors characterizing the beam are disposed over a cone with a given vertex angle. Impulse responses for the case of axial symmetry, and for the situation in which axial symmetry is lacking, are then given. The results obtained for acoustic bullets are then contrasted with those of other types of beams with large depth of field recently presented in literature (an example is the so-called “sinc wave”). Finally, the topic of the realization of practical approximation of these kind of beams, which, in principle, requires infinite apertures, is addressed. Numerical simulations are presented in order to analyze the departures from the ideal case when using a finite-dimension aperture.

Session 2aAA

Architectural Acoustics: Electronic Enhancement and the Traditionally Unamplified Art Form

Scott D. Pfeiffer, Chair

Kirkegaard and Associates, 801 West Adams Street, Chicago, Illinois 60607

Chair's Introduction—10:00

Invited Papers

10:05

2aAA1. Recent experiences with electronic acoustical enhancement in concert halls, opera houses, and outdoor venues. David Griesinger (Lexicon, Inc., 3 Oak Park, Bedford, MA 01730) and Steve Barbar (LARES Assoc., Columbia, MD 21045)

This paper gives a brief summary of acoustical theory based on human perception. It then uses this theory to discuss the design criteria and performance data of electronic acoustic-enhancement systems installed in a number of venues. The installations include the Deutsches Staatsooper in Berlin, the Hummingbird Centre in Toronto, the Circle Theatre in Indianapolis, the Adelaide Festival Centre Theater in Adelaide, Australia, and Morbisch FestSpiele in Austria. Solutions to the problems of maintaining optimum clarity while providing optimum envelopment are given.

10:25

2aAA2. The philosophy of electronic room enhancement. Mark A. Poletti (Industrial Res. Ltd., P.O. Box 31-310, Lower Hutt, New Zealand)

The acoustic design of rooms involves the determination of surface materials to provide a desired reverberation time, loudness, and diffusion. These acoustic properties are similar for sources and receivers throughout the room, and are therefore global. They are derived from the later part of room impulse responses. Reflecting surfaces are typically included to direct early energy from the stage area onto the audience areas. They operate locally for sources in the stage area, and produce more localized effects at different receiving positions. The acoustic effects are determined from the early part of room impulse responses. This paper argues that the design of electronic room enhancement systems should follow a similar philosophy to that of passive rooms. Those properties that are global, such as the reverberation time, should be enhanced using a system which can maintain this global property. This is ensured by using non-in-line systems which use microphones spaced throughout the room. Local properties are most effectively controlled using in-line systems which use microphones close to the stage area. A review of non-in-line and in-line enhancement systems will be presented, and a hybrid system introduced which aims to provide both global and local control of room acoustics.

10:45

2aAA3. Appreciation of the SIAP system by the audience, performers, and theater management. Ben Kok and Wim Prinssen (SIAP, P.O. Box 720, NL-5400 AS Uden, The Netherlands, siap@pbri.nl)

Over the past decade, numerous acoustic enhancement systems based on SIAP technology have been installed. Although the design concept and implementation have been proven by the objective results, this itself does not guarantee artistic acceptance by audience and performers. Most SIAP installations are being used without much comment and often the performers and/or the audience are not aware of the electronic modifications to the natural acoustics of the auditorium. This makes it difficult to measure the artistic acceptance. In some situations, however, the installation of the SIAP system was a turning point in the appreciation, thus exploitation, of the theater. After a brief introduction on what sets the SIAP system apart from other enhancement systems, this paper will give some examples of projects where the installation of an SIAP system had a positive influence on the audience attendance and appreciation by the performers.

11:05

2aAA4. Electronic architecture: Issues of use and acceptance in the musical community. Paul H. Scarbrough, David W. Robb, and J. Christopher Jaffe (Jaffe Holden Scarbrough Acoust., Inc., 114A Washington St., Norwalk, CT 06854, pscarbrough@jhsacoustics.com)

Electronic architecture has become a mature technology in recent years, particularly through the wide application of digital signal processing to the problem of creating uniform, stable acoustical enhancement. Despite vast improvements in the impact and stability of such systems, a considerable segment of the musical community continues to express reservations about the use of this technology in the concert hall or opera house. It is clear that some musicians and music critics fail to understand how such systems function as

well as the capabilities and limitations of the technology. This paper will outline the degree to which current state-of-the-art systems do indeed succeed in producing measurable acoustical improvements and discuss the varying degrees of acceptance such systems are receiving in different parts of the world.

11:25–11:45

Panel Discussion

WEDNESDAY MORNING, 31 MAY 2000

GEORGIAN ROOM, 8:25 A.M. TO 12:05 P.M.

Session 2aAB

Animal Bioacoustics: Animal Hearing Sensitivity and Characteristics

Larry L. Pater, Chair

U.S. Army Construction Engineering Research Laboratory, 2902 Farber Drive, Champaign, Illinois 61821

Chair's Introduction—8:25

Invited Papers

8:30

2aAB1. Developing noise–dosage weighting functions for animals. Ann E. Bowles (Hubbs-Sea World Res. Inst., 2595 Ingraham St., San Diego, CA 92109, annb1@san.rr.com) and Larry Pater (U.S. Army Construction Eng. Res. Lab., Champaign, IL 61821)

Human A- and C-weighting functions have been used to calculate noise dosage estimators in studies of wild birds and mammals. It would be better to develop functions specifically designed for them, perhaps taking into account psychophysical differences taxon by taxon. Unfortunately, methods that yield human weighting functions cannot be generalized to most animals. First, human weighting functions are based on equal-loudness contours, but these cannot be measured directly in wild animals. Possible alternate weighting functions include (1) the auditory threshold function of the species/taxon, standardized to 0; (2) the auditory threshold function shifted by a scalar quantity to reflect varying sensitivity; and (3) equal-latency functions, which fit equal-loudness curves well in humans. Second, it is not clear that such weighting functions yield efficient predictors of effect. While human A-weighting is used to calculate noise dosage in studies of both nonauditory and auditory effects, a single weighting function need not yield the best dosage estimator for effects in other animals. In studies of impact on behavior, equal-response-threshold functions (frequency-specific thresholds for behaviors) may be more efficient. Such alternative measures must be compared empirically before dosage estimators can be standardized.

8:50

2aAB2. Estimating equal-loudness contours in animals. Micheal L. Dent, Elizabeth F. Brittan-Powell, Robert J. Dooling, Bernard Lohr (Dept. of Psych., Univ. of Maryland, College Park, MD 20742), and Larry L. Pater (US Army CERL, Champaign, IL 61821)

In humans, loudness does not increase with intensity in the same way across frequencies. In animals, it is somewhat more difficult to determine the relative loudness of suprathreshold sounds, so comparable equal-loudness contours have not been obtained. This kind of information is important, however, for understanding how changes in frequency affect loudness in natural environments and for specifying which environmental noises are potentially hazardous. Here, we review two types of data from a number of animal species: reaction time data from psychophysical experiments, and amplitude-intensity functions from evoked potential experiments, in an attempt to estimate equal-loudness contours. Both types of data, from a number of species, can provide estimates like equal-loudness contours obtained in humans. At low SPLs, equal-latency contours closely parallel threshold curves, while at high SPLs the contours become shallower. A similar trend is seen in generating constant amplitude evoked-response contours across frequency. These results should be useful in considering aspects of acoustic communication in birds such as sound transmission, the design of vocal signals, and the frequency-specific sound-attenuating properties of the environment. [Work supported by NIH DC00198 and NRSA DC00046, and SERDP.]

9:10

2aAB3. Detection threshold of echo signals in a noise-free environment by the Atlantic bottlenose dolphin. Whitlow W. L. Au, David Lemonds (Hawaii Inst. of Marine Biol., P.O. Box 1106, Kailua, HI 96734), Stephanie Vlachos (Sea Life Park, Waimanalo, HI 96795), and Paul E. Nachtigall (Hawaii Inst. of Marine Biol., Kailua, HI 96734)

The Atlantic bottlenose dolphin has a sonar detection range of 72 m for a 2.54-cm solid sphere and 113 m for a 7.63-cm water-filled sphere in a noise-limited environment. However, there are many natural environments in which the ambient noise is low and will not mask a dolphin auditory system. The dolphin's detection range for the same 2.54- or 7.62-cm sphere in a nonmasking environment should be considerably longer than 72 and 113 m, respectively. In order to perform a sonar experiment in an environment where the ambient noise is low, a pool must be extremely large, or an outdoor facility with low ambient noise conditions would need

to be available. Another way of approaching this problem is to perform a passive listening experiment in which the signal presented to a dolphin simulates an echo from a specific target. We performed a listening experiment using both pure tones and simulated broadband echo from a 7.62-cm water-filled sphere. The detection threshold for narrow-band tonal signals (audiogram) is compared with the detection threshold for a broadband echo signal in order to determine the relationship between an audiogram and sonar detection. The results obtained with the broadband echo also represent the sonar detection range of a dolphin in a nonmasking environment. [Work supported by ONR.]

9:30

2aAB4. Auditory sensitivity in woodpeckers. Bernard Lohr, Elizabeth F. Brittan-Powell, Robert J. Dooling (Dept. of Psych., Univ. of Maryland, College Park, MD 20742, blohr@psyc.umd.edu), and Larry Pater (US Army CERL, Champaign, IL 61821)

While there are numerous behavioral and physiological studies of hearing in small birds, there are no studies of hearing in woodpeckers. Woodpeckers are particularly interesting because they exhibit a variety of cranial adaptations that may influence auditory sensitivity. The present study examined auditory sensitivity in small woodpeckers using the cochlear microphonic (measured subdermally and in the cochlea) as a means of estimating hearing thresholds. Subjects were presented with both rectangular-pulse clicks and tone-burst stimuli. Clicks were 0.1 ms in duration and were presented in ascending order from 30–100 dB SPL. Tone bursts were 5 ms in duration with a 1-ms rise/fall, and were presented in ascending order from 30–100 dB SPL at frequencies ranging from 500–8000 Hz. Thresholds were computed from latency-intensity and amplitude-intensity functions at all frequencies. Results suggest that hearing abilities of woodpeckers are similar to those of other small birds. [Work supported by SERDP.]

9:50–10:05 Break

Contributed Papers

10:05

2aAB5. Development of auditory sensitivity in budgerigars. Elizabeth F. Brittan-Powell and Robert J. Dooling (Dept. of Psych., Univ. of Maryland at College Park, College Park, MD 20742)

We examined the course of vocal development in budgerigars from hatch to 4 weeks postfledging. During this period, calls undergo dramatic changes that culminate in a stereotyped FM patterned food-begging call by the time of fledging. A shortened version of this call becomes the bird's first adult vocalization. Auditory feedback influences the development of these vocalizations, but little is known about nestling hearing. The present study tracked the development of auditory sensitivity in nestling budgerigars using the auditory brain stem response (ABR). Adult ABR audiograms are similar in shape to behavioral audiograms but are elevated by 20–30 dB. Thresholds for week-old nestlings are above 90 dB pSPL for tones and clicks, but improve markedly approaching adult levels by fledging. ABR waveform shape also changes dramatically over development. Because budgerigars use auditory feedback to learn and modify their calls throughout life, knowing how the auditory system develops and what and when the animal hears provides insights into the role that hearing plays in the development of different types of vocalizations. [Work supported by Grants DC-00046 from NIDCD to E.F.B.-P. and DC-00198 from NIH to R.J.D.]

10:20

2aAB6. Discrimination of Shepard tones by a song bird. Jeffrey Cynx (Psych. Dept., Vassar College, Poughkeepsie, NY 12604, chaos@vassar.edu)

We know relatively little about how songbirds perceive harmonic structures. This research investigated whether songbirds, like humans, can perceive pitch classes when pitch height is experimentally controlled. Humans, when presented with tone intervals controlled for pitch height (Shepard tones), perceive the intervals as either rising or falling according to closeness on a pitch class or octave circle. A special interval with these stimuli is the half octave (or tritone), as it is equidistant in the rising and falling directions on the octave circle. Songbirds, trained to discriminate rising versus falling intervals, showed an ability to perceive the Shepard tones also as rising or falling.

10:35

2aAB7. Hearing and vocalizations in the orange-fronted conure (*Aratinga canicularis*), a small parrot. Timothy F. Wright, Robert J. Dooling (Dept. of Psych., Univ. of Maryland, College Park, MD 20742, tw98@umail.umd.edu), Kathryn A. Cortopassi, and Jack Bradbury (Cornell Univ., Ithaca, NY 14850)

The orange-fronted conure (*Aratinga canicularis*) is a small Neotropical parrot that is the subject of an ongoing study of vocal communication and social organization. Three captive-bred orange-fronted conures were trained to detect pure tones using operant conditioning with a food reward and tested using the method of constant stimuli. These birds were tested at seven frequencies ranging from 0.5 to 8 kHz. The greatest sensitivity was found between 1 and 4 kHz and lower sensitivity outside this region. The shape of the audiogram for this species was similar to those previously found for three other species of small parrots, but the absolute sensitivity for orange-fronted conures was generally lower than these other parrot species at most frequencies. The range of greatest sensitivity in the orange-fronted conures corresponded fairly well with the spectral characteristics of their contact calls. In a sample of 288 contact calls recorded from eight wild birds in Costa Rica, the mean peak frequency of the calls was 3371 Hz with a SD of 341 Hz, and the bulk of spectral energy fell between 2295 and 5231 Hz. [Work supported by NIH Grant No. R01 DC00198 (RJD), NRSA Fellowship Nos. DC00046 and MH12111 (TFW), and a NSF grant (JB and KAP).]

10:50

2aAB8. Masked hearing thresholds for a bottlenose dolphin (*Tursiops truncatus*) and white whale (*Delphinapterus leucas*) in the far field and hydrodynamic near field. James J. Finneran and Sam H. Ridgway (Marine Mammal Prog., Space and Naval Warfare Systems Ctr., 49620 Beluga Rd., San Diego, CA 92152-6266, finneran@spawar.navy.mil)

A behavioral response paradigm was used to measure masked underwater hearing thresholds in a bottlenose dolphin (*Tursiops truncatus*) and white whale (*Delphinapterus leucas*) within the far field and hydrodynamic near field of an underwater sound projector. Masked hearing thresholds at 100 and 300 Hz were measured at source distances of 1, 2, and 4

m; over these distances the specific acoustic impedance ratio varied from approximately -10 to 0 dB at 100 Hz and -3 to 0 dB at 300 Hz. Individual hearing test trials consisted of a 3.5-s noise burst alone (N) or a noise burst along with a 1-s tone (S+N). Subjects were trained to produce an audible whistle in response to S+N trials and to remain silent for N trials. Thresholds were estimated using an up-down staircase procedure. Preliminary results indicate no significant differences between thresholds measured at the three different source distances for either test subject. [Work supported by ONR.]

11:05

2aAB9. Masked underwater audiogram for a Pacific white-sided dolphin. Jeanette Thomas (Western Illinois Univ. Reg. Ctr., 3561 60th St., Moline, IL 61265, Jeanette.Thomas@ccmail.wiu.edu), Ken Ramirez (Western Illinois Univ. Reg. Ctr., Moline, IL 61265 and Shedd Aquarium, Chicago, IL 60605), Mersedeh Jalili, and K. Kristine Grimm (Western Illinois Univ. Reg. Ctr., Moline, IL 61265)

Little is known about abilities of odontocetes to hear in a noisy environment. Using a descending staircase method and go/no-go response, underwater hearing of a female, Pacific white-sided dolphin masked by two levels of white noise (30 dB apart) was measured. The same dolphin had an unmasked audiogram [Tremel *et al.* (1998)]. Frequencies were tested in octaves between 500 Hz and 64 kHz, with 15 to 30 trials/frequency. Masked thresholds at maximum noise were similar to those on a bottlenose dolphin [Moore and Au (1982)], a beluga [Johnson *et al.* (1989)], and a false killer whale [Thomas *et al.* (1990)]. At 8 kHz, masked thresholds were more sensitive than unmasked thresholds, perhaps because masking and pool noise were similar levels. The highest critical ratio (43) was at 500 Hz and the lowest critical ratio (14) at 8 kHz. Critical ratios at 8 kHz or higher were comparable to other species. However, critical ratios at 500 Hz (43) and at 1 kHz (40) were higher than in the beluga (20 and 18, respectively). [The authors are grateful for support by the marine mammal staff at the Shedd Aquarium and financial support by ONR.]

11:20

2aAB10. Mapping acoustic sensitivity about the dolphin's head: A look at the peripheral hearing system. Randall L. Brill, Patrick W. B. Moore, David A. Helweg (SPAWAR Systems Ctr., San Diego D351, 49620 Beluga Rd., Rm. 204, San Diego, CA 92152-6505, brill@spawar.navy.mil), and Lois A. Dankiewicz (SAIC Maritime Services, San Diego, CA 92110)

The hypothesis that echolocating dolphins best receive acoustic signals over the pan bones of the lower jaw is widely accepted. Studies in echolocation and hearing have assumed that those areas serve as the dolphin's peripheral hearing system. The research that established that model, however, does not exclude other potential sound reception sites and suggests that additional areas of the head may be acoustically sensitive and perhaps frequency dependent. Using jawphones, relative hearing thresholds for representative frequencies (10, 30, 60, and 90 kHz) were behaviorally measured at over 40 sites on a dolphin's head. Iso-sensitivity curves were constructed and projected onto the image of a dolphin's head based on these measurements. The results suggest sensitivity to high frequency along the lower jaw with greater sensitivity forward of the pan bone area, sensitivity to low frequency around the external auditory meatus, and an acoustic asymmetry with greater sensitivity favoring the right side of the

head. These results may correlate to underlying anatomical features and suggest a more complex peripheral hearing system than has been previously assumed.

11:35

2aAB11. Effects of underwater noise on auditory sensitivity of a cyprinid fish. Amy R. Scholik and Hong Y. Yan (School of Biological Sci., 101 TH Morgan Bldg., Univ. of Kentucky, Lexington, KY 40506, arschol1@pop.uky.edu)

The ability of a fish to interpret acoustic information in its environment is essential for survival. Thus it is important to understand how underwater noise, from anthropogenic sources such as boats, affects fish hearing. In this study, the fathead minnow, *Pimephales promelas*, was used to examine the following: (1) the immediate effects of white-noise exposure (300–4000 Hz, 142 dB *re*: 1 μ Pa) on auditory thresholds; (2) recovery time of hearing ability after exposure; and (3) the effects of noise exposure, from a small boat engine recorded in the field, on hearing thresholds. Audiograms were measured using the auditory brainstem response (ABR) protocol and compared to fathead minnows not exposed to noise. Immediately after exposure to white noise, five out of the eight frequencies tested showed a significantly higher threshold compared to the control fish. Recovery was found to depend on both duration of noise exposure and auditory frequency. Boat engine noise, peak frequency of 1292 Hz, also elevated thresholds in the fathead's most sensitive hearing range (1000–2500 Hz). These results support the hypothesis that auditory thresholds can be altered by noise in the environment and provide evidence to show that these effects can be long term (>14 days). [Work supported by the Kentucky Academy of Science.]

11:50

2aAB12. Pressure and pressure-gradient sensitivity of the tadpole medulla. Andrea M. Simmons (Brown Univ., Dept. of Psychology, Providence, RI 02912) and Sheryl Coombs (Loyola Univ. of Chicago, Chicago, IL 60626)

The dorsal medulla of anuran tadpoles undergoes significant anatomical reorganization over metamorphic development. To examine how functional properties alter in response to anatomical changes, we recorded single and multiple unit activity in the medulla of bullfrog tadpoles to a sinusoidal-oscillating dipole source that traversed the length of the animal's body. This dipole stimulus has been used previously to distinguish between pressure-sensitive auditory responses in the saccular nerve of the goldfish and pressure-gradient lateral line responses in the lateral line nerves of goldfish and other teleost fish. Recording sites in the developing dorsal lateral nucleus showed pressure-type responses to dipole stimulation near the animal's head and lung. Evidence for the role of the bronchial columella and the air-filled lungs acting as a sound pressure transducer in mediating these responses was equivocal. Recording sites in the lateral line nucleus and surrounding neuropil showed pressure-gradient-type responses to stimulation near identified lines of superficial neuromasts on the body or tail. These data show that dipole stimulation is effective in eliciting responses from the medulla, and can be used to elucidate changes in function as the lateral line system degenerates and an auditory system adapted to hearing airborne sounds develops.

Session 2aBB

Biomedical Ultrasound/Bioresponse to Vibration: Ultrasound for Disease Treatment and Diagnosis I

Christy K. Holland, Chair

Department of Radiology, University of Cincinnati, 234 Goodman Street, Cincinnati, Ohio 45219-2316

Invited Paper

8:30

2aBB1. Ultrasound mediated drug delivery—A new era using ultrasound contrast agents? Evan C. Unger (ImaRx Therapeutics, Inc., 1635 E. 18 St., Tucson, AZ 85719 and Dept. of Radiol., Univ. of Arizona, Tucson, AZ)

Ultrasound has been used therapeutically for decades. High energy ultrasound is used clinically in hyperthermia. Transdermal drug delivery is accomplished with ultrasound, perhaps more effectively at low frequencies because of more efficient cavitation. High intensity focused ultrasound (HIFU) is used to coagulate tissue for treatment of cancer and hemostasis. Recently, contrast agents have been developed for diagnostic ultrasound. These agents can also be used as exogenous cavitation nuclei for therapy. Compared to some of the other therapeutic ultrasound techniques, exogenous cavitation nuclei may enable use of lower power levels. Applications of ultrasound contrast agents as cavitation nuclei include sonothrombolysis and drug and gene delivery. Cavitation of contrast agents can be used to increase capillary permeability to deliver drugs. Co-administration of ultrasound contrast agents with transdermal ultrasound accomplishes sonothrombolysis. Acoustically active carriers (contrast agents) have been developed to carry drugs and genes; systemic administration produces local drug delivery within the field of insonation. Contrast agents may open a new era of therapeutic applications of ultrasound for drug delivery and sonothrombolysis.

Contributed Papers

9:00

2aBB2. Ultrasonic scattering from blood as a means of measuring hemolysis. Constantin C. Coussios and John E. Ffowcs Williams (Eng. Dept., Cambridge Univ., Trumpington St., Cambridge CB2 1PZ, UK)

The need for a method that would permit the on-line monitoring of hemolysis (the destruction of red blood cells) as blood flows through artificial circuits is becoming increasingly important, as new techniques in transplantation surgery and organ preservation aim at longer periods of extra-corporeal support. It has already been demonstrated experimentally that ultrasonic scattering from blood is a function of the hematocrit (red cell concentration), the cell size, and the degree of red cell aggregation. This paper investigates whether ultrasonic scattering might be used, not only to obtain on-line measurements of the hematocrit, but also to assess the degree of hemolysis and to detect malformed populations of cells. The study involves a reexamination of existing models of ultrasonic scattering from blood, in which red cells have been traditionally modeled as spheres (Rayleigh scattering), and their extension to consider the effect on scattering of disk-shaped particles.

9:15

2aBB3. Cyclic variation of the Doppler power from porcine blood under pulsatile flow in tubes of varying compliance. Dong-Guk Paeng and K. Kirk Shung (Grad. Prog. in Acoust. and Bioeng. Prog., Penn State Univ., 315 Hallowell Bldg., State College, PA 16802, paeng@psu.edu)

The cyclic variation of the Doppler power backscattered from porcine whole blood and red blood cell suspensions was investigated under pulsatile flow in a mock flow loop of varying compliance. Experiments were performed using five different hematocrit values at four different stroke rates. A significant cyclic variation of Doppler power was measured in all experiments but the pattern of the cyclic variation was different. At high stroke rates (40 beats/min or more), the cyclic variation of the Doppler power showed a small peak during acceleration of flow and reached a minimum at peak systole and another peak was observed during deceleration. This minimum power phenomenon during peak systole became more prominent as the stroke rate was increased up to 60 beats/min. At a low stroke rate (20 beats/min), the Doppler power increased during early sys-

tole and remained the same until late diastole. The cyclic variation using whole porcine blood and red cell suspension showed the same pattern and suggested that aggregation of RBC did not affect the pattern of the cyclic variation. In addition to these general features, the differences for different experimental scenarios (whole blood versus red cell suspensions, and rigid tube versus compliant tube) are discussed.

9:30

2aBB4. Effect of high-intensity focused ultrasound on platelet activation, aggregation, and adhesion. Sandra L. Poliachik, Pierre D. Mourad, Lawrence A. Crum (Appl. Phys. Lab., Univ. of Washington, 1013 NE 40th St., Seattle, WA 98105, poliachi@u.washington.edu), and Wayne L. Chandler (Univ. of Washington, Seattle, WA 98195)

High-intensity focused ultrasound (HIFU) has been examined as a noninvasive means for achieving acoustic hemostasis [Delon-Martin *et al.*, UMB **21**, 113–119 (1995); Hynynen *et al.*, UMB **22**, 193–201 (1996); Vaezy *et al.*, UMB **24**, 903–910 (1998)]. Our own efforts in acoustic hemostasis are directed toward using diagnostic ultrasound to locate a hemorrhage and HIFU to halt the bleeding. To enhance the imaging of blood, the use of ultrasound contrast agents (UCAs; gas-filled microbubbles that increase the echogenicity of fluids) has been proposed as a means to locate internal bleeding; however, the combination of UCAs and ultrasound has been found to cause bioeffects in whole blood [Miller *et al.*, UMB **23**, 625–633 (1997); Poliachik *et al.*, UMB **25**, 991–998 (1999)]. Our results have shown that HIFU can cause platelets in a platelet rich plasma (PRP) sample to activate, aggregate, and adhere to a collagen-coated surface. Furthermore, UCAs can increase the amount of cavitation induced by HIFU, and thus lead to an increase in platelet activity. Although HIFU exposure alone can induce platelet activity, the addition of UCAs increases the amount and the rate of cavitation (cavitation dose); therefore, cavitation is the likely mechanism of HIFU-induced platelet activity. [Research supported by DARPA.]

2aBB5. Quantitative measurement of cavitation activity during red blood cell lysis and sonoporation. Wen-Shiang Chen, Tyrone M. Porter, Pierre D. Mourad, Peter H. Chang, Maile Hadley, and Lawrence A. Crum (Appl. Phys. Lab., 1013 NE 40th St., Seattle, WA 98105, wschen@apl.washington.edu)

Ultrasound contrast agents have proven to be an effective agent for producing cavitation-associated bioeffects. A comparative analysis of the strength and duration of cavitation activity and the lifetime of nucleation sites during red blood cell (RBC) lysis and sonoporation was performed. The RBCs were treated with 1.1-MHz tone-burst ultrasound in the presence of contrast agents and analyzed for hemoglobin release and the uptake of FITC-dextran (MW = 4 kD). Pulse durations ranged from 100 to 10 000 ms and the total number of bursts for each pulse duration was selected to ensure a constant exposure. The extent and duration of cavitation activity during treatment was monitored using a passive cavitation detection system. The collected cavitation-activity data was analyzed quantitatively and correlated with the percent of hemolysis and the ratio of fluorescent cells to total cells treated.

10:00–10:15 Break

10:15

2aBB6. Ultrasonic release of insulin from implantable, bio-compatible polymers coated with self-assembling membranes. Pierre D. Mourad, Lawrence A. Crum (Appl. Phys. Lab., Univ. of Washington, Seattle, WA 98105, pierre@apl.washington.edu), Connie Kwok, and Buddy Ratner (Chemical Eng. Dept., Univ. of Washington, Seattle, WA 98195)

There has been some research in the late 1980s on the development and use of biologically compatible, drug-carrying polymers for the purpose of subcutaneous release of the drug. In those papers, ultrasound was used to release the drug. This intriguing method for temporally targeted drug release has been hampered by a large, non-ultrasonic drug release rate. In other words, even without the application of ultrasonic stimulation, an excessive amount of the drug leaches out of the drug-carrying implant. We have developed a bio-compatible, drug-carrying polymer that can release a drug of interest upon ultrasonic stimulation, but whose background leaching rate is negligible. The advance in the present study over previous work is the coating of the drug-carrying polymer with a self-assembling membrane (SAM). Before and after the application of ultrasound, the SAM acts as an effective barrier. During and shortly after the application of ultrasound, the SAM transiently disassembles—thereby releasing the drug—then reassembles, which cuts off the drug flux. We present *in vitro* results demonstrating this effect, with insulin and ciprofloxacin as candidate drugs. [This work was supported in part by DARPA and NSF.]

10:30

2aBB7. Enhanced release of drugs from a novel polymeric film coated with self-healing, ordered methylene chains induced by hydrodynamic shear. Tyrone M. Porter, Pierre Mourad, Lawrence A. Crum (Appl. Phys. Lab., 1013 NE 40th St., Seattle, WA 98105), Connie S. Kwok, and B. D. Ratner (Univ. of Washington, Seattle, WA 98105)

The effect of numerous stimuli upon the release of drugs and other molecules from polymeric substrates has been of great interest in the drug delivery community. Ultrasound has proven to be an effective stimulus for the controlled release of drugs from polymer films. However, the mechanism by which this controlled response occurs has yet to be fully understood. In this study, the ability of shear forces generated by microstreaming around a single ultrasonically stimulated bubble to reversibly increase the release of the drugs from a coated polymer film is demonstrated. The polymer film is loaded with the drug ciprofloxacin and then coated with methylene chains consisting of 12 hydrocarbon chains. The leaching rate of the drug thus depends upon the extent of surface coverage by the

methylene chains. A single oscillating bubble in a fluid medium has the capacity to drive streaming at the surface of the polymer film and disrupt the methylene chain coating. An increase of drug concentration in suspension as high as ten times the baseline and controls was achieved, which implies an increase in the leaching rate. After treatment, the leaching rate returns to baseline levels, suggesting the methylene chains reorganize upon the polymer surface.

10:45

2aBB8. Acoustic mechanisms as an enhancer for transdermal drug delivery. E. J. Park, K. I. Jung, and S. W. Yoon (Dept. of Phys., SungKyunKwan Univ., 440-746 Suwon, Republic of Korea)

An acoustic enhancer for transdermal drug delivery has been studied for several decades in the field of engineering. However, its delivery mechanisms are not well understood. In this study, simple experiments were proposed to simulate a human skin system with a rat back skin for transdermal insulin delivery. Unfocused 0.5-, 1-, and 3-MHz ultrasound transducers were used for acoustic enhancers, and a heating wire loop was used for a nonacoustic enhancer. Ultrasound worked as a very effective delivery enhancer. The acoustic enhancer even gave better enhancement of insulin transportation than the nonacoustic enhancer, such as the heating wire loop, while almost the same amount of heating from both enhancers was applied to the diffusion cell. These results show that the thermal effect in the acoustic enhancer works as a major mechanism for transdermal insulin delivery. However, the nonthermal effects of ultrasound was also observed to contribute to enhance the transdermal insulin delivery.

11:00

2aBB9. Focused ultrasound opens the blood-brain barrier *in vivo*. Pierre D. Mourad, Lawrence A. Crum (Appl. Phys. Lab., Univ. of Washington, Seattle, WA 98195, pierre@apl.washington.edu), Ali Mesiwala, H. R. Winn, and Daniel L. Silbergeld (Univ. of Washington, Seattle, WA 98195)

The blood-brain barrier (BBB) consists of specialized endothelial cells that line the capillaries within the brain whose overlapping regions are held firm by “tight junctions.” In the central nervous system (CNS) the BBB plays a protective role, restricting the movement of molecules into the brain. It is also the major impediment to the flux of therapeutic drugs into the CNS. High-intensity focused ultrasound (HIFU) has been shown to induce the displacement of ordinarily impermeable vital dyes from the blood stream into the brain without damage to the CNS in the proximity of the flux. However, there has been no elucidation of the biological response to ultrasound that allows this flux. We show using a rat model that, by opening the tight junctions or by causing the endothelial cells to retract, HIFU can selectively disrupt the BBB in a reversible, controlled manner without injury to the surrounding CNS. Moreover, this effect reverses within 96 h. We believe that this work will lay the foundation for the development of therapies designed to optimize and take advantage of this novel biological response to ultrasound.

11:15

2aBB10. Effects of static pressure on ultrasound-mediated transdermal delivery of tagging compounds into migratory fish. Jane A. Young (CDNSWC, Code 725, Bethesda, MD 20084 and COMB, Ste. 236 Columbus Ctr., 701 East Pratt St., Baltimore, MD 21202), Joseph A. Clark (CDNSWC, Bethesda, MD 20084 and COMB, Baltimore, MD 21202), Victor Frenkel, and Yonathan Zohar (Univ. of Maryland Biotechnology Inst., Baltimore, MD 21202)

Research is currently being conducted with the US Fish and Wildlife Service investigating the use of ultrasound to enhance the transdermal delivery of tagging compounds into migratory fish. There is a need for noninvasive, low-cost, and efficient methods of tagging large numbers of fish in order to follow their migratory patterns. Current methods require long periods of immersion that can create health problems and are limited

in the amount of uptake of the compound. Previously reported research [J. A. Clark *et al.*, J. Acoust. Soc. Am. **101**, 3138(A) (1997)] showed cavitation levels of ultrasound (2–3 W/cm²) at 71 kHz significantly enhanced the transdermal delivery of the tagging compound calcein (fluorexon) into rainbow trout (*Salmo gairdneri*) yoke-sac larvae. However, attempts to repeat the experiment at higher altitudes revealed that static pressure criti-

cally affected the results. A further investigation into the effect of static pressure on the enhanced delivery of calcein by ultrasound into rainbow trout is reported here. A system for generating controlled doses of cavitation level ultrasound at various levels of ambient pressure will be described. The effect of static pressure on calcein uptake will be shown via fluorescent microscopy.

WEDNESDAY MORNING, 31 MAY 2000

CANADIAN ROOM, 8:00 TO 10:35 A.M.

Session 2aEAa

Engineering Acoustics and Signal Processing in Acoustics: Audio Engineering

P. K. Raju, Cochair

Department of Mechanical Engineering, Auburn University, Ross 201, Auburn, Alabama 36849-5341

David I. Havelock, Cochair

IMS/ASP, National Research Council, M36 Montreal Road, Ottawa, Ontario K1A 0R6, Canada

Chair's Introduction—8:00

Contributed Papers

8:05

2aEAa1. Internal noise in miniature electret microphones. Stephen C. Thompson (Knowles Electronics, LLC, 1151 Maplewood Dr., Itasca, IL 60143)

Modern miniature electret microphones normally include an internal amplifier to buffer the very high electrical impedance of the electret motor. This amplifier, usually an FET source follower circuit, generates a part of the noise in the microphone output signal. The electrical noise originates in the semiconductor and in the resistive components of the electrical circuit. Recent developments have reduced the electrical noise considerably so that it is no longer the dominant source. The other noise sources are the resistive components of the mechanical and acoustical impedances within the microphone. This paper describes recent work and special measurement methods that have helped to identify and reduce the acoustical noise components.

8:20

2aEAa2. Primary microphone calibration has become an easy task. Erling Frederiksen (Danish Primary Lab. of Acoust. and Customized Product Unit, Brüel & Kjaer, 2850 Naerum, Denmark)

New automated calibration equipment and software have made the quite complex pressure reciprocity calibration of microphones an easy and very accurate method. Today, the method can be applied not only by the specialists of National Calibration Laboratories but also by technicians of other laboratories. Since its invention in 1940, the method has become highly refined and is now used for essentially all primary calibrations. Many physical, electrical, and mechanical parameters of relatively small importance are now taken into account by the present international standard IEC61094-2 and by the software of the new equipment. Therefore, the system works with a high accuracy, better than ± 0.03 dB in the best range. As all measurements and calculations are automated, this is obtained together with a short calibration time and a simple operation. Another improvement is due the applied plane wave couplers, which can be used up to high frequencies (10 kHz with 1-in. and 25 kHz with $\frac{1}{2}$ -in. microphones) without the hydrogen filling required by the large-volume

couplers, which were commonly used in the past. This eases the calibrations significantly and eliminates safety problems. System features and calibration results will be discussed.

8:35

2aEAa3. Acoustic modeling of miniature loudspeakers. Guillermo Herrera and Scott D. Sommerfeldt (Dept. of Phys., Brigham Young Univ., Provo, UT 84602, guillermo_herrera@byu.edu)

There is currently considerable interest in developing models that more accurately predict the response of systems with small loudspeakers, such as in cellular telephones. There has also been a trend towards developing hands-free communication with cell phones, thus requiring a relatively high acoustic output from a small source. Work will be reported on the development of a model for a cell phone to help address such research issues. It has been shown that boundary layer effects can cause significant frequency shifts in the resonance peaks associated with systems having elements with dimensions on the order of the boundary layer thickness. This phenomena has been observed in the presence of ports with diameter of 1 mm and less. These effects are being investigated in a cell-phone system, and the results obtained to date will be presented.

8:50

2aEAa4. A real-time binaural loudness meter. Olli Tuomi and Nick Zacharov (Nokia Res. Ctr., Speech and Audio Systems Lab., Visiokatu, P.O. Box 100, FIN-33721 Tampere, Finland)

The alignment of sound intensity is of great importance, particularly in the field of psychometrics. When sounds have different spectral content, loudness alignment becomes a more suitable tool. However, traditional models assume the source either to reside in a free field at 0 azimuth or in a diffuse field. To assess or align sound-source loudness at different locations or with different radiation characteristics is more problematic. This paper presents the development of a real-time loudness meter based around the Moore loudness model. The tool has been developed in c and

MATLAB and supports both diffuse and free-field microphone input on a minimal hardware PC platform, with a generic sound card. Furthermore, binaural input to the model has been developed for a head and torso simulator (HATS). Use of a HATS allows for compensation of directional loudness properties caused by the head and pinna. It is thus possible to assess or align reproduction systems of different directivities or with different number of sources in either a monaural or binaural sense. Automatic calibration allows for easy alignment for calibrated use with either microphone or HATS. The system is presented and verification results given under different test conditions.

9:05

2aEAa5. Eyes 'n ears: A system of attentive teleconferencing. Bill Kapralos, Michael Jenkin, John Tsotsos (Dept. of Comput. Sci., York Univ., North York, ON M3J 1P3, Canada), and Evangelos Milios (Dalhousie Univ., Halifax, NS B3H 1W5, Canada)

Various teleconferencing systems exist, including systems intended for multiple speakers where a speaker must be localized. Although many sound-localization systems are available, not only are they expensive and nonportable, most require extensive audio arrays resulting in substantial computational processing. A simple, economical, and compact method of sound localization for use in a teleconferencing system is being investigated. In order to localize a sound source, this system relies solely on ITD measurements between microphone pairs, which allows the sound source to be determined as one of many possible locations on the surface of a cone. By taking the intersection of a sufficient number of cones, the location of a sound source in three-dimensional space may be determined. Our goal is to overcome the problems associated with the sound-localization systems currently available. In particular, this work seeks to develop an affordable, limited maintenance and compact sound-localization system for use with a teleconferencing system capable of locating and tracking a speaker in a multiple speaker setting.

9:20–9:35 Break

9:35

2aEAa6. Abstract withdrawn

9:50

2aEAa7. Speech/nonspeech discrimination in reverberant teleconferencing environments for acoustical source location. Christopher V. Alvino, Deborah M. Grove, James T. Kleban, and Shashank Sathyanarayana (Ctr. for Adv. Information Processing, Rutgers Univ., 96 Frelinghuysen Rd., Piscataway, NJ 08854, chrsvino@caip.rutgers.edu)

A real-time acoustical source location system implemented at the CAIP Center at Rutgers University presently tracks all sounds within a room. The system aims a video camera and audio sensors at the determined sound source for transmission of video and audio information to a remote teleconferencing location. It is desired to only direct attention to speech and to ignore nonspeech such as the shuffling of feet or the moving of papers. Speech detection is implemented to allow the system to aim sensors only at talkers. Present detection methods depend upon energy levels and the pitch present in voiced sounds to detect speech. In a reverberant environment the accuracy of speech detection degrades. Speech detection accuracy in reverberation is tested. Measures of false acceptance of nonspeech signals and false rejection of speech signals are used to determine which speech detection methods are optimal in different reverberant rooms. Speech and nonspeech in reverberant rooms are simulated using impulse responses generated by CATT-Acoustic. The test environments range from anechoic to heavily reverberant. A look-back buffer and adaptive tracking of energy and zero-crossing levels are used to allow speech detection to work optimally with the current source location system.

10:05

2aEAa8. Active source location and beamforming. Ramani Duraiswami, Dmitry Zotkin, Eugene A. Borovikov, and Larry S. Davis (Inst. for Adv. Computer Studies, Univ. of Maryland, College Park, MD 20742, ramani@umiacs.umd.edu)

A fundamental problem in designing user interfaces that incorporate speech and vision is locating the user(s) and enhancing the quality of their speech. We present a novel algorithm for actively doing this. Two complementary coarse-to-fine strategies that divide both the space being searched and the received signal in the frequency domain are used to arrive at an efficient dynamic search algorithm that is capable of both locating and beamforming multiple sources distributed in space. The algorithm makes no assumptions on noise statistics, and has a deterministic performance bound. It is able to make use of prior information on the possible location of sources obtained via means such as video. The algorithm is particularly applicable to locating speech sources in rooms of regular size, and is thus suitable for videoconferencing and user-interface applications. The algorithm and details are presented. [Support from the W. M. Keck Foundation and ONR Contract No. N000149510521 is gratefully acknowledged.]

10:20

2aEAa9. Dynamic sound localization using spatial awareness networks. Parham Aarabi (Gates 4B, Dept. of Comput. Sci., Stanford Univ., Stanford, CA 94305, parham@stanford.edu)

This paper proposes a sound localization algorithm that utilizes spatial awareness networks (SAN). SANs are an integrated network consisting of a probabilistic location inference model along with an adaptive observability network. This structure yields a synergic mechanism in which individual sensors are configured independently. This leads to a sound localization system that has no limitations on the number, orientation, or position of the microphones.

Session 2aEAb**Engineering Acoustics and Committee on Archives and History: History of Engineering Acoustics**

Kim C. Benjamin, Chair

*Naval Undersea Warfare Center, Newport, Rhode Island 02841***Chair's Introduction—11:00*****Invited Paper*****11:05****2aEAb1. History of engineering acoustics.** Stanley L. Ehrlich (Stan Ehrlich Assoc., P.O. Box 3679, Newport, RI 02840-0305)

At the Providence, RI meeting in December 1955 it was proposed to add three new technical committees to the original seven established about a year earlier. Two of the new ones, Audio Engineering and Electroacoustics, and Sonic and Ultrasonic Engineering, are precursors of the present Engineering Acoustics Technical Committee. In 1961, the name of the former committee was shortened to Electroacoustics, and I began to attend their meetings. In 1964, the scope of the committee was broadened, and its present name of Engineering Acoustics adopted. Early repeatable experiments in engineering acoustics have been attributed to Benjamin Franklin, and mostly anecdotal evidence exists before that. Modern developments involve interdisciplinary involvements with all acoustics, and specifically with the discovery of new materials, applications of their properties, and inventions of a wide variety of devices for producing, receiving, and using acoustics. Henry, Joule, Bell, Berliner, Edison, the Curie brothers, DeForest, and Fessenden will be among those noted, with some related recognition of them by the Engineering Acoustics TC.

A NOTE ABOUT THE ASA HISTORY LECTURE SERIES

In 1997, the ASA Committee on Archives and History conceived a plan for a series of invited lectures on each of the technical areas of the Society which would memorialize the significant achievements and milestones of each of its twelve technical committees and one interdisciplinary technical group during the first three quarters of the Society's first century.

With the cooperation of the technical committees, distinguished individuals are selected to review the history of their particular technical specialty and present a lecture which shows how that activity has developed and has contributed to the Society at large and to the broad field of acoustics as well. At the 138th meeting in Columbus, Ohio, the first two History Lectures were presented: Gabriel Weinrich on Musical Acoustics and Robert Beyer and David Blackstock on Physical Acoustics. At each subsequent meeting two additional lectures will be scheduled including those in Architectural Acoustics and Engineering Acoustics at this meeting in Atlanta.

The invited lecturers have been asked to prepare a written manuscript of their lectures which will be published in a commemorative book for the 75th Anniversary of the Society to be celebrated in 2004. The Archives and History Committee and the individual technical committees/group welcome comments and suggestions on both the History Lecture Series and on the proposed ASA Diamond Anniversary Book. Volunteers to assist the committees would be most welcome too. Contact Henry Bass, Chair, Committee on Archives and History, pabass@sunset.backbone.olemiss.edu

2a WED. AM**Session 2aMU****Musical Acoustics: Modeling and Perception of Musical Sound Sources I**

Antoine J. Chaigne, Chair

*Department TSI, ENST, CNRS URA 820, 46 rue Barrault, Paris, Cedex 75634, France****Invited Papers*****9:00****2aMU1. Modeling and control parameters of percussive sounds.** Antoine J. Chaigne (ENST, Dept. TSI, CNRS URA 820, 46 rue Barrault 75634 Paris Cedex 13, France)

The sounds of percussive instruments are essentially the result of free oscillations of a structure subjected to impact. For a given impact, their qualities depend on both the geometry and material of the structure. One efficient strategy for simulating such sound sources consists of modeling the propagation of elastic and acoustic waves in the various parts of the instruments. This yields a great variety of sounds governed by a relatively small number of parameters. The realism of the simulation is linked to the degree of

refinement in the modeling. Current challenging points in such an approach are due to the difficulties in measuring the physical parameters of the model, including the action of the player, and in the modeling of complicating factors, such as the various causes of inhomogeneities in the vibrating structure. In this paper, some aspects of the mathematical and physical connections between the geometry, the material properties, and the spectral parameters of simple and complex sound sources will be presented. The question of the appropriate “refinement” of the modeling will be addressed with regard to the practical use of the modeling: physics, instrument making, and experiments in psychoacoustics.

9:25

2aMU2. The psychomechanics of real and simulated sound sources. Stephen McAdams (IRCAM-CNRS, 1 place Igor Stravinsky, F-75004 Paris, France)

Recent advances in the physical acoustics of musical instruments and the development of physically based signal-processing models have opened a new path to the psychological investigation of sound-source perception. In particular, it is now possible to rigorously and systematically study the perceptual results of changes in physical properties that affect the mechanics of a musical instrument. As a first step in this direction, experiments from several laboratories have been conducted on the recognition and classification of sounds produced by real percussive instruments (bars and plates). These results show a surprising sensitivity to geometric properties. However, this sensitivity depends on the materials composing the sources and on the kinds of judgments listeners have to make. Other experiments have been performed on the perceptual dissimilarities among sounds produced by physical models of bars and plates whose material properties such as density, damping, and elasticity are varied. The results also show a very strong sensitivity of listeners to these latter mechanical dimensions, with the relevant perceptual dimensions often being transformations of the mechanical ones. The scope of the psychoacoustics of sound signals appears to be widening to include what might be called the “psychomechanics” of sound sources.

9:50

2aMU3. Modeling stiffness in virtual bowed-string instruments. Stefania Serafin and Julius O. Smith (CCRMA, Dept. of Music, Stanford Univ., Stanford, CA 94305)

Physical models of bowed-string instruments have achieved a degree of completeness that enables understanding and simulation of most of the phenomena that appear in real instruments. Moreover, improvements in hardware technology and the development of efficient signal processing algorithms enable the implementation of these models in real-time platforms, including also improvements and refinements which were not possible before. One of these elements is the stiffness of strings, whose main effect is to disperse the sharp corners that characterize the ideal Helmholtz motion. In this paper, a method is proposed for estimating numerical filters that model dispersion in strings. The phase dispersion of the string is modeled using allpass filters whose coefficients are obtained by minimizing the L -infinity norm of the error between the internal loop phase and its approximation by this filter cascade. This algorithm, which turns out to be a nonlinear version of the Remez exchange algorithm, can successfully design filters of low order which are suitable for real-time implementation, and also accurate enough to model strings of different materials.

10:15

2aMU4. The direct and inverse problems in 3D modeling of horns: Current results and future directions. Noam Amir (Holon Academic Inst. of Technol., 52 Golomb, Holon 58102, Israel, noamot@wine.cteh.ac.il) and Jonathan Kemp (Univ. of Edinburgh, Edinburgh, UK)

Modeling of acoustic horns is a subject that has been studied extensively for at least a century. It is only with the advent of the digital computer that it has become feasible to study the coupling between the various transversal modes in such horns. In this paper we examine the influence of the higher-order modes, obtained through a numerical solution. We show that though they may influence the acoustic field considerably, their influence on the input impedance is much smaller. We demonstrate that for many horns, taking only one higher-order mode into account improves accuracy to a great extent. We also show how the higher-order modes can influence parameters used in the plane wave model, such as radiation impedance and the end correction. Preliminary results in solution of the inverse problem, taking higher-order modes into account, are also discussed. Current algorithms incorporate plane waves only, and thus are inaccurate at areas of large flare. Finally, we discuss the more practical sources of inaccuracy in measurements: the relationship between spatial and temporal discretization, and transformation of measurements and simulated results between frequency and time domains. We summarize with some philosophical musings on the meaning of “accuracy” when dealing with musical instruments.

Contributed Papers

10:40

2aMU5. Inclusion of wave-front steepening in a frequency-domain model of trombone sound production. Michael W. Thompson and William J. Strong (Dept. of Phys. and Astron., Brigham Young Univ., Provo, UT 84602, mwt2@email.byu.edu)

This research investigates the modeling of harmonic generation occurring in trombone sound production at loud dynamic levels. The profile of

a tenor trombone was approximated as a set of contiguous cylindrical tubes. A linear model of the trombone was obtained by requiring continuity of pressure and flow between adjacent tubes. The mouthpiece and radiated pressure spectra were measured in an anechoic chamber for several pitches played in first position at several dynamic levels. The measured mouthpiece spectra were processed by the model, taking into account harmonic generation due to wave-front steepening. Results comparing the calculated radiated spectra with the measured radiated spectra will be reported, along with an overview of the algorithm.

2aMU6. Determination of the frequency vibrato extent based on the linear predictive coding parameters. Jose Diaz (Univ. of Carabobo, Valencia-Venezuela), Antonio Arroyo, and Howard Rothman (Univ. of Florida, Gainesville, FL)

Many time-domain signals can be modeled as a sum of sinusoids whose main parameters are amplitude and frequency. The frequency can be obtained from the linear predictive coding (LPC) parameters by the pole location in the Z-domain. Although intuition dictates that the amplitude could also be obtained from the LPC parameters, there appears to be no information available that would allow its computation. The purpose of

this research is to investigate the relationship between the LPC parameters and the amplitude of a wave in the time domain. In order to accomplish this, the power spectrum obtained using the LPC method was plotted against the FFT spectrum of vibrato samples. They were then compared, and several parameters of the LPC spectrum were changed until both spectra matched. From these matched spectra, the wave amplitude was obtained. This new method was then used in an application to determine the time-varying amplitude of the vibrato wave. This served two purposes: first, to verify the accuracy of the method, and second, to show the usefulness of the time-domain parameters. The results show that there is a direct relationship between the LPC parameters and the amplitude of a wave in the time domain.

WEDNESDAY MORNING, 31 MAY 2000

TOWER 1204, 9:00 TO 10:25 A.M.

Session 2aNS

Noise and Education in Acoustics: Educating the Public on the Perils of Overexposure to Noise

Daniel R. Raichel, Chair

The City College of the City University of New York, New York, New York 10031

Chair's Introduction—9:00

Invited Papers

9:05

2aNS1. Comparison of percentile and equivalent sound level criteria for construction projects. Gregory C. Tocci (Cavanaugh Tocci Assoc., Inc., 327F Boston Post Rd., Sudbury, MA 01776, gtocci@cavtocci.com)

Community ordinances and construction specifications widely use equivalent sound levels, and more often maximum sound pressure levels, to establish limits on environmental noise. In the case of construction, limits on noise expressed as maximum sound levels or equivalent sound levels can be either unreasonably permissive or unreasonably restrictive. It will be argued and shown that expressing construction noise limits as percentile sound levels offers a more realistic assessment of noise impact and can be more useful in understanding construction noise impacts. This presentation will discuss the advantages and disadvantages of limits expressed as equivalent sound levels, maximum sound levels, and percentile sound levels. Methods of analysis and results using each type of descriptor will be discussed. These methods will be compared using actual probability density functions and equivalent sound levels measured for construction operations. The presentation will close with the author's suggestions for managing construction noise control projects on the basis of his firm's 20 years of experience conducting construction noise and vibration analyses on a wide variety of construction projects including highway, transit, and building projects.

9:25

2aNS2. The importance of community involvement in a successful construction noise control program. Erich Thalheimer (Central Artery/Tunnel Project, 185 Kneeland St., Boston, MA 02111)

The Central Artery/Tunnel Project in Boston is generally recognized as the largest and most ambitious urban reconstruction project ever undertaken in the United States. The most politically charged of the required mitigation programs is to successfully control construction noise. Toward this end, the importance of the Project's active interaction and relationship-building with the affected communities cannot be overstated. The CA/T Project's construction noise control program attempts wherever possible to be proactive to avoid noise complaints. Part of the proactive strategy is to fully inform the affected neighborhoods through regularly scheduled community meetings. These meetings are organized by the Project's Community Liaison Group, with support from applicable technical staff. All interested parties are encouraged to attend. A statistical analysis has been performed for all noise-related complaints received by the Project for the time period of May 1994 through November 1999. The trends show that the "rate" of noise complaints has remained relatively steady even though construction activities have increased some fourfold over this time period. The good news is that through the lessons learned and precedents set by the CA/T Project, large-scale urban construction projects can be advanced and noise can be successfully mitigated to the extent necessary to satisfy the affected communities.

9:45

2aNS3. Noise in the news: What is and is not being covered by the media. Les D. Blomberg (Noise Pollution Clearinghouse, Box 1137, Montpelier, VT 05601)

More than 100 newspapers were monitored for 4 years for articles concerning noise. A database of more than 7000 newspaper articles concerning noise was created. The articles were summarized and categorized by noise source and area of interest. The data provided a unique overview of what the public has been reading about noise in print media. Coverage of noise was seen to follow hot-button issues. Aircraft noise, airport noise, and airport expansion dominated news coverage. Some very important noise issues, such as occupational noise, received almost no coverage. And, some emerging trends and issues such as environmental justice and noise-related violence were observed. Finally, conclusions regarding media coverage and the role of the media in educating the public about noise were discussed.

10:05

2aNS4. Protecting pedestrians from open construction activities. Daniel R. Raichel (The City College of the City Univ. of New York, New York, NY 10031) and Michael Dallal (Cooper Union Res. Foundation, New York, NY 10003)

During the course of a project that was conducted for New York City's Department of Design and Construction, it was observed that many urban construction activities occur in open areas, involving street repaving, overhaul of utility lines, sidewalk repairs, etc. Equipment such as backhoes, jackhammers, air compressors, and pavement saws generate extremely loud noises, often exceeding 100 dB(A) at 5 or 6 m. Passersby, exposed to the cacophony without the benefit of any intervening barriers, are as close as 0.4 m to the noise-generating equipment. An additional peril to passing pedestrians exists in debris flying from the breakup of concrete, asphalt, or metal. It was also noted that many pedestrians, even those suffering the threshold of pain from the noise, will not go out of their way to a quieter section of the street. Hence, a need exists to establish and enforce regulations to protect pedestrians from such open construction activities. Noise barriers can do double duty in preventing the scattering of debris and mitigating noise. Portability and reusability constitute the requirements for these barriers. But, contractors need to be educated on the proper deployment of these barriers and on the necessity for them.

WEDNESDAY MORNING, 31 MAY 2000

SPANISH ROOM, 8:45 A.M. TO 12:00 P.M.

Session 2aPA

Physical Acoustics: Thermo-, Photo-, Magneto-Acoustics I

Robert A. Hiller, Chair

Department of Physics and Astronomy, University of Mississippi, University, Mississippi 38677

Invited Papers

8:45

2aPA1. Acoustic Faraday effect and transverse waves in a superfluid. William P. Halperin, Yoonseok Lee, Thomas M. Haard, and James A. Sauls (Phys. Dept., Northwestern Univ., Evanston, IL 60208, w-halperin@nwu.edu)

Acoustic waves provide a powerful tool for studying the structure of matter. The speed, attenuation, and dispersion of acoustic waves give useful details of the molecular forces and microscopic mechanisms for absorption and scattering of acoustic energy. In solids, both compressional and shear waves occur, so-called longitudinal and transverse sound. However, normal liquids do not support shear forces and consequently transverse waves do not propagate in liquids with one notable exception, the quantum liquid phase of ^3He . We have observed rotation of the polarization of transverse sound waves in superfluid $^3\text{He-B}$ in a magnetic field. This magneto-acoustic effect is the direct analog to the magneto-optical effect discovered by Michael Faraday in 1845, where the polarization of an electromagnetic wave is rotated by a magnetic field along the propagation direction. Our experiment [Nature **400**, 431 (1999)], is the first demonstration of the existence of propagating shear waves in a liquid.

9:15

2aPA2. Thin film photoacoustics: From the laboratory to the real world and back. John A. Rogers (Lucent Bell Labs., 600 Mountain Ave., Murray Hill, NJ 07974), Alexei A. Maznev (Philips Analytical, Natick, MA 01760), and Keith A. Nelson (MIT, Cambridge, MA 02139)

Characterization of thin film anisotropic elastic moduli, thermal diffusivities, thickness, delamination, cure kinetics, residual stress, and other properties through impulsive stimulated thermal scattering (ISTS), a noncontact, nondestructive photoacoustic method, was demonstrated in the 1990s. The method, which involves two crossed "excitation" laser pulses used for generation of acoustic waves, a probe laser beam, and a diffracted signal beam whose time-dependent intensity is measured to reveal acoustic oscillations and damping as well as thermal diffusion, originally required large-frame lasers and painstaking optical alignment. A series of refinements resulted in a compact ISTS system including only miniaturized diode and diode-pumped lasers, using only wallplug electrical power, with automated tuning of the acoustic wavelength and frequency, and with no user-adjusted optical elements. Commercially available

ISTS systems are now used routinely in the microelectronics industry for metal film thickness measurement, yielding angstrom precision with 1-s data acquisition times. The ISTS method will be reviewed. Several new advances in methodology and new applications and thin films and bulk materials will be discussed. Finally, the connection between ISTS and related time-resolved measurements of collective vibrations will be illustrated.

Contributed Papers

9:45

2aPA3. Generation of photoacoustic sound for gases in an open environment. Serdar H. Yónak and David R. Dowling (Dept. of Mech. Eng. and Appl. Mech., Univ. of Michigan, Ann Arbor, MI 48109-2121, drd@engin.umich.edu)

Photoacoustics is the generation of acoustic waves due to unsteady heating from a light source. It was discovered over 119 years ago and is commonly used for trace gas detection and molecular spectroscopy when the photoactive gas or gas mixture is contained in a chamber. However, photoacoustic phenomena are less well characterized in an open environment. In this presentation, scaling laws for radiated photoacoustic sound pressure are developed from dimensional analysis and verified with experiments using a carbon dioxide laser tuned to 10.6 μm and small plumes of sulfur hexafluoride, a photoactive tracer gas. In addition, a forced wave equation that includes laser-gas heating is developed to predict the generation and propagation of photoacoustic sound. Using this forced wave equation, the shape and amplitude of photoacoustic sound pulses are determined based on the shape of the tracer-gas plume and the extent of laser-beam absorption through the Lambert–Beer law. The model results are compared to measured photoacoustic signal pulses and show favorable agreement. [Work supported by Ford Motor Company.]

10:00

2aPA4. Thermoacoustic microengines. Ran Yaron (Adv. Refrigeration Technologies, 188 S. Cedar Brook, Boulder, CO 80304), Ray Radebaugh (NIST, Boulder, CO 80303), Ezra Elias (Technion—Israel Inst. of Technol., Haifa, Israel), and Lev A. Ostrovsky (CU, CIRES/NOAA ETL, Boulder, CO 80303)

The high rate of heat transfer (up to 500 MW/m²) which can be achieved by explosive bubbles has long been recognized. This process finds a widening range of application from microsurgery to computer peripherals design. The best-known example is the thermal ink-jet printer, where an ultrahigh heat flux applied for short duration of about 6 μs induces rapid bubble generation and growth, which drives an adequate mass of ink forward through the exit nozzle. Explosive boiling is a phenomenon in which boiling is developed in an explosive manner at large superheat. Superheat conditions can be realized by absorption of a strong laser beam or by quick heating of a surface. Reversible phase transitions near the fluid's critical temperature by "acoustic explosive bubble" may provide a new way to continuously transfer the high thermal power associated with explosive boiling in a variety of heat exchange processes. For example, explosive bubbles could form the basis for fluid micropumping devices to be used in microrefrigerators. The present paper concentrates mainly on estimating the necessary conditions for realizing explosive boiling in liquids and how this boiling can be used to create a stable acoustic oscillation involving growth and collapse of the bubble.

10:15–10:30 Break

10:30

2aPA5. Design and construction of a solar thermal powered thermoacoustically driven thermoacoustic refrigerator. Jay A. Adeff and Thomas J. Hoffer (Phys. Dept., Code PH, Naval Postgrad. School, Monterey, CA 93943)

A compact, small temperature span, thermoacoustically driven thermoacoustic refrigerator powered by solar thermal energy has been successfully built and tested. An 18-in-diam Fresnel lens was used to focus sunlight at 475 °C onto the hot end of a 1-in-diam reticulated vitreous carbon prime mover stack, thereby eliminating the need for a hot heat

exchanger. The high intensity sound waves produced by the prime mover were used to power a thermoacoustic refrigerator to produce 4 W of cooling power at a cold temperature of 5 °C and a temperature span of 18 °C.

10:45

2aPA6. Experimental study of a heat-driven/solar thermoacoustic refrigerator. David G. Holmberg, G. H. Chen, Andrew M. Wo (Nat. Taiwan Univ., Inst. of Appl. Mech., 1 Roosevelt Rd., Sec. 4, No. 1, Taipei 106, Taiwan, david@spring.iam.ntu.edu.tw), and H. T. Lin (Chung Cheng Inst. of Technol., Tao-Yuan, Ta Shi 33509, Taiwan)

This work reports the design and testing of a prototype thermoacoustic refrigerator optimized by the code DeltaE. The key functional feature of this quarter-wavelength resonator is its dual power input capability electric heating elements and the sun. These can be utilized independently or complementarily to benefit all-weather operation. The machine operates near 270 Hz with helium (60%)–argon (40%) mixture at two bars mean pressure. Heat from either heat source is transported from the high-temperature end to the prime mover stack via thin copper plates that also serve as the hot-end heat exchanger. Both stacks utilize an uncoated automobile catalytic converter ceramic structure with 1.3-mm cell-to-cell square openings and 74% porosity. The ceramic structure provides a strong and easy-to-implement stack with minimal heat conduction. Both middle- and cold-end heat exchangers use simple copper fin design, with the middle heat exchanger using circulating water cooling. DeltaE results predict a 30 °C temperature drop across the heat pump stack at no-load condition with 30 W heating, 363 °C across the prime mover stack (work = 4.0 W) and pressure fluctuation level equal to 6.2%. Results from the effects of working fluid, changing pressure, heating temperature, and cooling load will be discussed.

11:00

2aPA7. A thermoacoustic traveling wave linear amplifier. Robert A. Hiller (Dept. of Phys. and Astron., and Natl. Ctr. for Physical Acoust., Univ. of Mississippi, University, MS 38677, hiller@olemiss.edu)

This paper describes an experiment to show linear amplification of traveling sound waves in a duct using a thermoacoustic regenerator. As noted by Ceperley [J. Acoust. Soc. Am. **66**, 1508–1513 (1979)] a Stirling engine-type regenerator should act as an acoustic gain medium for traveling waves in a duct. This principle is used in thermoacoustic traveling wave engines to transfer power from heat reservoirs to acoustic energy. However, it is difficult to produce finite gain for pure traveling wave impedance, since viscous losses in the channels of the regenerator overcome the gain, and previous workers have only been able to show reduced loss in such a system. Optimizing the regenerator design with numerical modeling and using a greater temperature difference suggest that a traveling wave thermal amplifier can produce 2 dB of real gain over two octaves for traveling waves in air. Such a device would amplify a broadband acoustic signal without electrical transducers. Design of the amplifier and experimental results will be shown.

11:15

2aPA8. Traveling-wave heat pumping in stack-spaced pores. Ray Scott Wakeland and Robert M. Keolian (Grad. Prog. in Acoust., Penn State Univ., P.O. Box 30, State College, PA 16804, wakeland@psu.edu)

Both standing and traveling waves cause heat flow near surfaces. Thermoacoustic refrigerators are based on two well-studied types of acoustically induced heat flow: (1) heat pumping concentrated about a thermal penetration depth from a surface due to a standing wave, and (2) heat pumping much closer to the surface due to a traveling wave. This talk

examines another thermoacoustic pumping mechanism, induced by a traveling wave but occurring mostly a thermal penetration depth from the surface. This sort of heat pumping, therefore, is greatest for pore sizes typical of stacks rather than regenerators. Unlike "Stirling pumping" (traveling-wave heat pumping that occurs very close to a surface), which is independent of surface temperature gradient, this mechanism is proportional to temperature gradient. It pumps heat *down* the temperature gradient, and is thus a sort of "thermoacoustic short circuit," detrimental to all types of working thermoacoustic devices. It appears to set an upper limit on the pore size of the regenerator in a thermoacoustic Stirling-type refrigerator that depends on temperature gradient. Analytical and graphical results will be presented, along with a Lagrangian "moving-parcel picture" that explains the results on physical grounds. [Work supported by ONR and NSF.]

11:30

2aPA9. High-PACE: A high performance acoustics computing environment. Brian C. Tuttle, Victor W. Sparrow (Grad. Prog. in Acoust., Penn State Univ., 157 Hammond Bldg., University Park, PA 16802), and Brian H. Dennis (Penn State Univ., University Park, PA 16802)

In physical acoustics it is now possible to solve some large problems via direct simulation on parallel computers due to the emergence of relatively inexpensive Beowulf class clusters of commodity personal computers. In this brief talk, an overview of the performance characteristics and

initial calculation results from such a cluster will be provided. Called High-PACE, for high performance acoustics computing environment, the new machine has 32 Pentium III processors at 450 MHz, 8.5 GB of RAM, over 200 GB of disk storage, and is interconnected by a Myrinet (Myricom, Arcadia, CA) 1.28 gigabit/s interface. The machine uses the Linux operating system and is programmed with FORTRAN 90 and the message passing interface, MPI. Initial work is focused on a deep understanding of nonlinear acoustic streaming for thermoacoustics applications. [Work supported by ONR.]

11:45

2aPA10. The effect of time-averaged velocity on boundary layer species separation. Roger Waxler (Grad. Prog. in Acoust., ASB 217, Penn State Univ., State College, PA 16804, rwax@acs.psu.edu)

Recently, Swift and Spoor observed acoustically induced separation of the components of a gas mixture. To explain their observations, they developed a theory by estimating the second-order time-averaged mass fluxes of each component, assuming the second-order time-averaged velocity to be zero. Here, the effect of the second-order time-averaged velocity is estimated.

WEDNESDAY MORNING, 31 MAY 2000

STATE ROOM, 8:00 A.M. TO 12:10 P.M.

Session 2aPP

Psychological and Physiological Acoustics and Speech Communication: Characterization and Possible Neural Bases of Age-Related Hearing Loss

Robert D. Frisina, Chair

Otolaryngology Division, University of Rochester Medical Center, 601 Elmwood Avenue, Rochester, New York 14642-8629

Chair's Introduction—8:00

Invited Papers

8:05

2aPP1. Age-related differences in temporal processing and speech perception in noise. Karen B. Snell and D. Robert Frisina, Sr. (Int. Ctr. for Hearing and Speech Res., Natl. Tech. Inst. for the Deaf, Rochester Inst. of Technol., Rochester, NY 14623, kbsncp@rit.edu)

Studies of age-related changes in speech perception have shown that older adults experience more problems understanding everyday speech, and require more favorable signal-to-noise ratios when compared with younger adults. Similarly, other studies suggest older adults show poorer temporal processing, especially when the complexity of the signal, the background, or the task is increased. While many age-related differences can be predicted by hearing loss, our studies of older subjects with essentially normal audiometric thresholds suggest that some age-related differences in temporal processing and speech perception in noise persist even in the absence of hearing loss. In addition, it appears that declines in speech perception in noise may follow changes in temporal processing that begin decades earlier. Results from studies of middle-aged subjects will be discussed that suggest age-related changes in auditory processing occur over many decades. [Work supported by NIA-NIH, and the International Center for Hearing and Speech Research, Rochester, NY.]

8:35

2aPP2. Beyond presbycusis: Non-hearing loss-related temporal processing deficits in the elderly. Pierre L. Divenyi (Exp. Audiol. Res., V. A. Medical Ctr., Martinez, CA 94553, pdivenyi@marva4.ebire.org)

A group of 56 elderly individuals were tested in a battery of tests including pure tone audiometry, various word- and sentence-level speech understanding tasks, and measurement of spectral, temporal, and spatial auditory abilities. Presbycusis hearing loss, when present, was symmetrical in the two ears and at most mild to moderate, with no significant deficit below 1.5 kHz. Frequency

resolution was estimated by calculating auditory filter width from notched-noise masking at 1 kHz; temporal resolution was estimated from gap discrimination and from auditory segregation of temporal patterns of harmonic complexes with components below 1.2 kHz (in the low envelope frequency range); spatial resolution in the horizontal plane was estimated from interaural time difference (ITD) thresholds for 500-Hz clicks. The largest differences between results of the elderly and those of normal-hearing young subjects were seen for speech understanding in spatially distributed babble, word recognition in reverberant rooms, ITD thresholds, and segregation of temporal patterns. The results portray the decline of temporal processing as an important component of auditory and speech understanding dysfunctions in aging and strongly suggest that a non-negligible part of these dysfunctions may be of central origin. [Supported by the National Institute on Aging and the V. A. Medical Research.]

9:05

2aPP3. Time and timing: Age-related differences in auditory, speech, language, and cognitive processing. M. Kathleen Pichora-Fuller (Audiol. & Speech Sci. Dept., 5804 Fairview Ave., Univ. of British Columbia, Vancouver, BC V6T 1Z3, Canada, kathy_fuller@audiospeech.ubc.ca)

Older adults have difficulty understanding language spoken in noisy conditions or at fast rates, even though they may have little difficulty with slow speech in quiet. Reduced audibility in the spectral domain cannot fully account for their difficulties. Differences in temporal processing may provide new explanations. Temporal processing includes synchrony coding (phase-locking), gap and duration coding (onset and offset detection), and coding of rhythmic patterns (envelope patterns). Particular speech cues rely on particular types of auditory temporal processing: voice cues arising from vocal-fold vibration rely on synchrony coding; some segmental contrasts rely on gap and duration coding; suprasegmental cues associated with syllabic rhythms rely on coding of higher-level patterns. The relationship between psychoacoustic and speech-perception measures will be considered at each of these levels of temporal processing. Evidence of age-related differences has been demonstrated in studies employing jittered and speeded speech, and speech with reduced stop-consonant gaps. While researchers have argued that perceptual and cognitive declines both affect aging listeners, apparent age-related differences in cognitive performance seem to be secondary to perceptual differences. Both time and timing seem to be affected, with aging auditory systems being slower and more asynchronous.

9:35–9:50 Break

9:50

2aPP4. Functional changes in the ear with old age: A review. Richard A. Schmiedt and Hainan Lang (Dept. of Otolaryngol., Medical Univ. of South Carolina, P.O. Box 250150, Charleston, SC 29425, SCHMIERA@MUSC.EDU)

To study the true effects of age on hearing, one needs a model wherein environmental noise and genetic influences can be controlled over the lifetime of the animal. Popular models include various mice and rat genotypes, gerbils, and rabbits. All models show varying degrees of hair-cell loss, spiral ganglion cell (neural) loss, and atrophy of the *stria vascularis* with age. Animals raised in quiet show little loss of hair cells, and cochlear nonlinearities remain intact, except for those genotypes that are known to have progressive hair-cell loss, like the C57BL/6 mouse. On the other hand, all models show stria atrophy and neuronal loss with age. Functional consequences of stria atrophy include a decreased endocochlear potential and an altered ionic homeostasis within the cochlear duct. Neuronal loss may preferentially affect cochlear afferent neurons with low spontaneous rates and high thresholds. Indeed, the loss of the endocochlear potential may correlate with the inactivity of these neurons, even in young animals. Thus, stria atrophy may induce changes in the activity of specific neuronal populations, thereby changing not only hearing thresholds, but also how sound is encoded at suprathreshold levels. [Work supported by NIA.]

10:20

2aPP5. Physiological correlates of presbycusis: Single-neuron recordings in awake animals. Joseph P. Walton (Otolaryngol. Div., Univ. of Rochester School of Medicine & Dentistry, 601 Elmwood Ave., Rochester, NY 14642-8629, Joseph_Walton@urmc.rochester.edu)

During aging, the brain undergoes senescence alterations causing a decrease in many behavioral responses, including the perception of complex sounds in difficult listening situations. Two neural deficits are likely responsible. One is a loss in sensory cells in the inner ear, while the second is likely due to deficits in neural processing within the central auditory system. This presentation will address the question as to whether evidence exists to support the idea of age-related deficits in auditory processing in the brain. Initial studies were devoted to the inferior colliculus (IC), an important midbrain nucleus that integrates information from many lower brainstem centers. We used two experimental approaches to assess neural encoding of complex signals. First, temporal processing of static (gap detection) and dynamic (amplitude modulation) changes in sound intensity in single neurons from young and old animals was examined. Results demonstrate that encoding of sound temporal features in many aged neurons is degraded. Second, we applied neurotransmitter antagonists microiontophoretically to young IC neurons to degrade temporal coding ability. The results display a wide range of complexities, but in some neurons temporal encoding is markedly disrupted. [Work supported by NIA–NIH, and the International Center for Hearing & Speech Research.]

10:50

2aPP6. Progressive hearing loss in mouse models of presbycusis: Influence of an augmented acoustic environment and of gender. James F. Willott (Dept. of Psych., Northern Illinois Univ., DeKalb, IL 60115, TJ0JFW1@WPO.CSO.NIU.EDU)

Progressive hearing loss associated with aging can be modulated by a number of genetic and environmental variables. Understanding how such variables work may suggest new approaches to ameliorate presbycusis. Two examples are long-term exposure to an augmented acoustic environment (AAE) and factors associated with gender (e.g., during the period of reproductive decline in females after 6 months of age). Research on inbred strains of mice that exhibit progressive hearing loss (C57BL/6J, BALB/cJ,

DBA/2J) has demonstrated that daily exposure to a 70-dB SPL broadband noise AAE slows the progressive elevation of ABR thresholds, enhances behavioral responses to tones (startle response; prepulse inhibition), and affects histopathology of the cochlea and anteroventral cochlear nucleus (AVCN). However, gender plays a role in both the progression of hearing loss in nonexposed mice and the effects of the AAE. In the absence of AAE exposure, young C57BL/6J and BALB/cJ males and females have similar ABR thresholds, but hearing loss deteriorates more rapidly in females after around age 6 months. The auditory systems of males and females also respond differently to the AAE. Depending on gender, both positive and negative effects of AAE exposure are observed, although these effects are complicated.

11:20

2aPP7. Neural substrates for presbycusis: Anatomy, chemistry, and neuroimaging of the central auditory system. Robert D. Frisina, Jr. (Otolaryngol. Div., Univ. of Rochester School of Medicine & Dentistry, 601 Elmwood Ave., Rochester, NY 14642-8629, rdf@q.ent.rochester.edu)

Overall, loss of hearing with age involves at least two main components: Reductions in sensitivity to sounds, starting with the high frequencies, and a decline in the ability to understand speech in the presence of background noise. The former is intimately linked to the age-related loss of: sensory cells, *stria vascularis* metabolic activity, and auditory-nerve fibers. The latter is affected by these peripheral changes, as well as by the pathologies of the central auditory system that occur with age. Recent findings concerning age-related plasticity in the structure, chemistry, and functional anatomy of the central auditory system will be presented to shed light on changes in the brain that may underlie the physiological and perceptual deficits in the processing of biologically relevant sounds presented by the other speakers of this session. Some changes include: declines in central auditory pathway connectivity, changes in the presence of calcium-binding proteins critical for intracellular calcium regulation and optimal neural functioning, reductions in GABA neurochemistry at the midbrain level, and changes in activity of the human brain during the processing of speech in background noise as seen in PET imaging experiments. [Work supported by NIA–NIH, and the International Center for Hearing & Speech Research, Rochester, NY.]

11:50–12:10

Panel Discussion

WEDNESDAY MORNING, 31 MAY 2000

FLAG ROOM, 8:30 TO 10:45 A.M.

Session 2aSAa

Structural Acoustics and Vibration: Surface Radiation

Lonny L. Thompson, Chair

Mechanical Engineering Department, Clemson University, 219 EIB, Clemson, South Carolina 29634

Contributed Papers

8:30

2aSAa1. The existence of internal acoustic radiation modes in a cylindrical acoustic cavity. Wayne M. Johnson (The G. W. W. School of Mech. Eng., Georgia Inst. of Technol., Atlanta, GA 30332, wj27@prism.gatech.edu)

The external acoustic radiation operator maps a boundary-valued quantity, such as the normal surface velocity of a radiator, to the pressure radiated into the field. It has been shown that the decomposition of the acoustic radiation operator results in a set of orthogonal singular velocity patterns or acoustic radiation modes. It has also been demonstrated that superposition of selected acoustic radiation modes can be used to prescribe surface velocity distributions resulting in poor radiation of pressure into the field. The development of acoustic radiation modes for the interior cavity problem may provide insight into the physics of structural acoustic coupling between, for example, a thin cylindrical shell and the enclosed acoustic volume. However, the existence of these radiation modes for the interior problem is uncertain. Also, if these interior radiation modes do exist, do they in fact coincide with the acoustic cavity or normal modes? This study intends to explore the existence of acoustic radiation modes for the interior cavity problem and their relationship to the cavity modes of a right circular cylindrical acoustic volume of air.

8:45

2aSAa2. High spatial density measurements of the vibro-acoustic response of some aircraft structures. Peter C. Herdic,^{a)} Brian H. Houston, and Earl G. Williams (Naval Res. Lab, Code 7130, Washington, DC 20375, herdic@lpsa2.nrl.navy.mil)

Experiments involving spatially dense interior pressure and surface velocity measurements (10–1000 Hz) on aircraft structures both in the laboratory and in-flight are reported. In the laboratory experiments, the structures were predominately excited by point forcing functions applied at a number of locations on the fuselage. The structural-borne energy couples energy into the interior at resonant frequencies of both the structure and the interior cavity. The normal modes of the cavity tend to dominate the problem at frequencies above 100–200 Hz where the density of states is very high. Further, we also report on in-flight measurements involving the use of a scanned microphone array (55 microphones at 32 axial stations) in the passenger cabin of a turbo-prop powered aircraft. The surface velocity was determined through near-field acoustic holography (NAH). The counter-clockwise rotating propellers sweep past the fuselage, resulting in a slap of air pressure on the aircraft sidewalls. This motion creates a dominant antisymmetric (left/right) surface velocity and interior pressure distribution at the blade passage frequency and its harmonics. The results associated with two laboratory experiments and one in-flight

experiment, along with some comparisons to complex three-dimensional finite element-infinite element numerical models, will be reported. ^{a)}Also with SFA, Inc., Landover, MFD

9:00

2aSAa3. Measurement of the backscatter from two pressure-release spheres using a near-field measurement system. Thomas Logan, Brian Prasse, and Peter Rogers (Georgia Inst. of Technol., School of Mech. Eng., Atlanta, GA 30332, thomas.logan@me.gatech.edu)

The backscattering of sound by two pressure-release spheres was studied using a near-field measurement system (NFMS) based upon the near-field calibration array (NFCA) concept developed by Trott and Van Buren. The NFMS consists of a planar near-field transmit array (NFTA) consisting of 34 real-shaded line elements comprised of 26 hydrophones and a synthetic cylindrical aperture near-field receive array (SA-NFRA) consisting of a single line element comprised of 26 hydrophones. The line elements of the NFTA were shaded to obtain a volume of plane wave uniformity of 1 m^3 in front of the NFTA. The synthetic aperture was formed by moving the single line element with a precision positioning system. The total acoustic pressure field was measured by the SA-NFRA. Since the incident pressure and the scattered overlapped in the time domain, background subtraction was used to obtain the scattered pressure. Far-field scattered beam patterns for the SA-NFRA were formed using the NFCA concept to extrapolate near-field measurements to the far field. Three configurations of pressure-release spheres were studied: a single sphere, two spheres arranged perpendicularly to incident field, and two spheres arranged parallel to the incident field. Experimental and numerical model results will be presented as well as NFTA uniformity measurements.

9:15

2aSAa4. A combined experimental/FEM model for predicting vehicle panel acoustic noise contribution. Teik Lim (Dept. of Mech. Eng., The Univ. of Alabama, 290 Hardaway Hall, Box 870276, Tuscaloosa, AL 35487-0276, tlim@coe.eng.ua.edu)

Radiated acoustic noise contributions from the structural panels that form the inner shell of the vehicle body to the interior sound pressure response are modeled using an approximate spectral formulation. The approach essentially utilizes the theoretical interior acoustic sensitivity terms derived from a finite element model and measured spatial-averaged structural-acoustic spectra. The finite element calculation is validated by comparison to a set of experimental acoustic transfer functions. In addition, an experimental setup for measurement of the sound intensity and structural-acoustic response is used to acquire the cross and auto power spectra needed to predict the relative mean-squared velocity term of each control plane near the panel surface. These velocity functions are then combined with the acoustic sensitivity terms to compute the individual panel contribution function. The proposed approach is applied to a specific case of a passenger car to compute the acoustic noise spectra in 1/12 octave band form at selected positions in the passenger compartment. The calculations show excellent match with the overall experimental results. Also through this example, the approximate computational scheme is proven to be generally quite robust and effective for analyzing higher frequency interior noise problems.

9:30

2aSAa5. Analysis of reconstruction of radiated acoustic pressure fields by using the HELS method. Nassif Rayess and Sean Wu (Dept. of Mech. Eng., Wayne State Univ., 5050 Anthony Wayne Dr., Detroit, MI 48202, swu@me1.eng.wayne.edu)

This paper presents a detailed analysis of reconstruction of the acoustic pressure fields radiated from a partially vibrating sphere by using the Helmholtz equation-least squares (HELS) method [Z. Wang and S. F. Wu, J. Acoust. Soc. Am. **102**, 2020–2032 (1997); **104**, 2054–2060 (1998)]. The objectives of this study are to gain insight into an acoustic inverse problem and to develop guidelines that may be helpful in the future reconstruction work. Because of the complexities involved, many questions

such as where and how many measurements should be taken, how many expansion terms are necessary for a given number of measurements, etc., remain unanswered. These questions will be tackled here. The reasons for selecting a partially vibrating sphere are twofold: (1) the resulting acoustic pressure field contains a very pronounced near-field effect; and (2) analytic results are known so the reconstructed acoustic pressures can be checked rigorously. The impact of a loss of the near-field effect on reconstruction will be investigated. The effects of measurement distance, area, and shape on the accuracy of reconstruction will be examined. The efficiency of reconstruction using the HELS method as the increase in frequency will also be studied. [Work supported by the NSF, Grant No. CMS-9802847.]

9:45

2aSAa6. Combined Helmholtz equation-least squares (CHELS) method for reconstructing radiated acoustic pressure fields. Sean Wu and Xiang Zhao (Dept. of Mech. Eng., Wayne State Univ., Detroit, MI 48202, swu@me1.eng.wayne.edu)

A combined Helmholtz equation-least squares (CHELS) method is developed for reconstructing acoustic pressures radiated from an arbitrary object. This method combines the HELS method [Wang and Wu, J. Acoust. Soc. Am. **102**, 2020–2032 (1997); Wu and Yu, *ibid.* **104**, 2054–2060 (1998)] and boundary element method (BEM)-based Helmholtz integral formulation. The main advantage of the HELS method is its simplicity in expressing the radiated acoustic pressure in terms of an expansion of orthonormal basis functions. However, the accuracy of reconstruction may deteriorate with the increase of complexity of source geometry. This is because the convergence of the basis functions becomes poorer for an irregular surface than for a smooth surface. The BEM-based integral formulation allows for reconstruction of the radiated acoustic pressures on an arbitrarily shaped object. However, the number of measurements required for reconstruction is often too large to make the process practical. In the CHELS method, only a finite number of measurements are taken to establish the HELS method. Once this is done, enough field pressures are generated using the HELS method, which are then taken as input to the BEM-based integral formulation to reconstruct acoustic pressures everywhere. [Work supported by NSF Grant No. CMS-9802847.]

10:00

2aSAa7. Accurate nonreflecting boundary conditions for time-dependent acoustic scattering. Lonny L. Thompson, Runnong Huan, and Dantong He (Dept. of Mech. Eng., Clemson Univ., Clemson, SC 29634, lonny.thompson@ces.clemson.edu)

Accurate radiation boundary conditions for the time-dependent wave equation are formulated in the finite-element method as an auxiliary problem for each radial harmonic on a spherical boundary. The method is based on residuals of an asymptotic expansion for the time-dependent radial harmonics. A decomposition into orthogonal transverse modes on the spherical boundary is used so that the residual functions may be computed efficiently and concurrently without altering the local/sparse character of the finite-element equations. The method has the ability to vary separately, and up to any desired order, the radial and transverse modal orders of the radiation boundary condition. With the number of equations in the auxiliary Cauchy problem equal to the transverse mode number, the conditions are exact. If fewer equations are used, then the boundary conditions form high-order accurate asymptotic approximations to the exact condition, with corresponding reduction in work and memory. Numerical studies are performed to assess the accuracy and convergence properties of the radiation boundary conditions. The results demonstrate dramatically improved accuracy for time-domain simulations compared to standard boundary treatments and improved efficiency over the exact condition. [Work supported by NSF PECASE.]

2aSAa8. Exact radiation conditions on spheroidal boundaries with sparse iterative methods for efficient computation of exterior structural acoustics. Lonny L. Thompson and Cristian Ianculescu (Dept. of Mech. Eng., Clemson Univ., Clemson, SC 29634, lonny.thompson@ces.clemson.edu)

Exact radiation boundary conditions are derived from harmonic expansions in spheroidal coordinates and formulated in a finite-element method for the Helmholtz equation in unbounded domains. The use of spheroidal boundaries enables the efficient solution of scattering from elongated objects. Extending the ideas of Malhotra for the exact Dirichlet-to-Neumann (DtN) map defined on a sphere, a matrix-free implementation of the non-local DtN map for spheroidal boundaries, suitable for iterative solution of the resulting complex-symmetric system, is described. Efficient SSOR preconditioners together with Eisenstat's trick based on the sparse matrix partition associated with the interior mesh and local part of the radiation boundary operator are used to accelerate convergence of Krylov-subspace iterative solution methods. Numerical examples are computed to demonstrate the efficiency and accuracy of the boundary treatments for high-frequency scattering from elongated structures. [Work supported by NSF PECASE.]

2aSAa9. Lumped-element analysis of a fluid-loaded pipe-hull coupling. Jayme J. Caspall, Minami Yoda, and Peter H. Rogers (Woodruff School of Mech. Eng., Georgia Inst. of Technol., Atlanta, GA 30332-0405)

This research addresses the question: how much acoustic energy is coupled into the fluid and the structure at the end of a fluid-filled pipe connected to an externally fluid-loaded shell? As a first approximation, the hull is modeled as a flat elastic plate with an infinite acoustic medium on one side and a vacuum on the other side. The pipe is modeled as a lumped mass with finite rotational inertia attached to the hull and founded on imaginary fixed points by torsional and linear springs. Energy transmission through the simple structure to the far field is examined by allowing one of the fixed foundation points to move while applying a force at that point. The radiated acoustic power is calculated as a function of frequency. The effect of the simple structure on the radiated power spectrum is determined by comparing the results to the response of a point driven plate. [This work was supported by the ONR Structural Acoustics Program under Grant No. N00014-98-1-0851.]

WEDNESDAY MORNING, 31 MAY 2000

FLAG ROOM, 11:00 A.M. TO 12:00 NOON

Session 2aSAb

Structural Acoustics and Vibration and Noise: Combining Active and Passive Control of Vibration and Noise I

Kenneth A. Cunefare, Chair

School of Mechanical Engineering, Georgia Institute of Technology, Atlanta, Georgia 30332-0405

Contributed Papers

11:00

2aSAb1. Optimal design of Stockbridge dynamic vibration absorber viscoelastically modified by applying the finite element method. Sergio Floody (Acoust. Dept., Univ. Tecnológica Vicente Perez Rosales, Brown Norte 290, Santiago, Chile) and Jose Espindola (Univ. Federal de Santa Catarina, Campus Trindade, Florianopolis, SC, Brazil)

The present paper deals with an extension in the study of vibration absorbers applied to complex mechanical systems that are independent of geometry, mass, stiffness, and damping distribution by introducing the concept of equivalent generalized quantities. The basic idea is to transform the mechanical impedance, that the absorber transfers in the primary system's point of connection, in generalized equivalent quantities of mass and damping that depends on the frequency. This methodology was applied in order to design a viscoelastically modified Stockbridge absorber by applying the finite element method extending the proposed theory of generalized equivalent quantities for several degrees of freedom that can be presented due to the characteristics of this absorber. Later, the genetic algorithm was used to minimize the frequency response of the joint system in order to obtain the optimal dimensions of this mechanical device.

11:15

2aSAb2. State-switched absorber for control of beams. Kenneth A. Cunefare (School of Mech. Eng., Georgia Inst. of Technol., Atlanta, GA 30332-0405, ken.cunefare@me.gatech.edu)

A system that has the capability to make instantaneous changes in its mass, stiffness, or damping may be termed a state-switchable dynamical system. Such a system will display different dynamical responses dependent upon its current state. State-switchable stiffness may be practically

obtained through the control of the termination impedance of piezoelectric stiffness elements. If such a switchable stiffness element is incorporated as part of the spring element of a vibration absorber, the change in stiffness causes a change in the resonance frequencies of the system, thereby instantaneously retuning the state-switched absorber to a new frequency. In between state switches, the operation of such a device is passive, being fundamentally a passive vibration absorber. This paper considers the application of such a device for the purpose of vibration control on beams.

11:30

2aSAb3. Active noise control on a large-scale ducted fan with inlet wave-number sensing. Jerome P. Smith, Ricardo A. Burdisso (Vib. and Acoust. Labs., Dept. of Mech. Eng., Virginia Tech, Blacksburg, VA 24061-0238), Daniel L. Sutliff, and Laurence J. Heidelberg (NASA Glenn Res. Ctr., Cleveland, OH 44135)

Active noise-control (ANC) methods for reduction of turbofan inlet noise have been previously demonstrated on a business-sized turbofan engine. The potential of ANC is investigated here on a large 4-ft-diameter ducted fan which has a geometry and tone frequencies which are more characteristic of current and future large production engines. The control inputs are implemented with a single circumferential array of acoustic drivers constructed with rare-earth magnets. Both far-field pressure transducers and an inlet wave-number sensing technique are investigated as error-sensing strategies to reduce the 2BPF tone of the ducted fan (approximately 1000 Hz.) The wave-number sensing strategy involves filtering and combination of inlet-mounted microphone signals to result in the observation of a particular wave number propagating axially out of the inlet. The ducted fan is equipped with a rotating microphone array in the inlet, which allows for modal analysis of the inlet acoustic field before and

after control. The in-duct mode measurements provide insight into the control mechanisms. The test facility is also equipped for making far-field pressure measurements. Experimental results for a single circumferential control source array in conjunction with far-field pressure and inlet wave-number sensing techniques will be presented and discussed. [Work supported by NASA Langley Research Center.]

11:45

2aSab4. An adaptive active noise control system in ducts. Shimin Peng (The Environment Protection Bureau, Loudi City, Hunan Province 417000, PROC)

The active noise control (ANC) technique is used to cancel original noise by actively generating a secondary sound signal. Actually it is a method using the interference effect of sound waves. The first part of this

paper will introduce and discuss the ANC technique, including the basic concepts and principles, history, current development, and application. Some application examples of ANC technique in duct noise control are given here. The second part will introduce the necessary digital signal processing knowledge and system identification technique which is used in this noise control system. The mathematic model of the sound feedback channel and error signal channel is established by using the system identification technique. The third part of this paper will introduce the ANC experiment results in an air-duct, and an important conclusion is given. Many factors related to ANC are investigated; these factors include the position of the microphones in the duct, the vibration of the shells of the duct, the reflection of the sound waves from the end of the duct, Very excellent results were achieved by using this system; it reduces the wide-band noise in the range of 0–800 Hz effectively; 30-dB reduction of sound power level was obtained to the white noise of 250 Hz oct.

WEDNESDAY MORNING, 31 MAY 2000

HENRY/GRADY ROOM, 8:30 TO 11:30 A.M.

Session 2aSC

Speech Communication: Cross-Linguistic Aspects of Speech Perception and Production (Poster Session)

James E. Flege, Chair

Department of Rehabilitation Sciences, University of Alabama, Birmingham, Alabama 35294-0019

Contributed Papers

All posters will be on display from 8:30 a.m. to 11:30 a.m. To allow contributors an opportunity to see other posters, contributors of odd-numbered papers will be at their posters from 8:30 a.m. to 10:00 a.m. and contributors of even-numbered papers will be at their posters from 10:00 a.m. to 11:30 a.m.

2aSC1. Development of pitch perception and memory of general psychological abilities in children speaking tonal languages. A comparative approach. Eduardo Castro-Sierra (Lab. of Psychoacoustics, Hospital Infantil de Mexico (HIM) and Natl. School of Music (ENM, UNAM), Dr. Marquez No. 162, 06720 Mexico, D.F., Mexico), Dinorah Romero-Garibay (Hospital Infantil de Mexico, 06720 Mexico, D.F., Mexico), and Marha Gomez-Gama (ENM, UNAM)

The development of pitch perception and memory in language and music among 6- to 15-year-old Otomi-speaking children of either sex and Zapotec-speaking children of the same age and of either sex was studied. These languages are characterized by the use of lexical tones. For the analysis of musical abilities, Bentley's Tests of Pitch Discrimination and of Tonal Memory were employed; for that of language abilities, special tests were devised that measured the perception and memory of tonal changes in either Otomi or Zapotec words. Results showed that development of abilities in music in Otomi-speaking children followed a course that was parallel to, although significantly below, that obtained by Bentley in his original study of 1966 in England. In the Otomi-speaking group, the course of development of abilities in language was above that of those in music. On the other hand, the same tests in Zapotec-speaking children provided data that were very irregular and further below the average obtained by the Otomi-speaking children. Interestingly, the application of Raven's Progressive Matrices and Koppitz's Test of Human Figure Drawings to these two groups of subjects indicated that, generally, Zapotec-speaking children had a higher cognitive level than Otomi-speaking children. [Work supported by Fondo de Neurologia, Hospital Infantil de Mexico.]

2aSC2. The intraspeaker variability of pitch in Mandarin and English subjects producing disyllabic words. Shawn Nissen and Christopher Gibbons (Dept. of Speech and Hearing Sci., The Ohio State Univ., 110 Pressey Hall, 1070 Carmack Rd., Columbus, OH 43210-1002, nissen.9@osu.edu)

Considering their use of pitch to convey lexical information, do tone language speakers develop greater control over pitch mechanisms, thus exhibiting less overall pitch variability than speakers of intonation languages? One previous researcher has argued that tone language speakers exhibit a "remarkably stable and precise absolute pitch template" when reading word lists [Deutsch *et al.*, J. Acoust. Soc. Am. **106**, 2267 (1999)]. The present study aims to investigate the intraspeaker pitch variability of Mandarin and English speakers when producing disyllabic words. Specifically, this research will examine (1) intraspeaker pitch variability between Mandarin and English speakers, (2) the degree of pitch stability exhibited by speakers of both languages, and (3) the effects of elicitation condition on a speaker's pitch variability. Subjects will produce initially stressed, disyllabic words (similar segmental structure) in their native language elicited through picture identification. Randomized blocks containing similar target items will be presented under two elicitation conditions. To measure possible "list effects" the second elicitation condition will contain an intervening masking task. Measures of fundamental frequency change of target nuclei will be obtained at four locations: start point, end point, high point, and low point. A three-way ANOVA will be used to analyze the pitch variability data.

2aSC3. Perception and production of English vowels by Spanish adult L2 learners: Effects of experience and proficiency. Ratree P. Wayland and James E. Flege (Prog. in Linguist., 209 Yon Hall, Univ. of Florida at Gainesville, FL 32611-5454, ratree@lin.ufl.edu)

This study examined the effects of experience and proficiency on the perception and the production of English vowels by native Spanish adults. The 60 subjects were assigned to subgroups according to the length of their residence in the United States (3 LOR groups) and overall degree of foreign accent in English (three proficiency groups). Experiment 1 assessed the perceived relationship between 11 English and 5 Spanish vowels. Experiment 2 examined the subject's categorial discrimination of pairs of the English vowels, and exp. 3 assessed the subjects' accuracy in producing eight English vowels. The native Spanish subjects' perception and production of English vowels changed as a function of their overall proficiency in pronouncing English. An examination of relation between perception and production suggested that perceived similarity of English vowels to Spanish vowels, but not the subjects' categorial discrimination of vowels, was correlated with subjects' vowel production accuracy. Also, a stronger correlation was found to exist between the perceived relation between English and Spanish vowels and subjects' overall degree of foreign accent than between their vowel discrimination scores and overall degree of foreign accent.

2aSC4. The effect of immersion on Japanese listeners' labeling of American English vowels. Janet W. Stack (Commun. Disord. Prog., Univ. of Virginia, 2205 Fontaine Ave., Charlottesville, VA 22903, jws8n@virginia.edu), Rieko Kubo (ATR Human Information Processing Res. Labs., Kyoto, Japan), and John S. Pruitt (Microsoft Corp., Redmond, WA 98052-6399)

A set of 118 synthetic stimuli was presented twice to three groups of listeners to be labeled as American English (AE) vowels. The stimuli were generated through a Klatt cascade formant synthesizer, where first ($F1$) and second ($F2$) formant frequencies were varied factorially. $F1$ ranged from 200–700 Hz in 50-Hz steps. $F2$ ranged from 600–2600 Hz in 200-Hz steps. The labeling of 11 AE vowels was compared for AE listeners, Japanese listeners living in the United States (immersed), and Japanese listeners living in Japan (nonimmersed). In addition to labeling, all listeners judged the goodness of the stimuli on a 1–7 rating scale. The modal responses of the immersed Japanese listeners were significantly more similar to those of the AE listeners than were those of the nonimmersed Japanese. The immersed Japanese also had higher ratings than the nonimmersed Japanese listeners, showing greater confidence in their labeling of AE vowels. As an additional task, the nonimmersed Japanese also labeled the stimuli as the five Japanese vowels, and demonstrated that Japanese vowel identities largely differed from AE vowels on $F1$ – $F2$. This difference may be a factor in Japanese listeners' difficulty in perceiving and producing AE vowels.

2aSC5. Does early speaking experience in Korean help adult relearners? Janet S. Oh (UCLA Dept. of Psych., Box 951563, Los Angeles, CA 90095-1563), Sun-Ah Jun, and Terry K. Au (UCLA, Los Angeles, CA 90095-1563)

Research indicates that early linguistic input is crucial for acquiring native-like phonology and morphosyntax. This study investigates whether early *discontinued* linguistic experience helps adult relearners master the phonology of a language. Students from first-year Korean classes were classified into three groups: no/minimal speaking experience, childhood speaking experience (until about age 5), and native speakers. Native speakers not enrolled in Korean classes also participated. Participants produced sentences that included words utilizing the three-way contrast in Korean stop consonants (lenis, tense, and aspirated). Measurements of stop closure duration indicate that only the two native speaker groups were differentially producing lenis and tense consonants. Measurements of VOT indicate that only the childhood and the native speaker groups were differentially producing the three consonants. Furthermore, the childhood

speakers' general pattern of closure duration and VOT for the three consonants was similar to those of the native speakers (this was not true for the no/minimal experience group). Results indicate that even early discontinued speaking experience can provide an advantage in relearning a language as an adult. [Work supported by an NIMH Developmental Cognitive Science Training Grant awarded to Janet Oh, and by NIMH Grant MH56118 and a UCLA Academic Senate Grant awarded to Terry Au.]

2aSC6. Persistent errors in the perception and production of word-initial English stop consonants by native speakers of Italian. Ian R. A. MacKay (Dept. of Linguist., Univ. of Ottawa, 70 Laurier Ave. E., Ottawa, ON K1N 6N5, Canada, imackay@uottawa.ca), James E. Flege, and Thorsten Piske (Univ. of Alabama at Birmingham, Birmingham, AL 35294)

This study examined the perception and production of word-initial tokens of English stops (/b d g/ and /p t k/) by native speakers of Italian. The native Italian subjects were assigned to one of four groups based on their age of arrival (AOA) to Canada from Italy and percentage of self-reported use of the native language, Italian. The results obtained here suggested that AOA was a more important predictor of the native Italian subjects' perception and production of word-initial English stops than L1 use was. The results also provided evidence of native versus non-native differences in segmental perception and production that persisted after decades of frequent second-language use. As hypothesized, the native Italian subjects produced English /p t k/ more accurately than /b d g/. In a perception experiment examining naturally produced English stops, the native Italian subjects misidentified short-lag tokens of English /b d g/ as /p t k/ more often than they misidentified long-lag /p t k/ tokens as /b d g/. Late bilinguals erred more often in identifying the voicing feature in /b d g/ than did early Italian–English bilinguals or native speakers of English, apparently because Italian /p t k/ are realized with short-lag VOT values.

2aSC7. Effect of childhood language experience on pronunciation in adulthood. Leah M. Knightly, Terry Kitfong Au, and Laura E. España (Dept. of Psych., UCLA, Los Angeles, CA 90095-1563, leah@psych.ucla.edu)

This study examined the effect of childhood experience with Spanish on Spanish pronunciation in adulthood. Twenty-eight native speakers of Spanish were asked to rate the pronunciation of Spanish consonants /p, t, k/ and /b, d, g/ produced by five groups of English-speaking adults: (1) Eighteen native Spanish speakers, (2) nine childhood Spanish speakers who had spoken Spanish fluently during childhood for at least three years, (3) six addressees who had regularly been addressed in Spanish during childhood but had spoken Spanish minimally, (4) sixteen overhearers who had overheard Spanish regularly as children but had spoken Spanish minimally, (5) seventeen late learners who had no experience with Spanish until high-school Spanish class. Childhood speakers, addressees, and overhearers all received reliably better accent ratings than those who first learned Spanish in high school [by HSD, $p < 0.05$]. Addressees obtained accent ratings comparable to native speakers for voiceless stops. However, only childhood speakers received accent ratings comparable to native speakers for both voiced and voiceless stops. Acoustical analyses revealed that those speakers who received more native-like ratings had shorter voice-onset time for /p, t, k/ and a higher percentage of lenited intervocalic /b, d, g/ [$r_s > 0.76$, $p_s < 0.0001$].

2aSC8. Acquisition of second language intonation. Sun-Ah Jun (Dept. of Linguist., UCLA, Los Angeles, CA 90095-1543, jun@humnet.ucla.edu) and Mira Oh (Yeojoo Inst. of Technol., Korea)

Foreign accents in second language (L2) production are caused by the phonological system and phonetic realization of the first language (L1), including both segmental and prosodic features. This paper examines the intonation structure of Seoul Korean and its realization by American En-

glish speakers. Four English speakers of Korean, differing in fluency, and two Korean speakers participated. Thirty-six sentences were designed to test the realization of prosodic structure, focus, and intonation patterns by varying the number of syllables within a word and a sentence, and conditions for the segment–tone interaction. Preliminary results show that, as in segmental data, more advanced L2 speakers produced more nativelike intonation pattern and prosodic structure than less advanced ones. Advanced L2 speakers were better in grouping words into phrases and producing the phrase-final high tone than less advanced speakers, but they still produced English-like pitch accents following English stress rules, suggesting that a lexically linked prosodic feature is hard to suppress. In addition, less advanced L2 speakers more often produced each word in one intonation phrase, reflecting a limitation of information processing as found in L1 acquisition. Finally, more advanced speakers distinguished the segmentally triggered tones better than less advanced speakers but with large variation.

2aSC9. Reevaluating the effect of length of residence (LOR) on adults' performance in a second language. James E. Flege and Serena Liu (Dept. of Rehabilitation Sci., VH503, Univ. of Alabama at Birmingham, Birmingham, AL 35294, jeflefe@uab.edu)

Previous work has shown that children but not adults make regular progress in a second language (L2) as length of residence (LOR) in an L2-speaking country increases. These findings might be interpreted to mean that adults' ability to learn an L2 is severely constrained. The present study examined the effect of LOR in a more controlled fashion than previous studies. Experiment 1 assessed the identification of word-final English consonants. Experiment 2 examined the scores on a 144-item grammaticality judgment test. Experiment 3 examined the scores on a 45-item listening comprehension test. Adult Chinese participants were assigned to one of four groups of 15 each based on LOR in the United States and academic status (students versus nonstudents). In all three experiments, significantly higher scores were obtained for students with a relatively long LOR than students with a shorter LOR. However, the difference between nonstudents differing in LOR was nonsignificant in every instance. When taken together, these results suggested that the negative findings for LOR that were obtained in previous studies might have been due to sampling error. Adults' performance in an L2 seems to improve measurably as LOR increases, but only if they receive a substantial amount of native-speaker input. [Work supported by NIH.]

2aSC10. Neural activation patterns during discrimination of a native versus non-native contrast by Japanese speakers. Tobey L. Doeleman (Dept. of Psych., Box 357920, Univ. of Washington, Seattle, WA 98195, doeleman@u.washington.edu), Patricia K. Kuhl, Ryan J. Conley, and Erica B. Stevens (Univ. of Washington, Seattle, WA 98195)

Discrimination of a native versus a non-native consonantal contrast was investigated using two experimental paradigms in a functional magnetic resonance imaging (fMRI) examination of neural activation patterns. Native Japanese speakers were tested on both a /b/–/d/ (native) contrast and an /r/–/l/ (nonnative) contrast in same–different and AXB conditions. All consonant tokens were paired with vowels [a, i, u, e] in two different phonological environments, syllable initial and intervocalic, and discrimination was measured behaviorally both in and out of the scanner. Stimuli in the same–different task consisted of VCV syllables, while CV syllables were presented in the AXB task. The use of different syllable structures in the two tasks ensured that the activation patterns common to both conditions were in fact due to processing of the abstract phoneme representations. Activation during discrimination was calculated as those voxels were significantly more active than both baseline and recovery periods on

two independent runs. A comparison of patterns for the native versus non-native contrast reveals differences in processing characteristics such as location and extent of activation.

2aSC11. Locus equation parameters in non-native speech. Eunjin Oh (Dept. of Linguist., Stanford Univ., Stanford, CA 94305-2150, eunjin@turing.stanford.edu)

The principled relation between the magnitude of CV coarticulation and the language-specific system of phonological contrasts was investigated by comparing slope values of the locus equation [Lindblom, J. Acoust. Soc. Am. **35**, 1773–1781 (1963)] for three stop places of articulation in English and Russian. The slopes of the locus equation were hypothesized to be shallower in Russian, indicating less vocalic influences on adjacent stop consonants, than in English, due to the requirement maintaining the phonological distinctions between plain and palatalized stops imposed only on Russian [Cohn, UCLA WPP **69**, 51–59 (1987)]. More importantly, this study compares the slope, y-intercept, and R-squared values of native speakers with those of English learners of Russian. Preliminary analyses on bilabial stops indicated the following. (i) The locus lines of Russian were characterized by substantially shallower slopes and larger y-intercepts than those of English. (ii) The locus lines of relatively experienced English speakers of Russian showed shallower slopes and larger y-intercepts than those of native English speakers. An advantage of investigating the locus equation parameters as indices of acquisition effects of CV coarticulation is that the locus equation parameters characterize the acquisition of coarticulatory regularities of a whole distribution of CV tokens, not just the acquisition of a single CV token.

2aSC12. Acoustics and perception of Russian plain-palatalized contrast. Alexei Kochetov (Dept. of Linguist., Univ. of Toronto, 130 St. George St., Rm. 6076, Toronto, ON M5S 3H1, Canada)

Acoustic and perceptual correlates of Russian syllable-final plain and palatalized coronal stops were investigated. The acoustic analysis based on the data from six talkers showed that the stops were consistently released both word-finally and before stops. Both consonants in final position were strongly affricated and aspirated. The VC transitions to palatalized stops extended for more than half of the vowel, substantially affecting its quality. The perception experiment (12 listeners) determined that both bursts and transitions were perceptually relevant to the identification of the two phonemes. The contrast in final and preconsonantal environments differed in the relative importance of these cues: VC transition was more important for the plain-palatalized distinction in final position, while burst was a major cue to the contrast in preconsonantal position. The findings suggest that the presence of the plain-palatalized phonemic contrast in perceptually less favorable syllable coda positions is compensated for by additional acoustic information that assists a listener in the recoverability of the contrast. [Work supported by Social Sciences and Humanities Research Council of Canada.]

2aSC13. Coarticulation and stress across word boundaries in Russian. Priscilla McCoy (1232 Milvia St., Berkeley, CA 94709, pmccoy@socrates.berkeley.edu)

There is a rule in Russian phonology that can be stated as follows: /i/ > [-i] / C__ . [i/ is realized as [-i] (barred i) in the environment after a nonpalatalized consonant.] This rule holds true word internally (e.g., prefixation), at prepositional boundaries, but shows significant variation across word boundaries, particularly when stress and timing are considered. Phrases focusing on two-word sequences were recorded by six speakers, where the timing of the phrases was varied by fast or careful speech. The constituent structure of these phrases is illustrated below (. = slight pause). (A) Stressed V1: 1. Eto moj brat "Igor." This my brother Igor. fast –[t-i]; careful –[t.i] 2. Vot idet "Ira." There goes Ira. fast –[t-i]; careful –[t.i] . (B) Stressed V2: 3. brat Iv'ana brother of Ivan fast –[t-i];

careful –[t-i]. (4) idet Ir'ina goes Irina fast –[t-i]; careful –[t-i]. Formants of [i] versus [-i] were measured to examine coarticulation, as evidenced by a barred i in the phonetic realization of /i/. Results showed that (A) stressed V1 of the final word has no coarticulatory effects in careful speech, whereas (B) stressed V2 adhered to the rule.

2aSC14. Effects of syllable structure and syllable boundary on segment duration in Seoul Korean. Byung-jin Lim (Dept. of Linguist., Indiana Univ., Bloomington, IN 47405, bylim@indiana.edu)

The present study reports results of an investigation into Korean durational patterns with respect to syllable structure and syllable boundary. The main questions raised in this study were (1) whether different preceding syllable structures contribute to the durational patterns of the following segmental durations: closure duration, voice onset timing (VOT), and vowel duration; and, if so, (2) how they are realized, especially over syllable boundaries determined by Korean orthography. Thirty-six two-syllable nonsense words varying in syllable structure were analyzed. For the durational pattern of internal elements of syllables over syllable boundary, the initial consonants of the second syllable were differentiated into lax, tensed, and aspirated stops. Results indicate that initial consonant duration showed no significant difference regardless of syllable structure. Initial vowel duration, however, reflected temporal compensation depending on the syllable structure. In addition, the effect of preceding syllable structure was reflected only on the closure duration, not on the VOT nor on the vowel durations of following syllables. These findings imply that temporal compensation tends to occur across segments over syllable boundaries, suggesting the importance of syllable boundary in durational patterns in Seoul Korean. The relevance of orthography to syllable boundary is also discussed.

2aSC15. The effect of neutral tone on the f_0 peak alignment in Standard Chinese. Zhiqiang Li (Dept. of Linguist., MIT, 77 Massachusetts Ave., Cambridge, MA 02139, zqli@mit.edu)

The f_0 peaks sometimes are realized after the right-hand boundary of their host syllables when followed by syllables that carry one of the four tones in Standard Chinese. Further, tone 1 (*H*) and tone 2 (*R*) exhibit different peak alignment patterns. This study investigates the effect of neutral tone on the alignment of f_0 peaks with their host syllables in tone 1 and tone 2 at normal speaking rate. When a syllable is unstressed, it does not carry any of the four tones, but is said to be in the neutral tone. The greatest amount of peak delay was found on tone 2 when it was followed by one or two neutral tone syllables. The f_0 peaks were aligned pretty close to the *C-V* boundary of the following syllable, and in many cases even across it. The peak on tone 1 was delayed the most in the context of two neutral tone syllables, always being realized after the right-hand boundary of its host syllable, and the least before a full tone syllable. The different alignment patterns are attributable to the inherent tonal properties of tone 1 and tone 2, and the phonological analysis of neutral tone.

2aSC16. Intrinsic versus interaction effects on F_0 . Mi-Ryoung Kim (Program in Linguist., Univ. of Michigan, 1076 Frieze Bldg., Ann Arbor, MI 48109-1285)

This study investigates the nature and magnitude of the effect of segments, initial and final consonants, vowel quality, and vowel length on F_0 contour in English and Korean. The two languages were compared with a special reference of the effect of initial consonants on F_0 . F_0 was measured at three temporal locations of each syllable rime onset, midpoint, and offset. For Korean, of the four segment types, initial consonants had the greatest effect on F_0 contour in both spectral and temporal magnitude, indicating that consonant types interacted with tone. The English data

showed that the segmental influence reflected intrinsic effects on F_0 . In particular, the effects of initial consonants on F_0 contours in English were significantly smaller in spectral and temporal extent than those of Korean. This study concluded that there are cross-linguistic differences between English and Korean concerning the effect of initial consonants on F_0 contour (which I call ‘intrinsic’ versus ‘interaction’). The consonant–tone interactions in Korean are viewed as consistent with tonogenesis: voiceless (aspirated and tense) consonants correlate lead H and phonologically voiced (lax, sonorants, and zero) consonants lead LH. This is the first study that accounts for the Korean interaction phenomena by a theory of tonogenesis.

2aSC17. Recoverability constraints on gestural overlap in Georgian stop sequences. Ioana Chitoran (Linguist. and Cognit. Sci., Dartmouth College, 6087 Dartmouth Hall, Hanover, NH 03755, ioana.chitoran@dartmouth.edu), Dani Byrd (Univ. of Southern California, Los Angeles, CA 90025, dbyrd@usc.edu), and Louis Goldstein (Yale Univ. and Haskins Labs., New Haven, CT 06510)

Recent investigations of the temporal organization of articulatory gestures have found that consonants in a cluster exhibit less temporal overlap when in word-onset position than when the cluster occurs word-finally or spanning a word boundary. A possible reason for this is that substantial overlap of obstruent gestures in utterance-initial position may threaten their perceptual recoverability. Recoverability considerations may also account for results showing that a front-to-back order of place of articulation in stop–stop sequences (labial–alveolar, alveolar–velar, labial–velar) allows more overlap than the opposite order. Presumably, the recoverability of C1 is hindered if the release of C1 has no acoustic manifestation due to the presence of a more anterior C2 being coproduced in time, hiding the C1 release. Data demonstrating both these constraints on gestural patterning have previously been drawn only from English, limiting the type and position of consonant sequences. The Georgian language offers a rich inventory of consonant sequences including stop–stop sequences in initial, medial, and cross-word-boundary positions. This work is an articulator movement tracking study of two Georgian speakers producing such sequences. The results provide evidence that the sequence’s word position and the component consonants’ constriction locations both constrain the patterns of gestural overlap produced. [Work supported by NIH.]

2aSC18. On the noncategorical perception of Cantonese level tones. Alexander L. Francis, Brenda K. C. Ng, and Valter Ciocca (Dept. of Speech and Hearing Sci., Univ. of Hong Kong, Hong Kong)

Previous research provides conflicting evidence about the categorical nature of lexical tone perception. Chan *et al.* [J. Acoust. Soc. Am. **58**, 119 (1975)] found that an f_0 continuum between a Mandarin level tone (high level) and a contour tone (high rising) was perceived categorically. In contrast, Abramson [J. Acoust. Soc. Am. **33**, 842(A) (1961)] found that a Thai level tone continuum ranging in f_0 from low level to high level was not perceived categorically. The present study examined whether level lexical tones are perceived categorically in Cantonese. Native Cantonese speakers were asked to identify and discriminate stimuli consisting of Klatt-synthesized Cantonese syllables ranging in f_0 from low level to high level in nine equal mel steps. For identification, syllables were presented medially in (synthesized) sentential context to facilitate talker-dependent pitch processing [P. Wong and R. Diehl, J. Acoust. Soc. Am. **104**, 1834 (1998)]. For discrimination, adjacent steps were presented in isolated pairs with 250 msec ISI. Although the identification function showed boundaries typical of categorically perceived continua, discrimination was poor

across the continuum. These results suggest that Cantonese level tones are not perceived categorically. Results and implications for theories of tone perception will be discussed.

2aSC19. A study of vowel length in Thai. Rungpat Roengpitya (Univ. of California at Berkeley, 1203 Dwinelle Hall, Berkeley, CA 94720)

In Thai, vowel length is contrastive (e.g., [cɪp] ‘‘to sip’’–[cɪ̃p] ‘‘to pleat’’). The main cue to distinguish short-long vowels in Thai is vowel duration [Abramson, 1962]. However, in some literature, there are other cues for vowel length in Thai, such as final nasal duration. This paper was

aimed to see how important vowel duration and final nasal duration are as perceptual cues for Thai listeners. A perceptual experiment was conducted. Eighteen (near-) minimal pairs of Thai words were used, consisting of initial voiceless unaspirated stop [p], nine pairs of Thai vowels, a final voiceless velar stop [k] or an alveolar final nasal [n]. Short vowels were lengthened and long vowels were shortened at 10–30-ms steps, using LPC analysis and resynthesis. The same procedure was done for short and long final nasals at 10-ms steps. All 1148 tokens were put into a Thai frame sentence and randomized. Forty native Thai listeners listened to these tokens and judged whether each token had a short or a long vowel. The results are that vowel duration is the main perceptual cue to distinguish short and long vowels in Thai.

WEDNESDAY MORNING, 31 MAY 2000

AMERICAN ROOM, 8:25 TO 11:30 A.M.

Session 2aSP

Signal Processing in Acoustics, Biomedical Ultrasound/Bioresponse to Vibration, Underwater Acoustics, and Acoustical Oceanography: Acoustical Imaging and Tomography

David H. Chambers, Chair
Lawrence Livermore National Laboratory, P.O. Box 5508, Livermore, California 94551-5508

Chair’s Introduction—8:25

Invited Papers

8:30

2aSP1. Quantitative three-dimensional imaging with diffracting wavefields. Anthony J. Devaney (Dept. of Elec. and Computer Eng., Northeastern Univ., Boston, MA 02115)

The general theory of quantitative imaging of three-dimensional, semi-transparent (weakly scattering) objects using scalar waves is presented. The theory is developed for the case of transmission type experiments appropriate to ultrasound diffraction tomography but is applicable with minor modifications to reflection geometries appropriate to pulse echo imaging. The imaging problem is formulated from first principles in terms of the three-dimensional scalar wave equation and is shown to reduce ultimately to an inverse scattering problem. By use of certain linearizing approximations the inverse problem is shown to reduce to a conventional (coherent) imaging problem having a well-defined point spread function (PSF) that can be computed in terms of the experimental parameters and imaging geometry. It is shown that this formulation allows quantitative three-dimensional imaging to be treated completely analogously to coherent optical imaging and, in particular, leads to a characterization of image quality in terms of the PSF and its spatial Fourier transform, the coherent transfer function (CTF). Inherent limitations of three-dimensional imaging are discussed based on the computed PSF for certain canonical geometries. These limitations are shown to be partially overcome by using suites of scattering experiments and/or beam scanning techniques such as focus-on-transmit and focus-on-receive. The talk includes a discussion of the validity of the weak scattering approximations that underlie the imaging model as well as a discussion of the use of the wave aberration function for characterizing image quality. The talk is illustrated with simulated and experimental results.

9:00

2aSP2. Computing resolution in acoustic tomography. James G. Berryman (LLNL, P.O. Box 808 L-200, Livermore, CA 94551-9900)

The resolution of images obtained using acoustic tomography is important both after the fact, as a means of evaluating the usefulness of the final image, and before the fact, when the measurement system of sources and receivers is being designed according to some preset criteria specified by the users. Resolution computations can actually be significantly more difficult and expensive than computing the image itself. Our purpose then is to show how resolution can be computed economically in almost all cases, and to provide a means of comparing the resolution characteristics of many of the common approximate inverse methods. Means of obtaining the effective resolution operator for such standard iterative inversion procedures as conjugate gradients, Lanczos, and LSQR are provided while avoiding the more common approach of producing a singular-value decomposition (SVD) of the operator being inverted. The methods discussed produce very easy-to-understand results for purposes of analysis, but require reorthogonalization of the Krylov vectors/operators produced by realistic iterative procedures in finite precision. Although the need for reorthogonalization increases the expense of the procedure somewhat, it still produces the desired results with much greater efficiency than could be obtained by computing an SVD of the full operator.

2aSP3. Diffraction tomography in acoustical imaging: Theoretical development, signal processing techniques, and applications. Hua Lee (Dept. of Elec. and Computer Eng., Univ. of California, Santa Barbara, CA 93106, huallee@ece.ucsb.edu)

Diffraction tomography has been playing a crucial role in acoustical imaging. In many applications, it has advanced far beyond the classical concept of tomography with narrow-beam illumination and intensity-mapping schemes to the level of multiple-angle multiple-frequency systems with monostatic and bistatic data acquisition configurations. In this presentation, in addition to an overview of the fundamental concepts and theoretical development of diffraction tomography in acoustical imaging, the practical signal processing techniques associated with image reconstruction will be discussed in detail. Furthermore, important subjects such as resolution limit, trade-off analysis between aperture span and signal bandwidth, and computation complexity will be fully analyzed. Responding to various physical constraints, the data acquisition process of a diffraction tomographic system can be achieved by using linear or circular receiver arrays. This presentation will also include the algorithm structures of the image formation algorithms for both linear and circular array systems to illustrate the commonalities and differences of these modalities. Results from various applications in microscopy, underwater sonar sensing, and medical imaging will be included.

10:00–10:15 Break

10:15

2aSP4. Fast, nonperturbative, transmission inverse scattering, quantitative imaging with laboratory and simulated acoustic data. Steven A. Johnson, David T. Borup, James W. Wiskin, Michael J. Berggren, Thierry Guillerum, and Scott C. Olsen (TechniScan, Inc., 350 W. 800 North, Ste. 305, Salt Lake City, UT 84103, s.a.johnson@techscan-systems.com)

Inverse scattering provides an ultimate method for acoustic imaging for medical and other fields of application (because it is quantitative and is self-correcting for refraction, absorption, and diffraction artifacts). E. Wolf showed that, with plane wave incident excitation, spatial distributions of inhomogeneous acoustic speed of sound and absorption within bodies could be reconstructed under the Born approximation by use of a simple back propagation algorithm operating in spatial frequency space (k -space). This approach, as extended to include the Rytov approximation and/or to real-space, is known as “diffraction tomography” (DT). However, DT fails at frequencies above 1 MHz, for objects with the size and contrast of the human female breast. We have extended DT by replacing its approximate wave equations with an exact wave equation. We present for the first time images of laboratory objects (agar cylinders, excised tissue, etc.) made by inverting transmission data. Present algorithms are at the boundary of real time operations or for 3D imaging with PC level computer networks. We discuss problems characteristic of inverse scattering such as multiple minima solutions and the need for receiver and transmitter element beam patterns. Input of these images for computer aided diagnosis is discussed.

Contributed Papers

10:45

2aSP5. Acoustic 3D computed tomography for demining and archeological applications. Stergios Stergiopoulos (Defence and Civil Inst. of Environ. Medicine, 1133 Sheppard Ave. W., Toronto, ON M3M 3B9, Canada), Robert Alterson (York Univ., Toronto, ON, Canada), David Havelock, and Julius Grodskil (Inst. for Microstructural Sci., Natl. Res. Council, Ottawa, ON K1A 0R6, Canada)

The 3D acoustic tomography concept of this paper combines computerized tomography image reconstruction algorithms using acoustic diffracting waves together with depth information to produce a novel, three-dimensional image of an underground section. The new method illuminates the ground with an array of acoustic sources along equally spaced points on the circumference of a surface of interest. For each transmitted pulse, the reflected–refracted signals are received by a line array of acoustic sensors located at a diametrically opposite point from the acoustic source line array. For a given depth, which is represented by a time delay in the received signal, a horizontal tomographic 2D image is reconstructed from the received projections. Integration of the depth dependent sequence of cross-sectional reconstructed images provides a complete three-dimensional overview of the inspected terrain. The method has been tested with an experimental system that consists of a line array of 4 acoustic sources, providing plane waves, and a receiving line array of 32 acoustic sensors. Real results indicate that the performance characteristics of the new method are affected by the reflective acoustic energy from above the ground objects and the very poor signal-to-noise ratio of the received signal due to the poor air–ground coupling.

11:00

2aSP6. Topographical image reconstruction via differential geometry in the ultrasonic tactile sensing system. Kenbu Teramoto (Dept. of Mech. Eng., Saga Univ., Saga 8408502, Japan)

This paper describes a topographic image reconstruction via differential geometry using a proposed acoustical tactile sensing probe. The proposed tactile sensing system has an ability to identify the principal curvatures of the object surface by detecting reflected wave fronts. Two major difficulties, however, exist in the principal curvature identification. The first reason is that the wave front reflected by the paraboloidal surface cannot be described in the combination of the plane waves nor the spherical waves strictly. The second is that the principal curvature identification process becomes ill-posed due to the nonlinear relationship between the principal curvatures and propagation time of flight. Avoiding these difficulties, the proposed sensing system seeks the unique point over the manifold which satisfies the Snell’s law. After the principal curvature estimation at several points over the object surface, the topographic image of the surface can be reconstructed via geometrical connection. Connecting identified principal curvatures and principal direction, the differential geometrical method then provides a smooth surface description and discrimination following plane and paraboloidal surfaces: (1) plane, (2) paraboloid, (3) elliptic paraboloid, (4) hyperbolic paraboloid. Further, several acoustical experiments show that the tactile sensor can identify the principal direction and curvatures of the paraboloidal surfaces.

2aSP7. Passive acoustic thermotomography—Correlation and noncorrelation variants. Victor I. Passechnik (ELDIS IRE RAS, Starosadskiy per., 8, 101 000, Moscow, Russia, passechn@bhome.msk.ru), Andrej A. Anosov, and Konstantin M. Bograchev (ELDIS IRE RAS, Starosadskiy per., 8, 101 000, Moscow, Russia)

Passive thermoacoustic tomography is a method for reconstructing temperature, 2D and 3D distributions inside objects, using the intensity of ultrasonic thermal acoustic radiation emitted from within the object. This radiation can be measured by piezotransducers (PT) both in noncorrelation and correlation mode. For noncorrelation mode, *a priori* information about ultrasound absorption coefficient distribution is needed for reconstruction.

In this case, the methods for the reconstruction of temperature distributions were analyzed by computer simulation. The standard error of reconstruction did not exceed 0.25 K (for about 300 measurements and fluctuation error of every measurement, 0.1 K). The physical experiments confirmed the predictions of the theory. In correlation mode, two spatial distributions—of the temperature and of the absorption coefficient—could be potentially reconstructed. The spatial correlation function of sound pressure from a source of thermal acoustic radiation was measured experimentally by a two-probe acoustic interferometer. The acoustic blackbody made from a narrow strip of a material with a large absorption coefficient served as the source of radiation. The spatial correlation function was computed; this function turned to be in good agreement with experimental data. [Work supported by Russian Basic Research Foundation Grant No. 98-01-00100.]

WEDNESDAY MORNING, 31 MAY 2000

ENGLISH ROOM, 8:25 TO 11:45 A.M.

Session 2aUW

Underwater Acoustics: Propagation Modeling and Analysis

Stewart A. L. Glegg, Cochair

Department of Ocean Engineering, Florida Atlantic University, P.O. Box 3091, Boca Raton, Florida 33431

Stanley A. Chin-Bing, Cochair

Naval Research Laboratory, Stennis Space Center, Mississippi 39529-5004

Chair's Introduction—8:25

Contributed Papers

8:30

2aUW1. Precise acoustic measurements of seafloor transponder and towed-vehicle positions. Aaron D. Sweeney, Carl D. Chadwell, John A. Hildebrand, and Fred N. Spiess (Scripps Inst. of Oceanogr., Marine Physical Lab., 8602 La Jolla Shores Dr., La Jolla, CA 92093, aaron@mpl.ucsd.edu)

We describe a project using acoustic techniques to determine the pattern of seafloor deformation between annual measurements. Using cross-correlation between transmitted and received signals, round-trip travel times precise to a few microseconds were achieved between a towed vehicle and seafloor transponders. The technique is sensitive to sound-speed changes as small as 0.01 m/s. The positions of the transponders and towed vehicle are calculated in a least-squares sense from a ray-trace model, providing relative positions for seafloor transponders with an accuracy of a few centimeters. Two experimental sites were surveyed. The first is in the interior of the Juan de Fuca plate, where the temporal variability of the near-bottom sound speed was 0.05 m/s, with little range dependence. Deformation at the few-centimeter level was neither observed nor expected due to remoteness from plate boundaries. The second site straddled the Juan de Fuca spreading center. Measurements at this site suggest temporal sound-speed variations of about 0.15 m/s, and range dependence due to the presence of localized hydrothermal vents. These measurements are also consistent with little or no deformation over a 2-year interval, in

agreement with other measurements, although the spatio-temporal sound speed changes create larger overall error in position determination.

8:45

2aUW2. Travel time and pulse dispersion in range-dependent waveguides. Kevin D. LePage (SACLANT Undersea Res. Ctr., VI. San Bartolomeo, 19138 La Spezia, Italy)

The evolution of time series through range-dependent ocean media is of substantial interest for a variety of reasons. Here the group speed and pulse dispersion of range-dependent modes is obtained under the narrow-band approximation. The results for range-dependent group speed show that in most cases the adiabatic approximation fails to capture even the most fundamental characteristics of individual mode travel time in range-dependent environments. Instead it is seen the group speed is strongly controlled by the balance of amplitudes in the modes and the way in which this balance is changed by bathymetry and sound speed fluctuations. In up-slope situations, as evanescent modes are locally energized by the cut-off and penetration of propagating modes, the local group speed of these modes becomes effectively equivalent to the group speed of the cutoff mode. In down-slope situations, where the adiabatic approximation fails entirely to assign a group speed to newly propagating modes, the current

approach gives a physically justified basis for assigning a value. Time series computed using the new range-dependent narrow-band approximation are compared to Fourier synthesized results for a variety of scenarios.

9:00

2aUW3. Sound propagation through a complex bubble-filled medium using a parabolic equation method. Trudy L. Philip (Grad. Prog. in Acoust., Appl. Res. Lab, The Penn State Univ., State College, PA 16804) and David L. Bradley (The Penn State Univ., State College, PA 16804)

A 2D parabolic equation (PE) code is used to predict sound propagation through a complex bubble field. The wake of a surface ship is the basis of the bubble distribution that is modeled as a simplistic geometric structure. Propeller motion and ship hydrodynamics create a dynamic field and the modeled simplistic structure is an estimation of this field as a “snapshot” in time. The complex features of the bubble field impact the sound field by variations in the sound velocity, density, and attenuation—all contributing loss mechanisms from the ship wake. A PE is used so that the environment can be changed to follow the spatial variation of the wake. Predictions are made for two different representations of the wake at three time frames, or distances from the source ship. For each location, the distribution of the bubbles is significantly different. Propagation is predicted for three pathways through the wake: down the wake (toward or away from the ship), across the wake (perpendicular to), and diagonally across the wake. [Work supported by ONR.]

9:15

2aUW4. Measurements of three-dimensional sound fields over different bottom types. Stewart Glegg and Joe Cichock (Ctr. for Acoust. and Vib., Dept. of Ocean Eng., Florida Atlantic Univ., Boca Raton, FL 33431)

This paper will describe measurements of three-dimensional underwater sound propagation over a sloping bottom. For this study a laboratory scale model was used with a typical slope of 2 deg. In the first series of measurements the sediment was modeled using concrete, providing a rigid lower boundary. In the second series, the concrete was coated with an epoxy layer, giving a softer lower boundary. The sound fields in these two cases showed significant differences that will be illustrated. Earlier work has shown that the sound field over the rigid bottom included interference effects caused by bathymetric refraction. The results for the softer lower boundary did not show these features. The measurements will also be compared with $N \times 2D$ PE calculations. [Work supported by ONR.]

9:30

2aUW5. Travel time bias in 2D modeling of 3D sound propagation. Oleg A. Godin (NOAA/Environ. Technol. Lab., 325 Broadway, Boulder, CO 80303, ogodin@etl.noaa.gov)

Limitations of the available mathematical models and/or computational resources often require invoking the so-called uncoupled azimuth approximation and substituting the 3D problem of sound propagation in horizontally inhomogeneous ocean by a set of 2D problems. In this paper, it is demonstrated that such approximation leads to a systematic error in ray and mode travel times. True travel times of ray arrivals as well as phases of cw adiabatic normal modes are *less* than those predicted within the uncoupled azimuth approximation. The travel time bias and its dependence on environmental gradients and propagation range are studied for

random and deterministic inhomogeneities and evaluated under conditions of recent experiments in deep and shallow water. A simple criterion is suggested that allows one to assess applicability of the uncoupled-azimuth approximation for given environmental conditions and propagation geometry, without performing 3D propagation modeling. Moreover, explicit corrections to modal travel times and amplitudes due to horizontal refraction are derived in terms of range-averaged cross-range environmental gradients that enable a significant extension of the domain of validity of the 2D modeling of the acoustic fields in the horizontally inhomogeneous waveguide. [Work supported by NRC.]

9:45

2aUW6. Coupled perturbed modes. Chris J. Higham, Chris T. Tindle, and Lin M. O'Driscoll (Phys. Dept., Univ. of Auckland, Bag 92019, Auckland, New Zealand, cjh@phy.auckland.ac.nz)

Conventional coupled modes are combined with perturbation theory to allow fast range-dependent normal mode calculations in deep water. The mode-coupling method follows R. B. Evan's two-way coupled mode formulation [R. B. Evans, J. Acoust. Soc. Am. **74**, 188–195 (1983)], where conservation of pressure and particle displacement along the z axis yields an algorithm to couple the modes as sound propagates in range. The conventional approach requires calculation of a complete set of normal modes at each step in range and is computationally intensive. Here perturbation theory is used to find the normal modes and wave numbers of each step in terms of linear combinations of the modes of the previous step. The process can be repeated for many steps and allows rapid calculation to long ranges. The method of *coupled perturbed modes* is applied to benchmark and canonical problems.

10:00–10:15 Break

10:15

2aUW7. On ray trajectory instability in a range-dependent waveguide. George Zaslavsky (Courant Inst., 251 Mercer St., New York, NY 10012-1185) and Anatoly L. Virovlyansky (Inst. of Appl. Phys., Nizhny Novgorod, Russia)

The sensitivity of ray trajectories to initial conditions in a deep ocean environment, including internal wave-induced inhomogeneities, is considered. It is known from a numerical simulation [Simmen *et al.*, J. Acoust. Soc. Am. **102**, 239–255 (1997)] that for a given realization of inhomogeneities there is a strong dependence of ray sensitivity on average (unperturbed) sound-speed profile. So, there are “stable” and “unstable” waveguides. Besides, for a given perturbed profile there are “stable” (typically steep) and “unstable” (axial) ray trajectories. We discuss a mechanism of ray instability due to overlapping of nonlinear ray-medium resonances. In an environment with a periodic range-dependence, this mechanism plays a crucial role in the emergence of ray chaos. Approximating a real nonperiodic sound-speed perturbation by one with a finite number of spatial frequencies we study overlapping of ray-medium resonances. This approach allows one to get a quantitative description showing how the sensitivity to initial conditions depends on the propagation conditions. [Work supported by U.S. Navy, Grant N00014-97-1-0426.]

2aUW8. Chaotic ray motion manifestation in mode amplitude range dependencies. Anatoly L. Virovlyansky (Inst. of Appl. Phys., 46 Ulyanov St., Nizhny Novgorod, Russia) and George Zaslavsky (Courant Inst., New York, NY 10012-1185)

An analytical ray-based description of normal mode amplitude range variations has been obtained by projecting the wave field taken in the geometrical optics approximation onto eigenfunctions of the boundary value problem. The mode amplitude at the given range is presented as a sum of contributions from several rays (which can be called eigenrays for the mode) with values of action variables equal to that of the mode. Under the condition of ray chaos, the number of rays contributing to a particular mode should grow exponentially with range, leading to complicated range variations of modal structure. This approach provides a convenient way to study how chaotic behavior of ray trajectories reveals itself in a range dependence of mode amplitudes, and, hence, in wave phenomena. An analog to nonlinear ray-medium resonance for normal modes has been investigated and the impact of this phenomenon on modal structure is discussed. It is argued that overlapping of different mode-medium resonances causes a complicated range dependence of mode amplitude in almost the same manner as the overlapping of ray-medium resonances leads to ray chaos. [Work supported by U.S. Navy, Grant N00014-97-1-0426.]

2aUW9. Ray theory and waveforms at caustics. Chris T. Tindle (Phys. Dept., Univ. of Auckland, Bag 92019, Auckland, New Zealand, c.tindle@auckland.ac.nz)

Waveforms for deep-water pulse propagation to a vertical array are calculated using a "measurement front" which is a continuous curve of arrival time as a function of depth. Each arriving pulse has a phase and amplitude determined by interpolation of the measurement front. The expressions for the phase and amplitude are derived directly from the wave equation using WKB expressions for the vertical dependence of the wave field. The measurement front contains caustics and the front is continued into the shadow using complex rays. The acoustic field on both sides of a caustic is calculated in terms of Airy functions. Detailed agreement with reference solutions is obtained.

2aUW10. Analysis of time-reversing array performance for a broadband pulse propagation in the presence of ambient noise. David R. Dowling and Sunny R. Khosla (2019 Auto Lab., Univ. of Michigan, Ann Arbor, MI 48109)

The properties of time-reversing arrays (TRAs) are being investigated for both active and passive underwater applications. Previous computational and experimental investigations have shown that automatic spatial and temporal pulse compression are possible with a TRA even when the acoustic environment is unknown. It has also been found that ambient or electronic noise limits the narrow-band performance of TRAs. In this presentation, the results from a computational study of the performance of a TRA that receives and transmits broadband pulses in a noisy environment will be shown. Pulsed-signal retrofocusing is simulated in noisy free-space and sound-channel environments via Monte Carlo techniques. The noise field is treated as white in the signal band and omnidirectional in space. The sound-channel simulations are done with a Fourier superposition of calculations from a parabolic equation code (RAM by M.D. Collins). Theoretical scaling laws for the probability of a retrofocus occurring as the

noise level increases are also presented. At the time and place of the retrofocus, results from both approaches suggest that pulse compression allows broadband TRAs to outperform narrow-band TRAs in noisy environments. [Work sponsored by ONR, Ocean Acoustics.]

2aUW11. Computed horizontal time-reversing array retrofocusing in a dynamic shallow-water sound channel. Michael R. Dungan and David R. Dowling (Dept. of Mech. Eng. and Appl. Mech., Univ. of Michigan, Ann Arbor, MI 48109, drd@engin.umich.edu)

A time-reversing array (TRA) can retrofocus acoustic energy to the location of its original source even when the acoustic environment is complex and unknown. Previous computational and experimental studies have utilized a vertical TRA to determine performance in shallow-water sound channels. However, other TRA deployment configuration possibilities exist. Here, results, based on narrow-band computations made with a parabolic equation code (RAM by M.D. Collins), are presented for a horizontal array in various orientations with respect to the source. In addition to characterizing the influence of range, acoustic frequency, and horizontal-TRA orientation on the quality of the retrofocus, the sensitivity of horizontal-TRA retrofocusing to internal wave dynamics in the water column of a shallow-water sound channel is also shown. The internal wave model is based on a superposition of linear internal waves whose temporal spectrum has been matched to spectra measured during the SWARM '95 experiment. Some performance differences are noted between end-fire and broadside array-source orientations. If time allows, results from fully 3D simulations will be shown. [Work supported by ONR, Ocean Acoustics.]

2aUW12. Dynamic coupling of primitive equation soliton models and acoustic propagation models via acoustic mode analysis. Stanley A. Chin-Bing, David B. King, Alex C. Warn-Varnas, and Robert A. Zingarelli (Naval Res. Lab., Stennis Space Center, MS 39529-5004, chinbing@nrlssc.navy.mil)

Simulation studies of acoustic mode conversions resulting from underwater acoustic fields propagating through solitons in the Strait of Messina have indicated a novel way of coupling oceanographic models with ocean acoustic models using dynamic feedback from the acoustic analysis. A weakly nonhydrostatic primitive equation model was used to generate simulations of oceanographic soliton fields in the Strait of Messina. The soliton fields were used to study acoustic mode conversions as the acoustic field passed through the solitons and propagated upslope. Conversion into modes having higher bottom loss have shown resonance effects similar to that observed in the Yellow Sea. Oceanographic parameters were adjusted so that amplitude and phase of the first two wavelengths of the simulated soliton fields were in agreement with measured soliton data. However, identical acoustic signals propagated through the simulated and measured soliton fields produced different acoustic mode structures. Oceanographic parameter values were varied until the differences in acoustic predictions were minimized. Thus, a feedback approach was used to dynamically couple the oceanographic and acoustic models. Initial results and the suitability of a two-layer oceanographic model for detailed acoustic studies will be discussed. [Work supported by ONR/NRL and by a High Performance Computing DoD grant.]

Session 2pAA**Architectural Acoustics: Stadium and Arena Acoustics**

Jack E. Randorff, Chair

*Randorff and Associates, 11 West Canyon View Drive, Ransom Canyon, Texas 79366-2206***Chair's Introduction—1:00*****Invited Papers*****1:05****2pAA1. Sound system and acoustic design for arena-sized performance venues.** Craig Janssen (Acoustic Dimensions, 15505 Wright Brothers Dr., Addison, TX 75001, cjanssen@acousticdimensions.com)

The rapid rise of “mega” churches in the USA over the last decade has led to the construction of numerous arena-sized church auditoriums seating over 7000 people. The programmatic usage of these spaces generates inherent design conflicts between the need for acoustics to support congregational singing, while at the same time supporting high-power concert quality sound reinforcement. These design challenges have precipitated the development of new design approaches for both room acoustics and the sound reinforcement system. A case study of a 7200-seat church will be presented to illustrate these core design approaches and techniques. The presentation will include discussion of specular versus reverberant field intelligibility predictions, the usefulness or lack of usefulness of early reflections, loudspeaker system design options, creation of a useful left, center, right loudspeaker system, and other related issues.

1:30**2pAA2. Managing high-frequency losses in large venue sound-reinforcement systems.** Stephen J. Siegel (Eastern Acoust. Works, One Main St., Whitinsville, MA 01588, stephen@eaw.com)

The myriad of challenges associated with sound-reinforcement systems for arenas, stadiums, and outdoor festivals presents significant obstacles to the audio professional. One of the most significant factors, which can affect both speech intelligibility and tonal accuracy of sound reproduction, is loss of projection due to air absorption. On the surface it might seem that this is a very high-frequency phenomenon which would only be a concern for certain types of musical program. Closer examination reveals that it is, in fact, a significant factor at frequencies in the vocal range. The types of losses encountered in the arena, stadium, and festival configurations are examined, and their impact on both speech and music program is discussed. Methods to implement large-scale sound systems to minimize these effects are presented, with several applications shown as examples.

1:55**2pAA3. Acoustics of political conventions—A review.** Jack E. Randorff (Randorff and Assoc., 11 W. Canyon View Dr., Ransom Canyon, TX 79366)

The acoustical challenges encountered in recent political conventions is reviewed. Examples are drawn from The 1992 Republican National Convention in the Astrodome in Houston, The 1996 Republican National Convention in the San Diego Convention Center, and the siting study prepared for The 2000 Republican National Convention scheduled for Philadelphia. The acoustical needs of the delegates and the broadcast audience are discussed. The technical performance requirements of convention sound reinforcement and media network broadcast feed are outlined. The necessary technical and performance tradeoffs are enumerated with respect to the physical constraints, schedule requirements, budget limitations, and planning committee expectations.

2:20–2:35 Break**2:35****2pAA4. RT60 and speech intelligibility comparisons in stadia with movable roof systems.** David E. Marsh (Pelton Marsh Kinsella, 1420 W. Mockingbird Ln., Ste. 400, Dallas, TX 75247)

Open-air stadia are usually less reverberant than enclosed ones and they provide the possibility of natural turf for sporting events, which is generally preferred over synthetic turf. However, enclosed stadia have the financial advantage of year-round event bookings with heating or air-conditioning as appropriate. In recent years, a number of stadia with movable roof systems have been built, providing the best of both worlds. Reverberation time (RT60) measurements and speech-intelligibility measurements were made in some of these facilities and the results were compared for the two conditions of “roof open” and “roof closed.”

2pAA5. Loudspeaker directivity as a function of frequency, and the importance thereof to sound system performance in large and reverberant spaces. James Long (EVI Audio, 600 Cecil St., Buchanan, MI 49107, jlong2@compuserve.com)

Control of a loudspeaker's directivity, or coverage pattern, at high frequencies can be readily achieved with devices of modest dimensions. Extending directivity control into the vocal midrange and below requires devices of significantly larger size and possibly greater complexity. However, directivity control at lower frequencies can significantly increase musical clarity and vocal intelligibility in large and typically reverberant spaces such as stadiums and arenas. The basic physics behind achieving extended frequency directivity control, including the use of signal processing, will be outlined. Typical existing, practical examples will be presented. A two-channel computer auralization comparing transducers in a reverberant space with and without extended frequency directivity control will be demonstrated.

Contributed Paper

3:25

2pAA6. Acoustics and sound-system design for the National Ice Centre, Nottingham, UK. Peter Mapp (Peter Mapp & Assoc., Colchester CO3 4JZ, UK, petermapp@btinternet.com)

The National Ice Centre, Nottingham is a unique indoor arena within the UK, as it will house two full-size ice pads both capable of hosting national and international ice-skating competitions. One of the ice pads will be permanently available for both local community (public) and professional ice-skating activities. The other pad, located in the main arena, can be covered over and the arena used to seat audiences of up to 10 000

spectators for pop concerts, major artist performances, sports events (e.g., tennis or boxing), and large conferences/meetings/exhibitions/conventions. The main arena will also host national and international ice dance and skating competitions. With such a wide range of activities, considerable attention has been paid to both the acoustics of the place and the installed sound systems. The paper describes both the acoustic treatment of the arena and the design philosophy and testing of the complex sound systems. The sound-system design is unusual for such a venue in that two separate systems are installed to cater to both the performance sound and voice alarm (fire alarm) needs without compromising the strict requirements relating to public safety in a building of this nature and size.

WEDNESDAY AFTERNOON, 31 MAY 2000

GEORGIAN ROOM, 1:30 TO 4:50 P.M.

Session 2pAB

Animal Bioacoustics and Signal Processing in Acoustics: Auditory Temporal Processing in Animals

Cynthia F. Moss, Cochair

Department of Psychology, University of Maryland, College Park, Maryland 20742

Whitlow W. L. Au, Cochair

Hawaii Institute of Marine Biology, P.O. Box 1106, Kailua, Hawaii 96734

Chair's Introduction—1:30

Invited Papers

1:35

2pAB1. Estimating echo-delay acuity from timing of neural responses in echolocating bats. James A. Simmons, Mark I. Sanderson, and Kyler M. Eastman (Dept. of Neurosci., Brown Univ., Providence, RI)

Big brown bats resolve delays of overlapping FM biosonar echoes down to 2–4 μ s. They must use a time-frequency representation because integration time is $\sim 350 \mu$ s. The bat may have $\sim 25\,000$ auditory-nerve fibers from ~ 1000 inner hair cells tuned to frequencies from 10 to 100 kHz. If auditory-nerve latency jitters of 50–100 μ s are averaged across frequencies, estimated timing accuracy could be as good as $\sim 0.5 \mu$ s. Monaural timing pathways to the inferior colliculus have jitters no worse than the auditory nerve and also could convey a timing accuracy as good as $\sim 0.5 \mu$ s. In the inferior colliculus, frequency-domain responses to interference notches register delay separations down to 5–6 μ s for two-point resolution. Time-domain latency jitter for tone bursts increases sharply in the IC, but over half of the cells respond to FM pulses and echoes in synthetic pursuit sequences with latency jitters of only 100–500 μ s. This ~ 2 – $5\times$ increase in intrinsic jitter in the IC is accompanied by a disproportionate increase in the number of active cells, thus preserving a physiologically plausible acuity of ~ 0.5 – 1μ s. Moreover, if the whole time-frequency representation in the IC is taken into account, delay changes of ~ 0.05 – 0.1μ s might be detectable.

2:00

2pAB2. Pitch perception versus target ranging and identification in echolocating bats. Shihab Shamma (Dept. of Elec. Eng., Univ. of Maryland, College Park, MD 20742)

When echolocating, the brown bat emits a train of FM pulses while listening to the echoes from the target. It then estimates the target range and identity based on the time intervals and interactions between the pulses and their echoes. These remarkable abilities are often thought to arise from unique neural computations and networks, since most other mammals do not echolocate. Furthermore, related physiological data are most commonly interpreted as indicative of temporal algorithms involving delay-tuned cells, auto-

correlators, and other temporally accurate structures. In this talk, we shall argue that target ranging and target identification are in fact computationally identical to pitch estimation in other mammals; specifically, they are analogous to the perception of repetition pitch and of periodicity (or musical) pitch, respectively. We shall also discuss how these estimates can be derived from so-called spectrally based (rather than temporally based) algorithms, invoking harmonic series templates and instantaneous coincidences among cochlear outputs that do not require arrays of delay-tuned cells. These results offer alternative interpretations for the physiological evidence from bats and suggest close parallels between them and other nonecholocating mammals.

2:25

2pAB3. Autocorrelation model for pitch processing in the frog's auditory nerve. Andrea M. Simmons (Dept. of Psych., Brown Univ., Providence, RI 02912)

Bullfrogs' reactions to simulated vocalizations are sensitive to disruption of the 100-Hz 1st harmonic periodicity caused by the presence of inharmonic frequency components or changes in component phase. This behavioral preference for a specific range of periods indicates that representation of periodicity "pitch" may be the basis for evaluation of conspecific calls. Anuran auditory-nerve fibers phase-lock to the 1st and sometimes the 2nd or 3rd harmonic of multiple harmonic signals. Interval histograms of fiber responses reflect the complex pitch of both harmonic and inharmonic stimuli. We developed a pitch model that estimates pitch-period autocorrelation functions from all-order interval histograms across frequency channels by implementing model neurons designed to replicate the frog's auditory nerve. Histograms are aligned in time across channels before being combined into a single pitch image ("strobing" in the Patterson pulse-ribbon model; "dechirping" in bat sonar target-ranging). The resultant output yields a sharp pitchlike image, consistent with both behavioral and physiological results. Robust 1st-harmonic pitch images are produced for multiple harmonic signals, and even two-component sounds require only a few parallel fibers to yield a robust 1st-harmonic pitch.

2:50

2pAB4. Characterization of time-varying responses to dynamic broadband spectra in primary auditory cortex. Jonathan Z. Simon, Didier A. Depireux, David J. Klein, and Shihab A. Shamma (Inst. for Systems Res., Univ. of Maryland, College Park, MD 20742)

The neural encoding of time-varying, spectrally rich sounds (such as speech and music) in primary auditory cortex (AI) is still poorly understood. In this research, the time-varying responses of units in AI to dynamically and spectrally modulated noise are characterized. The most elementary form of these stimuli, a "moving ripple," has a spectro-temporal envelope which is modulated sinusoidally in both spectral and temporal dimensions. More complex sounds have spectro-temporal envelopes decomposable into multiple sinusoidal components. Responses to both simple and complex stimuli allow independent measurements of the Spectro-Temporal Response Field (STRF). The STRF provides both a quantitative and intuitive measure of how a unit encodes dynamic sound features into its resultant time-varying spike train. The STRF is a linear response characteristic, but for a reasonably broad class of sounds, the STRF predicts the response of units to novel stimuli. Constraints on the structure of STRFs, such as whether they are separable (the vast majority appear to be quadrant-separable), are also constraints on the connectivity of the neuronal inputs. [Supported by grants from ONR, NIDCD, and NSF.]

3:15–3:35 Break

3:35

2pAB5. Cellular models of coincidence detection. Jonathan Z. Simon (Inst. for Systems Res., UMCP, College Park, MD 20742), Suchitra Parameshwaran, and Catherine E. Carr (UMCP, College Park, MD 20742)

The avian brainstem contains a circuit for detection of interaural time differences (ITDs). Bilateral projections from the cochlear nucleus magnocellularis (NM) form maps of ITD in the nucleus laminaris (NL). A computational model of a NL neuron investigates the cell's ability to detect coincidences between ipsilateral and contralateral inputs. The model, constructed in the NEURON programming environment, is biophysically detailed and includes recent information on the distribution of high threshold potassium conductances in NL neurons. The role of dendritic processing is explored and emphasized, including the effects of dendrite size on postsynaptic potentials, parallel (multiple) dendritic processing, and the benefits of having dendrites at all. The model makes several predictions, some of which confirm those of a previous model [H. Agmon-Snir, C. E. Carr, and J. Rinzel, *Nature* **393**, 268–272 (1998)]. Active potassium channels greatly increase the ability to detect coincidences at high frequencies. Rate-coded output is much more robust than vector-strength-coded output at distinguishing coincidences from partial coincidences. There is an optimal number of dendrites for a given stimulus frequency, dendritic length, and number of synaptic inputs. The most efficient dendritic length decreases with best frequency.

4:00

2pAB6. Temporal resolution in birds: Discriminating temporal fine structure in harmonic complexes. Robert J. Dooling (Dept. of Psych., Univ. of Maryland, College Park, MD 20742), Marjorie R. Leek (Walter Reed Army Medical Ctr., Washington, DC 20307), Otto Gleich (Univ. of Regensburg, 93042 Regensburg, Germany), Micheal L. Dent, and Jeff M. Cucina (Univ. of Maryland, College Park, MD 20742)

Recent behavioral experiments on auditory masking and waveform discriminations using harmonic complexes as stimuli have shown significant differences in performance between birds and humans. In humans, but not in birds, Schroeder-phase harmonic complexes constructed with monotonically increasing (positive Schroeder) or decreasing (negative Schroeder) component phases are differentially effective as maskers, even though they have identical temporal envelopes and long-term spectra. The similarity in masking effectiveness of these harmonic complexes notwithstanding, birds can easily discriminate the fine-structure differences between the positive and negative Schroeder waveforms even with harmonic periods as short as 2 ms. By contrast, humans are unable

to make these discriminations unless the periods are greater than about 5–6 ms. These behavioral discrimination and masking results have now been replicated by measuring evoked-potential responses in budgerigars. Amplitude and waveform shapes of the compound action potentials reflect the behavioral discriminations between the two Schroeder stimuli, but cochlear microphonics do not. The similarity in results between the behavioral and the electrophysiological studies, using the same stimuli, suggest differences in both temporal and spectral processing in avian versus mammalian cochleas. [Work supported by NIH R01 DC00198, DC00626, and NRSA DC00046.]

4:25

2pAB7. Auditory temporal computation in fish. John Crawford (Dept. of Psych., Univ. of Pennsylvania, 3815 Walnut St., Philadelphia, PA 19104, jud@psych.upenn.edu)

Fishes offer valuable opportunities for investigating auditory temporal computation. They use simple communication sounds. The ears and central auditory pathways are similar to those of other vertebrates, but the periphery is not specialized for frequency analysis as it is in mammals, and acoustic temporal analysis predominates. In the fish *Pollimyrus*, a small gas-filled bladder is coupled to each sacculus and translates acoustic pressure to motion of the hair cells. Behaviorally, these fish are sensitive to small differences in the temporal structure of sounds. Neurophysiological recordings from the auditory nerve show that sounds are faithfully encoded by the intervals between phase-locked spikes. This representation is transformed as it ascends the pathway, and is used in neural computations. In the auditory midbrain neurons are activated by particular temporal patterns. The selectivity of the auditory feature detectors results from neural computations that use synchronized trains of spikes as input. Computational modeling supports the hypothesis that a simple network, with three neurons and inhibitory gating, underlies the emergence of temporal tuning in the auditory midbrain. The slow kinetics of inhibition and postinhibitory rebound provide a basis for tuning to the long time intervals (10–50 ms) that are important in communication sounds. [Work supported by NIH Grant No. R01 DC01252.]

WEDNESDAY AFTERNOON, 31 MAY 2000

FRENCH ROOM, 1:45 TO 4:45 P.M.

Session 2pBB

2p WED. PM

Biomedical Ultrasound/Bioresponse to Vibration: Ultrasound for Disease Treatment and Diagnosis II

Pierre D. Mourad, Chair

Applied Physics Laboratories, University of Washington, Seattle, Washington 98195

Contributed Papers

1:45

2pBB1. Microbubbles and medical ultrasound: Translation dynamics and optimal driving. Anil J. Reddy and Andrew J. Szeri (Dept. of Mech. Eng., Univ. of California, Berkeley, CA 94720-1740)

If bubbles are to be used in a medical setting, for example as contrast agents in the imaging of tissue, or in the delivery of a drug to a blood clot or tumor, it is necessary to characterize the translational motion of a bubble subjected to acoustic excitation in the form of ultrasound. Further, because bubbles remain undissolved in a fluid for only a short period of time, it may be necessary to transport a bubble to its target location as quickly as possible. The acoustic radiation force, known as the primary Bjerknes force, and the recently derived expression [J. Magnaudet and D. Legendre, *Phys. Fluids* **10**, 550–554 (1998)] for drag force acting on a bubble with time-dependent radius, are taken into account in developing an equation governing the translation dynamics of a bubble. This is used together with the Rayleigh–Plesset equation for volume oscillations and optimal control theory to design the ideal acoustic excitation for transporting a bubble to a particular location, or for maintaining the position of a bubble at the target location despite the adverse flow of the surrounding fluid, as in the case of a bloodstream. [Work supported by NSF.]

2:00

2pBB2. Low-frequency response of the submerged human lung. James S. Martin, Peter H. Rogers (School of Mech. Eng., Georgia Inst. of Technol., Atlanta, GA 30332-0405, james.martin@me.gatech.edu), Edward A. Cudahy, and Eric L. Hanson (Naval Submarine Medical Res. Lab., Groton, CT 06349-5900)

A study was undertaken to determine the response of the human lung to low-frequency underwater sound in the range of 20 to 500 Hz. Experiments were conducted in a 1100 gal tank inside a hyperbaric chamber. Lung responses were measured by three independent techniques: a hydro-

phone close to the chest, an accelerometer attached to the chest, and a noninvasive ultrasound measurement of the lung surface. Ten subjects were tested. Each subject was tested twice at ambient pressures of 0, 10, 60, and 120 ft of sea water (FSW). The dominant feature of the response was a strong resonance that occurred, on average, at 39 Hz at surface pressure and 71 Hz at a depth of 120 FSW. This appears to be the fundamental resonance of the lungs. The depth dependence indicates that lung stiffness is dominated by gas in the lungs with a small contribution from the chest wall. Resonance Q's were in the range of 5 to 7. There is evidence in the data of other low-frequency resonances that can be attributed to intestinal gas. Other than these, the lung acceleration response to incident pressure above resonance appears to be flat. [Work supported by ONR.]

2:15

2pBB3. Parameter dependence of 20-MHz passive inertial cavitation detector output. Alyssa Bonnoit and E. Carr Everbach (Dept. of Eng., Swarthmore College, Swarthmore, PA 19081)

Passive acoustic monitoring of inertial cavitation emissions has been a useful tool to quantify cavitation activity *in vivo* and *in vitro*. In particular, the time-averaged root-mean-squared voltage of a narrow-band-filtered 20-MHz transducer output has correlated well with cellular bioeffects, when the insonifying transducer is in the low-megahertz range [E. C. Everbach *et al.*, *Ultrasound Med. Biol.* **23**(4), 619–624 (1997)]. A thorough study of parameters affecting this inertial detector output has not been undertaken, however. The results of such a study will be presented, showing time-averaged rms detector output as a function of insonifying frequency, pulse parameters, and other acoustic measures of inertial cavitation activity. Limitations of the method will be discussed.

2pBB4. Correlations between UCA-destruction-induced bioeffects and inertial cavitation dose. Wen-Shiang Chen, Peter P. Chang, Thomas J. Matula, and Lawrence A. Crum (Appl. Phys. Lab., Univ. of Washington, 1013 NE 40th St., Seattle, WA 98105)

Certain therapeutic applications of ultrasound may be enhanced by the use of ultrasound contrast agents, whose controlled destruction may produce a desired bioeffect. It is therefore important to understand the mechanisms of microbubble destruction, and to quantify the relationship between bioeffects and the acoustical signature of these agents. In previous studies, we found two distinct thresholds for the ultrasonic destruction of shelled microbubbles: a fragmentation threshold (P1) at which the microbubble shell is disrupted, creating smaller bubbles, and an inertial cavitation [IC] threshold (P2) for sustained, vigorous IC activity. The inertial cavitation dose (ICD) was used to monitor continuous changes in IC activity. The samples, containing mixtures of Optison with either a buffer solution or whole blood, were insonified with a 1.1-MHz focused transducer (up to 4.4 MPa peak-to-peak). Below P2, the ICD increased slowly with increasing pressure amplitude; hemolysis accumulated more rapidly. Above P2, the ICD decreased. Increased hemolysis was observed between P1 and P2, and was associated with short bursts of IC activity. At fixed PRF, or fixed total "on" time (by adjustment of pulse length and PRF), a pulse length threshold existed for the ICD and hemolysis measurements. [Work supported by WTC and EKOS Corp.]

2pBB5. Sonoluminescence as an indicator of cell-membrane disruption by acoustic cavitation. Stephen Cochran (Dept. of Biomed. Eng., Georgia Inst. of Technol., 315 Ferst Dr., Atlanta, GA 30332) and Mark Prausnitz (Georgia Inst. of Technol., Atlanta, GA 30332-0100, mark.prausnitz@che.gatech.edu)

Acoustic cavitation has been shown to reversibly disrupt cell membranes and thereby transport molecules into viable cells. The hypothesis was that the light emitted during inertial cavitation, i.e., sonoluminescence, correlates with these acoustic bioeffects. In this study, sonoluminescence was measured using a photomultiplier tube mounted at the base of an opaque cavitation chamber. The bioeffects measured on DU 145 prostate cancer cells placed in the cavitation field included the number of molecules delivered into each cell and overall cell viability. A 24-kHz cylindrical piezoelectric transducer was used in these experiments to generate cavitation over a range of pressure levels, pulse lengths, and total exposure times. The results show that cells were reversibly disrupted, yielding extensive uptake of calcein by some cells, little uptake by other cells, and loss of viability for a third population of cells. Correlation with sonoluminescence indicated that light emission correlated with these bioeffects.

2pBB6. Control of transient disruptions in cell membranes by acoustic cavitation at 500 kHz. Hector Guzman, Daniel Nguyen, and Mark Prausnitz (Dept. of Chemical Eng., Georgia Inst. of Technol., Atlanta, GA 30332)

Ultrasound-mediated drug targeting and delivery is a nonchemical, nonviral, and noninvasive alternative for the transport of drugs and genes into cells. Delivery of molecules is performed by using ultrasound to disrupt cell membranes by a mechanism believed to involve cavitation. To develop efficient drug and gene delivery methods based on ultrasound-mediated disruption of cell membranes, it is critical to understand the dependence of molecular uptake and cell viability on the acoustic parameters that govern the complex cavitation activity. This study examined these effects in cell suspensions using two model cell lines (aortic smooth muscle and DU145 prostate cancer cells) and two model compounds (calcein and bovine serum albumin) at 500-kHz ultrasound in the presence of Optison contrast agent. The acoustic parameters varied in this study were:

pressure, pulse length, and total exposure time. Cell membrane disruption increased with both pressure and exposure time, but was not influenced by pulse length over the range examined. A window of useful operating conditions was determined for delivering large numbers of molecules into viable cells. Correlations between membrane disruption and the acoustic spectra of cavitation were observed to support the proposed cavitation mechanism. [Work supported by Whitaker Foundation, NIH, and NSF.]

2pBB7. Ultrasound-guided thermo therapy. Robert M. Lemor and Bernhard V. Kleffner (Fraunhofer Inst. for Biomed. Technol., Dept. of Ultrasound, Ensheimer Str. 48, 66386 St. Ingbert, Germany)

Thermal therapies for cancer treatment have achieved a high level of interest in the medical society, especially for treatment of liver metastases. Laser-induced thermo therapy (LITT) is one of the most promising techniques for minimally invasive destruction of tumors through coagulation. The precision and efficiency of this treatment could be markedly improved by noninvasive monitoring of either temperature distribution in the tissue or by direct demarcation of the destructed zone. The purpose of this report is to demonstrate that parametric ultrasound imaging provides an efficient method for localizing the coagulated area during the treatment. An on-line time-of-flight velocimetry for temperature mapping as well as a mapping of the structural changes through measurement of changes in the attenuation coefficient are presented. The coagulation of porcine liver *in vitro* and *in vivo* has been performed by Nd:YAG laser radiation while access to the radio-frequency signals of a conventional ultrasound system served for collection of acoustical parameters from the treated zone. The introduced system is capable of acquiring 12 B-scans/sec with a mean spatial resolution of 1 mm. Quantitative parameters for the control of LITT treatments are derived.

2pBB8. 3D full wave ultrasonic field and temperature simulations in biological tissue containing a blood vessel. Francesco P. Curra, Pierre D. Mourad, Lawrence A. Crum (Appl. Phys. Lab., Univ. of Washington, Seattle, WA), and Vera A. Khokhlova (Moscow State Univ., Moscow, Russia)

In order to simulate ultrasound propagation and subsequent thermal effects in biological media in which blood vessels and other structures may be present, a three-dimensional model has been developed that eliminates the need for symmetry constraints. The model is based on the coupled solution of the full wave nonlinear equation of sound in a lossy medium and the bioheat equation obtained by a pseudospectral finite-difference method in the time domain. It includes nonlinear sound propagation, an arbitrary frequency power law for attenuation, and is capable of treating material inhomogeneities. Unlike other models based on parabolic approximations, it is not restricted to near-axis solutions and can account for reflections and backscattered fields. The program was used to simulate the application of high-intensity focused ultrasound (HIFU) in liver with a blood vessel placed perpendicular to the axis of the transducer and near the focus. This approach follows recent work by the authors [Curra *et al.*, IEEE Trans. Ultrason. Ferroelectr., Freq. Control (in press)]. Simulations are presented for different levels of driving pressure, sound nonlinearities, exposure times, and the relative position between the transducer focus and the blood vessel. [Work supported in part by DARPA and the CRDF Cooperative Grants Program.]

2pBB9. Ultrasound-induced transskull lesions using a large-area phased array. Greg Clement, Tonia Giesecke, and Kullervo Hynynen (Brigham and Women's Hospital, Harvard Med. School, 75 Francis St., Boston, MA 02115, clement@bwh.harvard.edu)

A transducer array for transskull focused ultrasound surgery is constructed and tested. The transducer design is motivated by numeric and experimental studies, resulting in a hemispherical array optimized in frequency and number of elements. The hemisphere shape is incorporated to

maximize the penetration area on the skull surface, thus minimizing unwanted heating. It is determined from the studies that 64 elements are sufficient for correcting scattering and reflection caused by transskull propagation. A resonant frequency near 0.7 MHz is chosen for the array using numeric and experimental thermal gain measurements between the transducer focus and the skull surface. An air-backed PZT array is assembled and characterized. The array is then used to focus high-intensity ultrasound through an *ex vivo* human skull. Thermally induced lesions are produced through a human skull in fresh tissue placed at the ultrasound focus inside the skull.

4:15

2pBB10. Ultrasound accelerates the healing of damaged peripheral nerves *in vivo*. Pierre D. Mourad, Francesco Curra, Lawrence A. Crum (Appl. Phys. Lab., Univ. of Washington, Seattle, WA 98105, pierre@apl.washington.edu), Daniel A. Lazar, and Michel Kliot (Univ. of Washington, Seattle, WA 98195)

Peripheral nerve injury is a common clinical problem that often debilitates by producing significant motor and sensory deficits, as well as pain and other unpleasant sensations. Recovery is often slow. Patients with proximal injuries that require axons to regenerate over long distances to reach their target muscles or sensory receptors can take up to 2 years to recover function. Moreover, slow recovery can often be incomplete, due to the associated muscle atrophy, decrease in axonal regeneration along the

pathway of the degenerating nerve, and decay in the axonal receptivity of the target tissues. In this work, we show that low-intensity therapeutic ultrasound can accelerate functional recovery in an animal model of a completely crushed peripheral nerve.

4:30

2pBB11. Using ultrasound to assess change in corneal stiffness. Mardi C. Hastings (Biomed. Eng. Ctr., The Ohio State Univ., 1080 Carmack Rd., Columbus, OH 43210)

Corneas of six intact donor eye globes from five different species were interrogated with continuous wave ultrasound (15 MHz) while being excited by audible sound waves. The corneal oscillation induced by the audible stimulus created a phase modulation in the received ultrasonic signal which was demodulated to obtain the amplitude and phase of the motion at discrete locations on the cornea as defined by the crossed beams of the ultrasonic transmitter and receiver. The complex stiffness parameters for each cornea were determined from a curve fit to the measured amplitude and phase of the transfer function relating the dynamic motion of the cornea to the local acoustic field. Data were recorded at three different intraocular pressures. Increasing intraocular pressure resulted in increasing corneal stiffness to produce measurable shifts in resonance peaks observed in the dynamic response. Noninvasive assessment of corneal stiffness as demonstrated in this study has potential application in diagnosing and treating glaucoma, and monitoring healing after injury or cornea surgery.

WEDNESDAY AFTERNOON, 31 MAY 2000

CANADIAN ROOM, 12:55 TO 3:30 P.M.

Session 2pEA

Engineering Acoustics and Physical Acoustics: Combustion Acoustics

Steven L. Garrett, Chair

Graduate Program in Acoustics, The Pennsylvania State University, P.O. Box 30, State College, Pennsylvania 16804

Chair's Introduction—12:55

Invited Papers

1:00

2pEA1. Combustion and acoustic waves—A damaging combination. Ann Dowling (Dept. of Eng., Univ. of Cambridge, Trumpington St., Cambridge CB2 1PZ, UK, apd1@eng.cam.ac.uk)

Instability can occur whenever combustion takes place within an acoustic resonator—essentially because the unsteady combustion generates sound, while the resulting acoustic waves perturb the combustion still further. Many combustors have been found to be susceptible to self-excited oscillations: currently the occurrence of instability in lean premixed combustion systems is proving to be a major technological challenge. By burning lean premixed and prevaporized, it is possible to reduce chemical emissions significantly. Unfortunately, these are just the conditions that make a flame particularly sensitive to acoustic disturbances. Coupling between the rate of heat release and the sound waves can lead to self-excited oscillations, and the resulting pressure oscillations and/or enhanced heat transfer may cause structural damage. Modeling of combustion instability requires application of ideas from acoustics, combustion, and systems engineering. A wave analysis will be presented which is useful to describe sound generation in complex geometries with mean flow. The complementary roles of low-order models for unsteady combustion and of computational fluid dynamics (CFD) will be discussed. Particular emphasis will be placed on the effective integration of CFD and acoustic modeling. These solution methods will be illustrated by investigation of oscillations in generic premixed ducted flames and in aeroengine combustors.

1:30

2pEA2. A review of active control of combustion instabilities. Ben T. Zinn (School of Aerosp. Eng., Georgia Inst. of Technol., Atlanta, GA 30332-0150, ben.zinn@aerospace.gatech.edu)

This paper presents an overview of active control of combustion instabilities. First, the paper will use a simple model to present the rationale that motivated the development of active control systems (ACS) for unstable combustors. Next, the design and performance of earlier ACS will be discussed to identify control system components and control approaches that offer the most promise for

practical applications. The bulk of the paper will discuss active control research at Georgia Tech, conducted by the author and colleagues. It will describe the architecture of the developed ACS, adaptive and nonadaptive active control approaches, and the performance of critical ACS components. Emphasis will be placed upon the performance of the observer that determines the state of the unstable combustor in real time and the fuel injector actuator that can modulate the injection rate of a fuel stream over a wide frequency range (e.g., 0–2000 Hz). This will be followed by descriptions of the performance of this ACS under different modes of operation when applied to different combustors. These results will show that the developed ACS can rapidly damp combustion instabilities whose characteristics vary with time and are not known *a priori*. The paper will close with a discussion of needed research and development efforts.

2:00

2pEA3. Pulse combustion. Jechiel I. Jagoda (School of Aerosp. Eng., Georgia Inst. of Technol., Atlanta, GA 30332-0150)

Pulse combustion has been gaining increased interest because of its potential for higher combustion efficiency, greater combustion intensity, and lower pollutant emissions. Unsteady combustion causes increased mass, momentum, and heat transfer. As a result, reactants mix faster, heat release is accelerated, and heat transfer is enhanced in unsteady reacting flows. Many of these phenomena were discovered long ago by engineers looking for the cause of often detrimental combustion instabilities. Much more recently some of these enhanced transfer properties have been used to design efficient and compact pulse combustors. Although, to date, successful commercialization on a large scale has been limited to home heating units (e.g., the Lenox Pulse Furnace) highly efficient pulse spray dryers (Bepex Unison Dryer), pulse calciners, and pulse waste incinerators have been designed. Pulsations have also been applied to carbon black fluidized bed gasifiers. Not all these designs will become economically viable. However, the development of tunable pulse combustors that can be acoustically matched to the changing resonance frequency of these pulse processes have made many of them more promising. Recent findings that pulsation can enhance burning even in turbulent flows lend further encouragement to the developers of novel pulse combustion devices.

2:30–2:45 Break

2:45

2pEA4. Combustion instabilities in propulsion systems. Vigor Yang (Propulsion Eng. Res. Ctr., Penn State Univ., 104 Res. Bldg. East, University Park, PA 16802, vigor@psu.edu)

Periodic acoustic motions in combustion chambers, a phenomenon commonly known as combustion instability, were discovered in air-breathing and rocket engines at about the same time in the late 1930s. Since then, unstable oscillations have occurred in most, if not practically all, new development programs. Indeed, because of the high density of energy release in volume having relatively low losses, conditions normally favor excitation and sustenance of oscillations in any combustion chamber intended for a propulsion system. This presentation will provide an overview of combustion instabilities in three types of propulsion systems (i.e., solid rocket motors, liquid rocket engines, and air-breathing engines), with emphasis focused on the state-of-the-art understanding and future research needs. Various research issues in acoustics, fluid mechanics, and chemistry related to oscillatory combustion in practical systems will be discussed. Both passive and active control techniques will be covered. In particular, the application of contemporary CFD schemes, approximate analytical methods, and experimental diagnostic tools to combustion instability studies will be addressed in detail.

Contributed Paper

3:15

2pEA5. Acoustocyclone filter concept for soot retention in automotive diesel engines. Thomas L. Hoffmann (Fraunhofer USA, 601 W. 20th St., Hialeah, FL 33010, tlh@kirkof.psu.edu)

Fine-particle soot emissions are an intrinsic problem of diesel-powered vehicles. A number of car manufacturers have already begun marketing diesel vehicles with particle emission control equipment in an effort to address the problem. While technical solutions exist to reduce the number of fine diesel exhaust particles, these systems tend to be bulky and poorly resistant to high mechanical stresses and high temperature operation. Fil-

ters tend to get blocked, causing unwanted pressure increases in the exhaust line and an eventual loss of engine power and fuel efficiency. The concept study presented here focuses on a compact particle retention device that will generate only a small pressure drop, while efficiently filtering fine and coarse particle exhaust emissions. Based on a technology developed for coal-fired power plants, the new filtration unit will make use of a two-stage collection system consisting of a conventional cyclone and an acoustic aerosol agglomerator. The approach will build on the concept of aerosol preconditioning to improve the retention efficiency of a conventional cyclone. This improvement will be achieved through initial agglomeration of the fine particles with larger soot particulates, and subsequent filtration of the preconditioned aerosol in a cyclone collector.

Session 2pMU

Musical Acoustics: Modeling and Perception of Musical Sound Sources II

Stephen E. McAdams, Chair

IRCAM, 1 Place Igor Stravinsky, Paris F-75004, France

Invited Papers

1:30

2pMU1. Physical and perceptual modeling of sound-source radiation. Olivier Warusfel, Nicolas Misdariis, and René Causse (IRCAM, 1 Pl. Stravinsky, 75004 Paris, France)

Radiation characteristics of sound sources are now considered important when reproducing the signal of a musical instrument or, more generally, when controlling a virtual auditory scene. Both applications have raised the need for analyzing, modeling, and synthesizing the radiation behavior of sound sources. The first part describes a general method for the reproduction of arbitrary radiation patterns with digitally controlled loudspeakers arranged to form a three-dimensional array. In order to evaluate the performance of the synthesis method, this section presents possible definitions of similarity criteria based either on a physical or a perceptual approach. In the second part, perceptual features are investigated based on the analysis of musical instruments or listening tests conducted with elementary radiation patterns such as the first spherical harmonics. The perceptual analysis is conducted first with the help of classical acoustic cues used in room acoustics, and second by analyzing the possible consequences of directivity characteristics on stream segregation mechanisms involved in auditory scene perception. The final part presents ongoing research on the reproduction of salient directivity cues of different musical instruments, as well as the voice.

1:55

2pMU2. Perception of shape and material from contact sounds. Dinesh K. Pai (Dept. of Computer Sci., Univ. of British Columbia, Vancouver, BC, Canada), Roberta L. Klatzky (Carnegie Mellon Univ., Pittsburgh, PA), and Eric P. Krotkov (Cytometrics, Inc., Philadelphia, PA)

The perception of sources of percussive sounds, as a function of variables that govern sound synthesis, was investigated. In the first set of experiments, perception of the source's material was investigated. Two types of experiments were conducted. In the first, subjects judged similarity with respect to material and length. The sounds corresponded to modal vibrations of struck clamped bars. They varied in fundamental frequency and frequency-dependent rate of decay. Differences between sounds in both decay and frequency affected similarity judgments, with decay playing a substantially larger role. In the second type of experiment, subjects assigned sounds to one of four material categories. Decay parameters for each category were estimated and found to correlate with measurements reported in the literature. The second set of experiments investigated perception of contact location on the object. Subjects were presented with pairs of sounds produced by striking a string at two locations. Two strings with different fundamental frequencies were used. In one experiment, subjects judged the distance between the strike points, and in another, judged which point was closer to the point of attachment. The results confirm that subjects are sensitive to the interstrike distance and can judge which point is closer.

2:20

2pMU3. Knowledge acquisition by listeners in a source learning task using physical models. Gary P. Scavone (Ctr. for Comput. Res. in Music and Acoust. (CCRMA), Stanford Univ., Stanford, CA 94305), Stephen Lakatos (Washington State Univ., Vancouver, WA 98686), and Perry R. Cook (Princeton Univ., Princeton, NJ 08544)

Two experiments examined listeners' ability to learn about the acoustic properties of physical models by varying the amount of structured information listeners received about the models. In experiment 1, 60 listeners participated in a repeated learning-testing paradigm, using a graphical interface to permit listeners to manipulate three acoustic properties for each of eight simulated percussive instruments. Information about the properties of the instruments varied across six conditions, ranging from highly structured (e.g., pictures of instruments, description of physical properties, sliders grouped by instrument with properties labeled) to weakly structured (e.g., unlabeled sliders arranged randomly without reference to instruments). Performance on discrimination and memory tasks suggests that learning can occur even with weakly structured acoustic information. To determine whether learning can occur even

when acoustic information is obtained “passively,” 30 listeners in experiment 2 were yoked to listeners in three conditions of experiment 1. These former listeners heard sounds produced by the manipulations of listeners in experiment 1, but were unaware of the nature of the information in the physical model interface. Both experiments suggest that mechanisms associated with acquiring knowledge about acoustic properties may be largely automatic.

2:45

2pMU4. Patterns of spectra and timbre in musical context. Roger A. Kendall (Music Cognition and Acoust. Lab., UCLA, 405 Hilgard, Los Angeles, CA 91406, kendall@ucla.edu)

Although there are a few exceptions, much of the research on musical timbre has involved single notes drawn from the center of the tessituras of the instruments employed. In single-tone research, multiple authors have found that the long-time-average spectral centroid maps well with the primary dimension of multidimensional perceptual spaces (ca. $r > 0.89$). The second dimension of perceptual spaces often maps to a measure of spectral flux or variability. The present study examines the patterning of spectral centroid and flux in quasi-musical contexts. Bb major scales consisting of quarter notes followed by quarter rests (M.M.=72) were recorded across the playing range of a variety of orchestral winds and strings. In addition, a folk song in Bb major was performed in normal tessitura for a given instrument. Spectral analyses (ninth-order FFT) were conducted for every steady state, and the unweighted spectral centroid, independent of frequency, calculated. Various statistical and mathematical analyses were performed, including the construction of a theoretical multidimensional map that incorporates pitch class. Perceptual analyses were conducted using verbal ratings of timbral similarity and quality. Results suggest that patternings of spectra are well connected to the fingerings employed, and hence driver-resonator characteristics.

3:10

2pMU5. Automatic identification of musical woodwind instruments using pattern recognition. Judith C. Brown (Phys. Dept., Wellesley College, Wellesley, MA 02181 and Media Lab., MIT, Cambridge, MA 02139, brown@media.mit.edu)

Computer recognition of musical instruments is a relatively unexplored and potentially very important research problem for its promise to liberate humans from time consuming searches on the Internet and the sequential identification of audio material on archival tapes. Pattern recognition with a cluster-based Gaussian mixture model has been used to identify a statistically significant number of short sounds from commercial recordings representing four musical classes. These are the oboe, flute, saxophone, and clarinet, all members of the woodwind family of instruments, and which share similar excitation and resonators. Features explored include cepstral coefficients, constant Q coefficients, spectral centroid, autocorrelation coefficients, and moments of the time wave. The most successful features for classification gave correct results in the range of 79%–84%. These were cepstral coefficients, bin-to-bin differences of the constant Q coefficients, and autocorrelation coefficients. The latter have not been explored previously in either speaker or instrument identification studies. Overall our results show great promise for automatic classification of musical sounds.

Contributed Papers

3:35

2pMU6. Semantic properties of short musical stimuli and their physical correlates. Eric D. Scheirer (Machine Listening Group, MIT Media Lab., Cambridge, MA 02143, eds@media.mit.edu)

When human listeners are confronted with musical stimuli, they rapidly and automatically orient themselves in the sound and begin to make semantic judgments about it. This ability—which might be termed perception of the musical surface—includes the perception of features such as genre, emotional content, and high-level descriptions. There is relatively little known today about the exact perceptual properties that listeners associate with actual musical sounds, or about the physical correlates of these properties. A preliminary study examined the perceptions of 40 musician and nonmusician subjects in response to 150 complex musical stimuli, each 5 s long, taken directly from a popular Internet music database site. Each listener rated each stimulus on six semantic scales: simple–complex, loud–soft, fast–slow, soothing–annoying, interesting–boring, and familiar–unfamiliar. Experimental results indicate that these ratings are made consistently between listeners and between segments of the same song. A psychophysical model, based on a pattern-recognition framework applied to a subband periodicity representation (the autocorrelogram), was able to predict a significant amount of the variance in re-

sponses on the first four scales. The regression weights and the psychophysical model suggest possible physical correlates that may underlie these semantic judgments.

3:50

2pMU7. Automatic classification of musical instrument sounds. Bozena Kostek (Sound Eng. Dept., Tech. Univ. of Gdansk, 80-952 Gdansk, Poland, bozenka@sound.eti.pg.gda.pl)

The aim of the present study is to show that the process of automatic classification of musical instrument sounds is possible on the basis of a limited number of parameters. However, due to the complexity as well as to the unrepeatable nature of musical sounds, both steady and transient states should be taken into account while creating feature vectors. For this purpose a database of musical instrument sounds was built containing various instrument sounds played with a different articulation. Then, this database was used in further experiments consisting of some stages, i.e., preprocessing, parametrization, and pattern recognition. The main subject of this study was the optimization of the set of parameters to be included in the feature vectors. Therefore, the quality of parameters with regard to automatic pattern recognition was analyzed both using statistical methods and learning algorithms.

Session 2pPA

Physical Acoustics: Thermo-, Photo-, Magneto-Acoustics II

Robert A. Hiller, Chair

Department of Physics and Astronomy, University of Mississippi, University, Mississippi 38677

Contributed Papers

3:45

2pPA1. Investigations of a thermoacoustic refrigerator prototype. Brian Minner, Luc Mongeau, and Jim Braun (Purdue Univ., 1077 Herrick Labs., West Lafayette, IN 47907-1077)

A thermoacoustic refrigerator prototype was built and tested. The system was designed to provide 140-W cooling power with a temperature lift of 40 °F, using a 55% mix of helium–argon at 20 bars pressure and operating at 180 Hz. Leaf springs were used to tune the mechanical driver system to be resonant at 154 Hz, such that the sum of mechanical and acoustical reactances is zero near 180 Hz. Initial testing achieved acoustic powers near 50 W, with an electroacoustic transduction efficiency of about 27% (the maximum efficiency is near 45%). Pressure amplitudes were near 3% of the mean in the stack region, and heat rejected to a secondary fluid was measured to be near 60 W. While this acoustic power level is near the target for the desired cooling load, initial tests indicate that the cooling power is well below 140 W. Results of continuing testing and investigations for determination of the cause of discrepancies in pressures, temperatures, and cooling power between measurement and model (Deltae) will be presented. The influence of DC flows, nuisance heat loads, gas mixture, and hardware (stack and heat exchanger) factors will be addressed. Accelerometer, thermocouple, and pressure sensor data will be presented and interpreted.

4:00

2pPA2. Characterization of a 300-W linear actuator. Brian Minner, Luc Mongeau, and Jim Braun (Purdue Univ., 1077 Herrick Labs., West Lafayette, IN 47907-1077)

The linear parameters of a moving magnet electrodynamic actuator were measured using several different procedures. The methods employed included blocked carriage tests, regression of free decay of voltage and acceleration for known mass addition, regression of complex impedance data resulting from peak-averaged voltage/current transfer functions in swept sine tests, and solution of nonlinear equations for impedance and velocity/voltage relationships for three or four data point sets. It was found that the mechanical system parameters are strong functions of the motional amplitude, and that electrical system parameters are functions of frequency and amplitude. A brief review of the literature and a comparison of the methods and results will be presented. Simulated data will be used to benchmark the methods, enabling clearer interpretation of the results.

4:15

2pPA3. Direct measurement of $F(\lambda_v)$ for a single circular pore. Andi Petculescu and Larry Wilen (Dept. of Phys. and Astron., Ohio Univ., 251 Clipping Labs., Athens, OH 45701, apetcule@helios.phy.ohiou.edu)

A new technique for the direct measurement of the viscous dissipation function $F(\lambda_v)$ in a circular pore will be described. The experiment is based on a loudspeaker-driven Helmholtz resonator geometry; the pore represents the aperture neck of the resonator. The technique involves measuring the acoustic impedance by simultaneously recording oscillatory

pressure and displacement and using lumped-element circuit models to extract $F(\lambda_v)$. Correction for end effects is attained by subtracting the impedances of two pores with different lengths prior to calculating $F(\lambda_v)$. Runs were taken between 0.5 and 96 Hz, which correspond to values of λ_v between 0.6 and 8 for the pore size used. There is good agreement with theory over most of the frequency range. The experimental setup offers versatility since it allows easy substitution of different samples. Time permitting, we will describe application of the technique to thermoacoustic stack geometries such as reticulated vitreous carbon (RVC) or pin arrays.

4:30

2pPA4. Experimental instability study in a thermoacoustic prime-mover. Emmanuel Bretagne, Maurice-Xavier François, and Ivan Delbende (Univ. of Paris 6, LIMSI-CNRS, B.P. 133, 91403 Orsay Cedex, France, mxfr@limsi.fr)

First, the stability properties of a linear thermoacoustic prime-mover are determined experimentally. For different values of the mean pressure between 0.5 and 10 bars, the damping rate is carefully obtained as a function of the temperature gradient $|\nabla T|$ enforced along the stack, up to the instability onset at $|\nabla T|_c$. These results are then confronted with the predictions of a numerical model based on Rott's theory used with complex frequencies, in which the prime-mover is seen as a feedback loop in the electrical analogy. The experimental results are found to comply with Rott's theory as soon as the mean pressure exceeds 2 bars. Below this value, which pertains to the merging of facing stack boundary layers, substantial discrepancies are found, as confirmed by the work of Yazaki. Second, the saturated wave amplitude has been measured as a function of $|\nabla T| - |\nabla T|_c$ for fixed pressure above the instability onset. The bifurcation to the nonlinear saturated wave may be subcritical, although thermal inertia effects make it look supercritical.

4:45

2pPA5. Instability in a thermoacoustic prime-mover: A numerical approach. Ivan Delbende (Univ. of Paris 6, LIMSI-CNRS, B.P. 133, 91403 Orsay, France, delbende@limsi.fr)

A numerical model aimed at investigating the development of instability in a thermoacoustic prime-mover is presented. It is based on a discretization of the two-dimensional compressible Navier–Stokes equations, and does not rest on the plane pressure wave assumption. The computational domain is made of a single stack interplate space extended up to both resonator ends. The working gas is assumed to be initially at rest, and a longitudinal temperature gradient is enforced. The linear instability properties of the system are then given by the eigenmodes of the discretized evolution operator linearized in the vicinity of the above basic state. Temporal growth-rate and oscillation frequency are thus obtained for all thermoacoustic modes, and the prime-mover instability onset is determined as a function of the various parameters (mainly temperature gradient magnitude and mean pressure). As far as linear instability is concerned, this formulation is attractive, since it avoids any direct numerical simulation and any problem involved by the presence of two very different time scales.

2p WED. PM

Session 2pPPa

Psychological and Physiological Acoustics: Discrimination and Pitch

Søren Buus, Chair

*Institute for Hearing, Speech and Language and Communication and Digital Signal Processing Center, ECE Department,
Northeastern University, Boston, Massachusetts 02115*

Contributed Papers

1:30

2pPPa1. Intensity-difference limens as a function of level for brief 2-kHz Gaussian-shaped tones. Lance Nizami, Walt Jesteadt, and Jason F. Reimer (Boys Town Natl. Res. Hospital, 555 N. 30th St., Omaha, NE 68131)

Intensity-difference limens (DLs) were measured for brief 2-kHz tones whose frequency spectra contained no sidelobes to 90 dB below the peak. The tone's pedestal levels spanned 30–90 dB SPL in 10-dB steps, where tone level was the SPL of the carrier. Tone duration was set by terminating the Gaussian envelope at 4 standard deviations from the mean. The DLs were obtained using Levitt's method for two men and three women (ages 18–40) having normal auditory thresholds. The tone durations, 32, 8, and 4 ms, were all integer multiples of the carrier's period. For the 32-ms tones, plotting DL vs pedestal level revealed the "near-miss to Weber's law" known for longer tones, although present DLs were higher. The DLs for 8 and 4 ms rose from about 2 dB at 30 dB SPL to 6–8 dB at 50–70 dB SPL, dropping to 1–2 dB at 90 dB SPL. (DLs will also be shown for 2 and 16 ms for three of the listeners.) A nonmonotonic DL also occurs for the acoustic click [Raab and Taub, *J. Acoust. Soc. Am.* **46**, 965–968 (1969)], which is surprising as the click's frequency spectrum is much broader than that of the Gaussian-shaped tone.

1:45

2pPPa2. Perceptual channel weights for level discrimination. Søren Buus (Inst. for Hearing, Speech, and Lang. and CDSP Ctr., ECE Dept. (442 DA), Northeastern Univ., Boston, MA 02115, buus@neu.edu), Reinier Kortekaas, and Mary Florentine (Northeastern Univ., Boston, MA 02115)

This study examines how listeners combine information for level discrimination (*aka* intensity discrimination) across auditory channels. Individual components' contributions to listeners' decisions were measured in a conditional-on-single-stimulus (COSS) paradigm. Stimuli were 7- and 15-tone complexes with components spaced one critical band apart around 1600 Hz. The mean component level was 60 dB SPL for the standard and was increased for the target to obtain 75%–80% correct responses. Results for four normal listeners show that the weights were highest for the two highest-frequency components when all components had equal level increments. Progressively increasing the level increment for components above the center generally had little effect on the weights. This indicates that listeners may be limited to using an optimal selection of channels for the decision. Adding 3-Bark-wide roving-level maskers above and below the complexes had little effect on the weights for the 7-tone complexes, but reduced the weights for the high components of the 15-tone complexes. The masker bands had near-zero weights. Altogether the results indicate that level discrimination depends on information across many auditory channels. Differences in component contributions are unlikely to reflect optimum weighting and are only partly explained by spread of excitation. [Work supported by NIH/NIDCD Grant No. R01DC00187.]

2:00

2pPPa3. Intensity discrimination of brief Gaussian-windowed tones by hearing-impaired listeners. Thomas Baer, Brian C. J. Moore, and Josephine Marriage (Dept. of Exp. Psych., Univ. of Cambridge, Downing St., Cambridge CB2 3EB, UK, tb107@cus.cam.ac.uk)

For normal listeners, intensity DLs for Gaussian-shaped tone pulses are largest at medium pulse durations when the pedestals are 10 dB above threshold, either in quiet or in a pink-noise background. One explanation is that worst performance occurs when the internal representation is most compact in time and frequency, affording minimal opportunity for multiple looks [N. H. van Schijndel, T. Houtgast, and J. Festen, *J. Acoust. Soc. Am.* **105**, 3425–3435 (1999)]. However, the mid-duration worsening is largest for medium overall levels, suggesting an involvement of the compression on the basilar membrane (BM), which is also greatest at medium levels [T. Baer, B. C. J. Moore, and B. R. Glasberg, *J. Acoust. Soc. Am.* **106**, 1907–1916 (1999)]. If so, the mid-duration worsening should be reduced when BM compression is reduced by outer hair cell damage. To test this, subjects with sensorineural losses were tested using 1-kHz or 4-kHz pulses, in quiet or in pink noise that raised thresholds by 10–20 dB. For subjects with mild losses, poorest performance was sometimes found for medium durations. For more severe losses, intensity DLs tended to improve monotonically or remain roughly constant with increasing duration. The results provide support for both multiple-look and BM-compression explanations. [Work supported by MRC.]

2:15

2pPPa4. Identification and similarity judgments of environmental sounds. Brian Gygi (Psych. Dept., Indiana Univ., Bloomington, IN 47405), Gary R. Kidd, and Charles S. Watson (Indiana Univ., Bloomington, IN 47405)

As part of an ongoing study of environmental sounds, identification thresholds were obtained for a collection of 100 naturally occurring familiar sounds embedded in noise. Threshold event-to-noise ratios covered a range of over 23 dB. Differences in total energy or in the spectral-temporal distribution of energy accounted for, at best, only 45% of the variance in identification thresholds. To determine the association between stimulus similarity and differences in identification thresholds, a series of three similarity rating studies was conducted. Separate groups of subjects judged the similarity of the sounds based on actual acoustic properties (sounds heard), imagined acoustic properties (sounds imagined), and semantic properties (*i.e.*, properties attributed to the source object or event). The MDS solutions for all three studies were quite similar and contained clusters suggesting that similarity judgments with these familiar sounds are based on a few very general acoustic properties including rhythmicity, continuity and spectral-temporal complexity. [Work supported by NIH.]

2pPPa5. Unexpected benefits of training frequency discrimination on the learning and generalization of interval discrimination. Beverly A. Wright and Jeanette A. Ortiz (Audiol. and Hearing Sci. Prog., 2299 North Campus Dr., Northwestern Univ., Evanston, IL 60208-3550)

It is generally accepted that training on one perceptual task does not yield learning on untrained tasks. If so, interleaved training on two tasks should not affect learning or stimulus generalization on either task. Contrary to this prediction, six listeners trained 2730 trials each on interval- and frequency-discrimination tasks using the same stimulus (100 ms at 1 kHz) learned on the interval task 30% more with the trained stimulus, and 134% more with an untrained stimulus (100 ms at 4 kHz), than ten listeners trained 2934 trials on the interval task alone. Furthermore, listeners trained on both tasks with the same stimulus learned on the trained and untrained stimuli, respectively, (1) 105% and 136% more than 6 listeners trained 2160 trials each on both tasks but with a different stimulus (200 ms at 4 kHz) for the frequency task, (2) 19% and 42% more than 20 listeners trained 9675 trials on the interval task alone, and (3) 71% and 579% more than 6 listeners trained 9000 trials on the frequency task alone. Thus practicing a second task with the same trained stimulus unexpectedly enhanced learning and particularly stimulus generalization on the first task, more so than performing additional trials on either task alone. These results may guide the development of theoretical models and clinical applications of learning. [Work supported by NIDCD.]

2pPPa6. The effect of interval order and highpass filtering on the pitch strength of shuffled click-train regular interval stimuli. Dan Mapes-Riordan and William A. Yost (Parml Hearing Inst., Loyola Univ. of Chicago, 6525 N. Sheridan Rd., Chicago, IL 60626, dmapes@phi.luc.edu)

Regular interval stimuli (RIS) contain temporal regularities that are known to produce a pitch percept. Previous experiments [Mapes-Riordan and Yost, ARO MidWinter meeting proceedings (2000)] showed that shuffled, wide-band first-order "kxkx" click-train RIS have a stronger pitch strength than the original, unshuffled stimuli, where "k" is a fixed

duration interval and "x" is a random interval. The primary difference between ordered and shuffled click-train sequences is that shuffled sequences contain multiple instances of regular "k" intervals. This result provides evidence for the importance of the short-term periodicity within RIS in determining its pitch strength. The current experiments investigated whether this phenomenon also occurred when the RIS were created from higher-order intervals and/or whose spectrum was limited to unresolved frequency channels. A set of trials was run in which listeners compared the pitch strength of an ordered click-train RIS and a shuffled version of the same sequence. Relative pitch strength judgments were made as a function of the regular interval duration, regular interval order, and high-pass filter cutoff frequency. The results of these experiments revealed that second-order interval click-train RIS require resolved frequency components to generate a salient pitch percept. [Work supported by a NIDCD Program Project Grant.]

2pPPa7. Quantifying the influence of dynamic intensity on changes in pitch. John G. Neuhoff (Dept. of Psych., Lafayette College, Easton, PA 18042-1781, neuhoffj@lafayette.edu)

Previous work has shown that dynamic continuous changes in intensity can influence judgments of pitch change in ways that are not predicted by classic equal-pitch contours derived with static stimuli. Rising intensity can elicit judgments of rising intensity even at frequencies where static increases in intensity yield judgments of lower pitch [J. G. Neuhoff and M. K. McBeath, J. Exp. Psychol. Human Percept. Perform. **22** (1996)]. However, the effects thus far have been measured in arbitrary perceptual units. Thus, we know relatively little about the absolute amount that one dimension influences the other in physical terms. In the current work, listeners heard tones that were simultaneously modulated in both frequency and amplitude at the same rate (2 Hz). When frequency and amplitude modulation were in phase, frequency was judged to change more than when frequency and amplitude modulation were 180 degrees out of phase. To quantify the effect in physical terms, pitch change in the experimental trials was compared to a control condition in which tones were only frequency modulated. The findings have implications for sonification and auditory display as well as structural modeling of pitch and loudness. [Work supported by NSF Grant No. 9905266.]

WEDNESDAY AFTERNOON, 31 MAY 2000

STATE ROOM, 3:30 TO 4:45 P.M.

Session 2pPPb

Psychological and Physiological Acoustics: Binaural Hearing

Leslie R. Bernstein, Chair

Center for Neurological Sciences, University of Connecticut Health Center, 261 Farmington Avenue, Farmington, Connecticut 06030

Contributed Papers

2pPPb1. Binaural coherence edge pitch—Manipulations and models. William M. Hartmann, Colleen D. McMillon, and Brad Rakerd (Michigan State Univ., East Lansing, MI 48824)

Binaural coherence edge pitch (BICEP) is a dichotic broadband noise pitch effect. The stimulus is made by summing spectrally dense sine components with random phases. The *interaural* phase angle is a constant (e.g., 0 or π) for components with frequencies below (or above) a chosen edge frequency, and it is a random variable for the remaining components. Pitch-matching experiments show that the BICEP exists for coherence

edge frequencies between 300 and 1000 Hz. It is matched by a pure tone frequency shifted from the edge frequency by 5% to 10%. The matching frequency falls on the incoherent side of the edge. The BICEP persists if the noise to one of the ears is delayed (< 3 ms). Some BICEP exists even when the component amplitudes are entirely independent in the two ears, so long as the phase coherence conditions of the stimulus are maintained. The existence of the BICEP is a challenge for current models of dichotic pitch, though the pitch shift is consistent with the equalization-cancellation model as previously applied to dichotic pitch effects by M. A. Klein and W. M. Hartmann [J. Acoust. Soc. Am. **70**, 51–61 (1981)]. [Work supported by the NIDCD of the NIH and by the NSF REU Program.]

2pPPb2. Distance judgments of nearby sources in a reverberant room: Effects of stimulus envelope. Scott G. Santarelli, Norbert Kopčo, and Barbara G. Shinn-Cunningham (Hearing Res. Ctr., Dept. of Cognit. and Neural Systems, Boston Univ., Boston, MA 02215)

Localization was measured for nearby sources with abrupt or slow rise/fall times in a reverberant space. A recent model of distance perception [A. W. Bronkhorst and T. Houtgast, *Nature* **397**, 517–520 (1999)] suggests that perceived distance is computed from the room impulse response. The model assumes that energy in the onset of the impulse response (primarily from the direct sound, varying with distance) is compared to late energy (primarily from the reverberation, roughly independent of distance). However, other results suggest that subjects are poor at deconvolving transfer function and sound source characteristics [Rakerd *et al.*, *J. Acoust. Soc. Am.* **106**, 2812–2820 (1999)]. Taken together, these results suggest that subjects cannot use the transfer function, but estimate source distance from some statistic closely related to that proposed in the model (e.g., the ratio of initial to late energy in the total waveform at the ear). For impulsive sounds, such a simpler statistic yields results similar to those of the model; however, distance judgments would be significantly degraded for sources with slow onsets. However, subjects were equally good at judging distance, independent of characteristics of the stimulus envelope. [Work supported in part by AFOSR Grant No. F49620-98-1-0108.]

4:00

2pPPb3. Human capabilities of dereverberation. Brad W. Libbey and Peter H. Rogers (Georgia Inst. of Technol., Atlanta, GA 30332-0405, gt1556a@prism.gatech.edu)

Humans listening to speech in a small room are frequently unaware of reverberation. It is unknown if neurological processes remove these echoes or if they are simply disregarded when speech is phonetically processed. In other words, is there a neurological mechanism that is capable of removing echoes to create a clean speech neurological signal before phonetic processing? Or is the brain capable of processing reverberant phonemes? Word intelligibility experiments examine these questions. Preliminary experiments investigate how characteristics of simulated reverberation such as room size, absorption of walls, source location, and listener position affect intelligibility. The results of these experiments are used in the design of primary experiments that address human capabilities. In the capability tests the effects of binaural listening, reverberation level, and deconvolution processing are investigated. These experiments approach the fundamental questions through the use of a three-factor experiment (the factors being binaural versus diotic, high versus low levels of reverberation, and simulated reverberation versus convolutional noise). Through the primary and interaction effects of these factors the data illustrates the extent of neurological dereverberation.

2pPPb4. Discrimination of brief interaural temporal disparities embedded within diotic bursts of broadband noise. Michael A. Akeroyd, Leslie R. Bernstein, and Constantine Trahiotis (Dept. of Surgery (Otolaryngol.), Ctr. for Neurological Sci., Univ. of Connecticut Health Ctr., Farmington, CT 06032)

Zurek (1980) measured listeners' sensitivity to interaural disparities conveyed by a 5-ms segment of a 50-ms burst of otherwise diotic broadband noise [P. M. Zurek, *J. Acoust. Soc. Am.* **67**, 952–964 (1980)]. He found that thresholds for interaural time delay (ITD) were markedly elevated when their onset was between 1 to 5 ms following the onset of the noise. Zurek postulated that the leading portion of the diotic burst of noise briefly inhibited sensitivity to subsequent binaural information and, in that manner, linked his findings to the "precedence effect." In our view, one implication of that reasoning is that the trailing portion of the diotic burst should have little, if any, effect on threshold ITDs. In order to test this hypothesis, we employed Zurek's general paradigm and included conditions in which we omitted either leading or trailing portions of the diotic burst of noise. We found that omitting *either* the leading *or* the trailing portion of the diotic noise greatly reduced the elevation in threshold ITDs observed in the original paradigm. Consequently, it appears that Zurek's original data, although interesting and important, may reflect effects and mechanisms apart from those responsible for the "precedence effect." [Work supported by NIH DC 04073.]

4:30

2pPPb5. Influence of source and echo separation on echo threshold of natural complex sounds. Miriam N. Valenzuela and Ervin R. Hafter (Dept. of Psychol., Univ. of California, Berkeley, CA 94720, miriam@ear.berkeley.edu)

Most studies of the precedence effect are based on stimuli that differ dramatically from "realistic" stimuli. Important differences include the following. (a) Stimulus types used in studies are often clicks, noise bursts or sinusoids; realistic sounds are mostly harmonic or inharmonic complex tones. (b) Durations of stimuli used in studies often avoid a temporal overlap between lead and lag; durations of realistic sounds are normally such that first and subsequent wave fronts overlap most of the time. (c) Lead and lag stimuli used in studies to simulate a primary sound and its reflection from a nearby surface are often signals of equal amplitude and identical wave form; reflections of realistic sounds differ normally from the primary sound in amplitude and wave form. As a first step toward understanding the precedence effect and its significance in "real world" situations, experiments with synthetic piano tones of durations longer than 100 ms (longer than echo delays in typical rooms) were conducted in a simulated reverberant environment. The dependence of echo threshold on the direction of sound incidence was studied using measurements of fusion and localization dominance. The results show a decrease of echo threshold with an increase of spatial separation between primary source and reflection.

Session 2pSA

Structural Acoustics and Vibration: Aeroelasticity and Flow-Induced Noise and Vibration

Kenneth D. Frampton, Cochair

Department of Mechanical Engineering, Vanderbilt University, Box 1592, Station B, Nashville, Tennessee 37235-1592

Sean F. Wu, Cochair

Department of Mechanical Engineering, Wayne State University, Detroit, Michigan 48202

Invited Papers

2:00

2pSA1. Nonlinear dynamics of aeroelastic systems. Earl H. Dowell (Dept. of Mech. Eng. and Mater. Sci., Duke Univ., Box 90300, Durham, NC 27708, dowell@mail.ee.duke.edu)

Aeroelastic systems are those that involve the coupled interaction between a convecting fluid and a flexible elastic structure. The nonlinear dynamical response of such systems is known to encounter limit cycle oscillations (LCO) in certain flight regimes and relatively simple experimental wind tunnel models have been designed to exhibit LCO as well. In the present paper, the results of these wind tunnel experiments are discussed and compared to comparable results from mathematical models. The physical models include (1) an airfoil and a control surface attached with an elastic spring including free-play, and (2) a delta wing with elastic geometrical nonlinearities due to bending and torsional deformations. Both self-excited oscillations such as flutter and LCO, as well as forced oscillations due to an aerodynamic gust, are discussed. The advantages of representing the unsteady aerodynamic flow field in terms of global modes for such studies are emphasized and illustrated.

2:20

2pSA2. Feedback criteria for edge tones and the multiple sources of choked jet screech. Alan Powell (Dept. of Mech. Eng., Univ. of Houston, Houston, TX 77204-4792)

Screech sound is generated by the interaction of jet instability waves with the spatially periodic cell structure of a choked supersonic jet, the sound initiating new instability waves at the nozzle. Its frequency is important because of possible structural damage to high-performance aircraft. It was hypothesized that the interaction resulted in a phased array of point sources, with an effective source at distance h' from the nozzle. This reduced the phase feedback criterion to that of edge tones with frequency $f = (U/h')(N+p)/(1+M)$, where U , M are, respectively, the velocity and Mach number of the instability waves, N is an integer, and p an ideally fixed constant, $0 \leq p \leq 1$. It was hypothesized that N takes the value yielding maximum feedback for maximum jet instability for edge tones, and that closest to perfect reinforcement, $f \rightarrow f_{PR} = (U/s)/(1+M)$, s = cell length, for screech. Multiple-source feedback is now considered, allowing for a different (effective) first cell length. The preceding concepts are fully confirmed for screech, but N is not limited to integer values. As illustrated by experimental results, and though $f \rightarrow f_{PR}$, absolutely perfect reinforcement is unlikely to occur.

2:40

2pSA3. Methods for calculating the response of structures to inhomogeneous turbulent flows. Stewart Glegg (Ctr. for Acoust. and Vib., Dept. of Ocean Eng., Florida Atlantic Univ., Boca Raton, FL 33431) and William Devenport (Virginia Polytechnic and State Univ., Blacksburg, VA 24061)

Problems involving the response of structures to turbulent flows are usually modeled using a linear approach. The incoming turbulence is described using a Von Karman or Liepmann spectrum using typical turbulence velocity and length scales. However, these models only apply to homogeneous turbulence and so inhomogeneous flows can only be treated in an approximate manner. This paper will describe an alternative approach in which the turbulent flow is decomposed into a set of uncorrelated modes using proper orthogonal decomposition. The response of the structure to each mode can then be considered independently and the results summed incoherently to give the complete response. However, to obtain the proper orthogonal modes the flow must be measured in great detail. It will be shown that this requirement can be reduced if measurements are limited to the most energetic parts of the flow.

3:00

2pSA4. Active control of turbulent-boundary-layer-induced sound radiation from aircraft style panels. Gary P. Gibbs, Kenneth W. Eure (NASA Langley Res. Ctr., Hampton, VA 23681, G.P.GIBBS@LaRC.NASA.GOV), and J. W. Loyd (Virginia Tech, Blacksburg, VA 23461)

The results of active control of TBL-induced sound radiation will be presented. In this test, a panel was constructed similar to the sidewall of an aircraft; however, the panel was flat to allow a simple incorporation into the wind tunnel wall. The static in plane stress normally associated with aircraft cabin pressurization (at 40 000 ft) was provided via a tensioning fixture. Thus the smooth "exterior" side was subjected to TBL flow inside the tunnel, and the other (interior) side radiated sound into the model preparation room. Accelerometers were mounted on the interior side of the panel, and microphones were positioned in the model preparation room to

monitor sound radiation. Experiments were conducted at both Mach 0.8 and Mach 2.5 flow conditions, and active control was performed on a single bay, and two bays. The results demonstrate reductions in total radiated sound power on the order of 10–15 dB at resonances, and 5–10 dB integrated over the band width of 150–800 Hz. This work represents the first known demonstration of ASAC of aircraft-style panels subjected to both subsonic and supersonic flows.

3:20

2pSA5. Controlling nonlinearity and chaos. Lucio Maestrello (NASA Langley Res. Ctr., M/S 463, Hampton, VA 23681-2199)

Several unrelated experiments were conducted in which nonlinear and chaotic behaviors of structural vibration and shock waves from free shear flow were observed. Since each experiment is sensitive to the initial condition, this makes them highly receptive to active control. Control of the vibration of the structure and free shock shear layer was achieved by redistribution of energy, such that the energy in the system is nearly preserved. Experiments are conducted to illustrate the role of the initial conditions in the active control strategy.

3:40

2pSA6. Application of reduced-order models to the design of adaptive aeroelastic structures. Robert L. Clark, John A. Rule, and Robert E. Richard (Dept. of Mech. Eng. and Mater. Sci., Duke Univ., Box 90300, Durham, NC 27708-0300)

Reduced-order models of potential-flow aerodynamics have been developed and greatly facilitate the analysis and design of aeroelastic systems in the early design phase. Models capturing 95% of the physics with 5% of the modeling effort can be realized to evaluate various active and passive design considerations. A delta wing model has been developed to determine the most effective locations for transducers required to provide gust alleviation and flutter control. The basic design philosophy will be presented, and both analytical and experimental results will be discussed. Results from this study indicate that a single sensor/actuator pair can be designed to significantly extend the flutter boundary.

4:00–4:15 Break

Contributed Papers

4:15

2pSA7. Compliant-backplate Helmholtz resonators. Stephen B. Horowitz, Toshikazu Nishida (Dept. of Elec. and Computer Eng., P.O. Box 116130, Univ. of Florida, Gainesville, FL 32611, steveh23@ufl.edu), Lou N. Cattafesta III, and Mark Sheplak (Univ. of Florida, Gainesville, FL 32611)

The results of a theoretical and experimental investigation into compliant-backplate Helmholtz resonators are presented. The motivation for this study is to develop a prototype energy-reclamation module based upon a Helmholtz resonator that uses a compliant piezoelectric backplate. To establish a baseline case, rigid backplate Helmholtz resonators are first designed using a lumped element model and then tested in a plane-wave tube. The plane-wave tube, in the Interdisciplinary Microsystems Lab at the University of Florida, permits characterization in a known acoustic field at frequencies up to 20 kHz. After this initial testing, the rigid backplate of the resonator is then replaced by a compliant piezoelectric diaphragm and tested in the plane-wave tube. The experimental results are compared to theory based on lumped element modeling of the Helmholtz resonator with a compliant diaphragm. The compliant-backplate configuration results in a lower resonant frequency relative to the rigid-backplate configuration. Finally, the potential for extracting useful energy from broadband pressure fluctuations outside the resonator (e.g., from a grazing turbulent boundary layer) and converting this energy to its electrical form via electromechanical transduction is evaluated. [Work funded by NASA Langley Research Center.]

4:30

2pSA8. An experimental and numerical investigation of sound produced by low-speed flow over a cavity with a leading edge overhang. Brenda S. Henderson, Homayun K. Navaz, and Raymond Berg (Mech. Eng. Dept., Kettering Univ., 1700 W. 3rd Ave., Flint, MI 48504-4898, bhenders@kettering.edu)

An experimental and numerical investigation is conducted on low-speed flow over a cavity. The production of discrete tones and development of complex flow structures are studied for deep cavities with leading edge overhangs. Experimental results indicate that multiple discrete fre-

quency tones are produced simultaneously or intermittently. The frequency characteristics and maximum sound pressure level depend on the leading edge overhang geometry. The onset of instability and production of multiple tones appear to be affected by boundary layer thickness. Numerical studies show a complex periodic vortex structure in the cavity mouth region, with a period equal to that measured for the dominant tone in the experiments. The vortex structure is affected by the leading edge overhang geometry. Mass addition and extraction from the cavity appear to be controlled by the vortex structure in the cavity mouth, which may explain the effect of lip geometry on peak sound pressure levels in the cavity.

4:45

2pSA9. On the existence of flexural edgewaves on fluid-loaded plates. I. David Abrahams (Dept. of Math., Univ. of Manchester, Oxford Rd., Manchester M13 9PL, UK) and Andrew N. Norris (Rutgers Univ., Piscataway, NJ 08854-8058)

Thin elastic plates *in vacuo* support flexural edge waves with wave speed proportional to and slightly less than the speed of flexural waves on a plate of infinite extent. This phenomenon has been known for 40 years, and has been experimentally verified. In this talk, the existence of edge-supported flexural waves on fluid-loaded plates is established theoretically. While the analogous *in vacuo* edge waves exist for all parameter values, submerged plates are shown to support such waves only under very light fluid-loading conditions. For example, thin plates of aluminum, brass, or Plexiglas will not support edge waves in water, although edge waves are permissible for each of these materials in air. More specifically, the nondimensional fluid-loading parameter must be less than 0.0462 for edge waves to exist on plates with Poisson's ratio of 0.33. The analysis is based on classical thin-plate theory and employs the Wiener–Hopf technique to derive the dispersion relation for the edge wave number as a function of frequency. In the limit of zero fluid loading, the dispersion relation predicts the well-known result of Kononov (1960) for edge waves on thin plates.

5:00

2pSA10. Noise generated from a flexible and elastically supported structure subject to turbulent boundary layer flow excitation. Michael J. Allen and Nickolas Vlahopoulos (Dept. of Naval Architecture and Marine Eng., Univ. of Michigan, 2600 Draper Rd., Ann Arbor, MI 48109-2145)

A numerical methodology based on the finite-element and boundary-element methods is presented for computing the noise radiated from a structure subject to boundary layer excitation. The new algorithm utilizes the fluctuating wall pressure in order to define the excitation on the structural acoustic system. The developments target wind-noise prediction for the sound radiated by the side glass window of an automobile. The glass-seal assembly is modeled as a flexible plate mounted on a foundation with stiffness and damping characteristics. Numerical predictions are compared successfully to wind-tunnel test data. Parametric analyses are performed in order to identify the characteristics of the seal that can lead to noise reduction. [Work supported by Ford Motor Co.]

5:15

2pSA11. Experiments on radiated noise and vibration from a lifting surface at high Reynolds numbers. David R. Dowling, Dwayne Bourgoynne, Steve Ceccio (Dept. of Mech. Eng. & Appl. Mech., Univ. of Michigan, Ann Arbor, MI 48109), and Thomas Mathews (Naval Surface Warfare Ctr., West Bethesda, MD 20817)

One of the main hydroacoustic noise sources from fully submerged lifting surfaces is the unsteady separated turbulent flow near the surface's trailing edge that produces pressure fluctuations on the surface and induces vibratory motions of the lifting surface itself. However, the hydrodynamic forcing and subsequent structural response of lifting surfaces are largely undocumented at the high Reynolds numbers typical of many marine propulsion applications. This talk describes a new experimental effort to identify and experimentally document the turbulent flow, induced surface pressures, structural response, and radiated noise of a hydrofoil at chord-based Reynolds numbers up to 60 million. The experiments are conducted at the

US Navy's Large Cavitation Channel with a two-dimensional test-section-spanning hydrofoil (2.1-m chord, 3.0-m span) at flow speeds from 0.5 to 18 m/s. The foil section is a modified NACA 16 with a flat pressure side. At a zero angle of attack, the lift load on the foil approaches 700 kn. The results presented here cover the first phase of the experiments and illustrate flow-structure coupling phenomena that will be investigated in greater detail in the second phase of experiments planned for later this year. [Work sponsored by ONR, Code 333.]

5:30

2pSA12. A formulation for modeling the operating propellers of an aircraft as a noise source in an acoustic boundary element analysis. Nickolas Vlahopoulos (Dept. of Naval Architecture and Marine Eng., Univ. of Michigan, 2600 Draper Rd., Ann Arbor, MI 48109-2145) and Sandy Liu (Bell Helicopter Textron, Fort Worth, TX 76101)

The propellers constitute a major noise source in turboprop/propeller subsonic aircraft. In order to optimize the interior noise reduction, it is important to have accurate information about the excitation exerted on the fuselage by the exterior acoustic field. Technology is developed for simulating the operating propellers of an aircraft as boundary conditions in an acoustic boundary element analysis (BEA). The propeller modeling utility program developed is called XSTREAM (eXterior STRuctural/Rotor Emission Acoustic Model). The geometry of the propeller and the operating conditions are utilized for generating the appropriate boundary conditions for the BEA computation. This approach fully accounts for scattering and diffraction effects due to the presence and the shape of the fuselage within the acoustic field created by the propellers. Detailed information about the acoustic loading over the entire structure is computed from the analysis. The developed formulation is implemented into computer software. The latter is utilized for simulating acoustic load exerted on the XV-15 tilt-rotor aircraft during flight. Numerical results for the acoustic pressure on the fuselage are compared successfully to test data for six flight conditions. Numerical results for the footprint of the aircraft are also compared successfully to test data for one flight condition. [Research supported by Rotorcraft Industry Technology Association.]

2p WED. PM

WEDNESDAY AFTERNOON, 31 MAY 2000

PEACH ROOM, 2:30 TO 4:50 P.M.

Session 2pSC

Speech Communication: Frank Cooper: A Legacy in Speech Acoustics

Anders Lofqvist, Cochair

Department of Logopedics and Phoniatrics, Lasarettet, Lund S-22185, Sweden

Michael F. Dorman, Cochair

Speech and Hearing Science, Arizona State University, Tempe, Arizona 85287-0102

Chair's Introduction—2:30

Invited Papers

2:35

2pSC1. Franklin S. Cooper: Pioneer and educator in speech research. Arthur S. Abramson (Haskins Labs. and Univ. of Connecticut, 270 Crown St., New Haven, CT 06511-6695)

Helped by the post-war availability of the sound spectrograph, invented at Bell Laboratories, rapid progress was made in the 1950s in acquiring an understanding of the acoustics of speech. As head of the Haskins Laboratories, then in New York City, Franklin Cooper developed a rich interdisciplinary environment for research into the information-bearing elements (the cues) embedded in the

speech signal. For him this was not only important basic research but also, as he sought to convince engineers, an indirect but vital approach to speech compression and processing. Cooper was also an educational pioneer as a teacher of the new acoustic phonetics to graduate students concerned with language but with little formal background in physics and higher mathematics. This began with his appointment as Adjunct Professor in the Department of Linguistics, Columbia University. Later, with the move of the Laboratories to New Haven, Connecticut, he oversaw the forming of links with the University of Connecticut and Yale University, which allowed for the formal participation of faculty members and a number of their graduate students in research at Haskins. Cooper's influence was deeply felt not only by academic colleagues but also by students of linguistics, psychology, and speech and hearing.

3:00

2pSC2. Muscles alive! Katherine S. Harris (Haskins Labs., 270 Crown St., New Haven, CT 06511 and The Grad. Ctr., CUNY, 365 5th Ave., New York, NY 10016, loumau@erols.com)

As the Haskins work on speech perception revealed the importance of the contribution to speech perception of time-varying aspects of the acoustic signal, Frank Cooper began a search for techniques to make detailed studies of the dynamic aspects of articulation. While some important work using cinefluorography, by Moll and his associates, had appeared in the 1960s, increasing concern about the hazards of radiation exposure severely limited the collection of substantial bodies of data by such a means. Frank suggested that recording the muscle signals underlying articulator movement, using the techniques suggested by Basmajian for various nonarticulatory systems, might have important advantages. To solve the problem of recording from tongue surfaces, he developed a suction electrode system. However, a later collaboration with researchers from the University of Tokyo using hooked wire electrodes provided better isolation of muscle fiber groups with complex relations to articulator surfaces. This work has led to present day modeling of the relationship of muscles to movement, although substantial problems remain for the work of future generations.

3:25

2pSC3. A large size in stockings is hard to sell. Lawrence J. Raphael (Dept. of Commun. Sci. and Disord., Adelphi Univ., Garden City, NY 11530)

Perhaps the most influential among Franklin S. Cooper's many contributions to the science and technology of speech research, was the design and construction of the Pattern Playback. The playback, devised originally as part of a project aimed at developing a reading machine for the blind, was significant for several reasons: It was the first speech synthesizer to be of practical use for experimental purposes. Indeed, for some time it (and its "offspring") was the only useful synthesizer. As a result, it attracted a distinguished cadre of phoneticians, linguists, and psychologists to Haskins Laboratories. This group of researchers produced a disproportionate amount of the early research on speech perception, employing playback-generated stimuli. The fact that most of the findings of this early research have withstood the test of time is a lasting testimony to the work of Dr. Cooper and his colleagues.

3:50–4:00 Break

4:00

2pSC4. Current research with cochlear implants in adults and children. Michael F. Dorman (Arizona State Univ., Dept. of Speech and Hearing Sci., Tempe, AZ 85287-0102 and Univ. of Utah Health Sci. Ctr., Div. of Otolaryngol./HNS)

Much of the early work at Haskins Laboratories, under the direction of Frank Cooper, focused on the development and evaluation of prosthetic devices. Near the end of his life, Frank Cooper again evidenced an interest in a prosthetic device—the cochlear implant. Part of Frank's interest stemmed from the fact that the most commonly used cochlear implant uses a scheme for stimulation which has a historical antecedent in the Peterson and Cooper (1957) peak-picking, vocoder (which, in turn, was based on the pattern playback). The other part of Frank's interest grew out of the many inter-related studies at Haskins Laboratories on deafness, phonetic processing, memory, and reading. In our last conversation, Frank quizzed me at length on cochlear implants in children—were the children intelligible, did they understand speech, did they acquire oral language, and did they read at grade level? In this talk I will continue my conversation with Frank and summarize recent research with cochlear implants in adults and children.

2pSC5. Goals and targets in speech production. Anders Lofqvist (Dept. of Logoped. & Phoniatrics, University Hospital, SE-221 85 Lund, Sweden, Anders.Lofqvist@logopedi.lu.se)

In a book chapter published in 1983, Franklin Cooper discussed past and current approaches to speech motor control [F. S. Cooper, "Some reflections on speech research," in *The Production of Speech*, edited by P. MacNeilage (Springer, New York, 1983), pp. 275–290]. Since he had been involved in what might nowadays be called the gestural approach to speech production at Haskins, he contrasted this approach with another one based on acoustic and/or perceptual targets. These different approaches are still being pursued today, and the presentation will focus on some of the theoretical and experimental issues in deciding on the proper coordinate system for speech planning. A survey of published experiments suggests that those studying vowels often end up with an acoustic coordinate system, while those studying consonants tend to favor an articulatory coordinate system. The notion of target will be examined and evaluated. It is suggested that this notion is more complex than usually presented, often involving multiple dynamic targets in both acoustic and articulatory space. Given that the production and perception of speech have evolved together, it may be artificial to make a sharp distinction between them and exclusively favor one coordinate system rather than the other.

WEDNESDAY AFTERNOON, 31 MAY 2000

AMERICAN ROOM, 1:25 TO 5:30 P.M.

Session 2pSP

Signal Processing in Acoustics and Noise: Time Frequency (TF) and Wavelet Processing (WP) in Acoustics

Leon H. Sibul, Cochair

Applied Research Laboratory, The Pennsylvania State University, P.O. Box 30, State College, Pennsylvania 16804

Charles F. Gaumond, Cochair

Naval Research Laboratory, 4555 Overlook Avenue S.W., Washington, D.C. 20375-5320

Chair's Introduction—1:25

Invited Papers

1:30

2pSP1. A probabilistic particle view of a propagating pulse. Leon Cohen (City Univ. of New York, 695 Park Ave., New York, NY 10471)

When time-frequency analysis is applied to a dispersive propagating pulse an interesting view is obtained. We will show how the Gram/Charlier/Edgeworth of standard probability theory can be generalized to time-frequency distributions and applied to study pulse propagation. Explicit expressions are given for the evolution of the pulse which appear to be more accurate than the standard asymptotic approximation, at least that is the case for some exactly solvable problems. Also, explicit expressions are given for the low order conditional moments and the spread of a pulse and we show that a very simple probabilistic particle model emerges that explains these results.

2:00

2pSP2. Methods and applications of time–frequency analysis. Patrick J. Loughlin (Dept. of Elec. Eng., Univ. of Pittsburgh, Pittsburgh, PA 15261)

Time–frequency analysis is a common technique for studying signals whose spectral characteristics change over time, the principal method being the short-time Fourier transform or spectrogram. In the past decade, the theory and methods of time–frequency analysis have witnessed extensive development and application, due to the effective use of Cohen's general formulation for time–frequency distributions. These methods have been found to be useful in a variety of areas, including acoustic scattering, biological signal analysis, and machine vibration analysis, among many others. A key development has been an understanding of the conditional moments of the distribution and their link to physical attributes of the signal such as its amplitude and phase. A review of the theory and methods of time–frequency analysis is presented, with an emphasis on applications in a variety of areas. [Work supported by ONR N00014-98-1-0680.]

2pSP3. Scaled window decompositions for reduced interference time–frequency distributions. William Williams (EECS Dept., Univ. of Michigan, Ann Arbor, MI 48109)

The reduced interference distribution (RID) is a time–frequency distribution (and a member of Cohen’s class) which enjoys a number of desirable properties. A simple design procedure, which allows one to start with a simple primitive function, $h(t)$ is provided. When $h(t)$ is equipped with certain constraints, a full RID kernel within Cohen’s class can easily be constructed. RID computation has traditionally been achieved by the “outer-product methods,” wherein the kernel is convolved with the local auto-correlation of the signal along time for each lag value. Amin, Cunningham, and Williams and others have provided alternative methodologies; the “inner-product methods.” This involves decomposition of the kernel into an orthonormal set of windows. Each window is used to form an STFT. We have found that a truncated set of windows can be used to closely approximate RIDs. Wavelet and scale-related windows can be used to generate a discrete time–frequency distribution that closely approximates RIDs. Illustrations of the contributions of each of the windows will be provided. This method of decomposition of the RID leads to new insights into the nature of time–frequency distributions as well as providing a method for very fast computation.

3:00–3:15 Break

3:15

2pSP4. When wavelet/scale processing is useful. Randy K. Young (Appl. Res. Lab., The Penn State, P.O. Box 30, State College, PA 16804-0030, rkyl@psu.edu)

“When wavelet/scale processing is useful” identifies specific acoustical problems where these tools is beneficial. The emphasis is on time-varying linear system modeling with a wavelet domain system representation. Systems that involve motion are emphasized. Both acoustic scattering and transmission problems will be presented and the unique features offered by wavelet/scale processing will be demonstrated. Time-frequency and wavelet/scale processing will be compared.

3:45

2pSP5. The analysis of signals containing multiple chirps. Laurie Linnett, Yuri Rzhanov, Brian Calder (Ctr. for Coastal and Ocean Mapping, Univ. of New Hampshire, Durham, NH 03824, llinnett@cisunix.unh.edu), and Chris Capus (Heriot-Watt Univ., Edinburgh, Scotland)

The analysis of signals containing multiple chirps occurs in many situations in acoustics, for example, in the bioacoustics field in analyzing the sounds of bats and dolphins. In sonar and seismic, chirp signals are used because of their pulse compression properties. Recently, the fractional Fourier transform has shown that linear chirp signals may be recognized and this has greatly facilitated the use of these signals. This paper shows a method similar to that of the fractional transform and shows how other types of chirp signals, polynomial, and other functional forms may also be analyzed. It also looks at the resolution that can be achieved when separating chirps both in rate and in time. This is of particular relevance when dealing with dispersion in the propagation of signals through different media. Examples will be presented, for the analysis of some bat signals and for some seismic signals. The analysis of signals in the time frequency domain using piecewise linear chirps will also be demonstrated, and some further examples from synthetic seismic signals will be demonstrated.

Contributed Papers

4:15

2pSP6. Performance results of noisy speech enhancement using undecimated wavelets and spectral peak enhancement. Min-sung Koh, Margaret Mortz (School of Elec. Eng. and Computer Sci., Washington State Univ., Spokane, WA 99202-1648, mkoh@eecs.wsu.edu), and Nancy Vaughan (Washington State Univ., Spokane, WA 99202-1648)

This paper presents a new speech enhancement algorithm based upon an undecimated discrete wavelet transform followed by spectral peak enhancement. Experimental data is presented which demonstrates improved automatic speech recognition scores in noisy speech. The new algorithm

first removes noise by soft-thresholding in the undecimated wavelet domain. However, the noise reduction process leaves residual noise that impairs intelligibility. In the new algorithm, residual noise removal after the undecimated wavelet denoising scheme is achieved by a nonlinear modification of the Fourier spectrum to enhance the spectral peaks relative to the valleys. Using 30 sentences at each SNR, evaluation of the new algorithm was conducted with the Connected Speech Test (CST), a procedure that ensures that the speech sentences have equal difficulty. For additive white Gaussian noise, significant improvements are observed in the experiment results and it is shown that an effective speech enhancement is obtained by spectral peak enhancement with small decomposition level. The results were variable with babble noise and office noise. Performance results will be presented.

2pSP7. Application of time-frequency analysis to landmine detection using a laser Doppler vibrometer. Vincent Valeau, James Sabatier, and Ning Xiang (Natl. Ctr. for Physical Acoust., Univ. of Mississippi, University, MS 38677, valeau@olemiss.edu)

The feasibility of detecting landmines acoustically was recently demonstrated by measuring insonified ground surface velocity using a laser Doppler vibrometer (LDV), since the presence of a mine enhances ground matrix velocity [J. M. Sabatier and N. Xiang, *J. Acoust. Soc. Am.* **106**, No. 4, Pt. 2 (1999)]. Although mines were initially detected using discrete LDV measurements, they were detected for this work by spatially continuous measurement, obtained by sweeping the beam onto the ground surface with a constant velocity in order to hasten the detection process. The ground was excited by single-tone acoustic waves. After LDV signal demodulation, a sine wave was obtained which was amplitude-modulated due to mine presence and corrupted by speckle noise “bursts” whose energy grew with the sweeping beam velocity. The signal was time-frequency analyzed by means of the smoothed Wigner distribution that provides straightforward physical interpretation since time is proportional to beam position. The mine was detected when the signal energy was enhanced at the excitation frequency. The wideband nature of speckle “bursts” allowed corrupted signal sections to be discarded. Results are provided for different mine depths, types of ground, and sweeping velocities. The limits of the analysis are investigated.

2pSP8. Forensic models for recorded acoustic gunshot signals. Steven D. Beck (BAE Systems, 6500 Tracor Ln. MS1-8, Austin, TX 78725, sbeck@tracor.com), Hirotaka Nakason (Federal Bureau of Investigation, Quantico, MD), and Joel T. Kalb (U.S. Army Res. Lab., Aberdeen Proving Ground, MD 21005-5425)

Forensic analysis of recorded acoustic gunshot signals requires knowledge of the acoustic sources and a fundamental understanding of the impulse waveforms and the physical factors affecting their characteristics. A data collection experiment was designed and conducted with the underlying purpose of determining if models can be developed to predict acoustic waveform characteristics under a set of controlled conditions. Weapons ranging from handguns to military style rifles were fired at an outdoor range and recorded under free-field conditions using microphones located at multiple range and bearing positions. Using these data, multiple discrete signal events are identified in a gunshot, and the waveform characteristics are measured. Traditional matched-filter and correlation-based matching techniques are shown to be not effective for analyzing and discriminating among gunshot signals. Instead, parametric models for these signals are developed and empirically tested. In general, the measured results conformed to the model predictions. A novel procedure is developed to decompose a superposition of signal events from a recorded gunshot waveform using adaptive physics-based wavelets. Finally, forensic procedures are developed that incorporate prior knowledge in the analysis of recorded acoustic gunshot signals.

5:00–5:30

Panel Discussion

WEDNESDAY AFTERNOON, 31 MAY 2000

ENGLISH ROOM, 12:55 TO 3:15 P.M.

Session 2pUW

Underwater Acoustics: Propagation: Experimental Data and Analysis

John A. Colosi, Chair

Woods Hole Oceanographic Institution, 266 Woods Hole Road, Woods Hole, Massachusetts 02543

Chair's Introduction—12:55

Contributed Papers

1:00

2pUW1. The North Pacific Acoustic Laboratory (NPAL) Experiment. B. D. Dushaw, B. M. Howe, J. A. Mercer, R. C. Spindel (Appl. Phys. Lab., Univ. of Washington, 1013 NE 40th St., Seattle, WA 98105, spindel@apl.washington.edu), J. A. Colosi (Woods Hole Oceanogr. Inst., Woods Hole, MA), B. D. Cornuelle, M. A. Dzieciuch, and P. F. Worcester (Scripps Inst. of Oceanogr., Univ. of California at San Diego, La Jolla, CA)

The North Pacific Acoustic Laboratory program augmented the existing ATOC acoustic network with a sparse, two-dimensional receiving array installed west of Sur Ridge, CA, close to an existing U.S. Navy

SOSUS array, during July 1998 to receive transmissions from the 75-Hz ATOC source north of Kauai. The NPAL array consisted of four 20-element vertical arrays, each with a 700-m aperture, and one 40-element vertical array with a 1400-m aperture. The arrays were deployed transverse to the 3900-km path from the Kauai source and had a total horizontal aperture of 3600 m. Data collected with the two-dimensional array and the U.S. Navy SOSUS receivers will be used to (i) study the temporal, vertical, and horizontal coherence of long-range, low-frequency resolved rays and modes, (ii) study 3D propagation effects, (iii) examine directional ambient noise properties, and (iv) to improve basin-scale ocean nowcasts via assimilation of acoustic data and other data types into models. Environmental data along the path from the Kauai source to the two-

dimensional array were acquired by two oceanographic subsurface moorings and by two XBT/CTD/ADCP transects along the path, one at the beginning and one at the end of the experiment. We describe the experiment and offer some preliminary data.

1:15

2pUW2. Frequency analysis of broadband acoustic signals propagating through nonlinear internal waves. Scott Frank, William L. Siegmann (Rensselaer Polytechnic Inst., Dept. of Mathematical Sci., Troy, NY 12180, franks2@rpi.edu), Mohsen Badiy, Yongke Mu (Univ. of Delaware, Newark, DE 19716), James F. Lynch (Woods Hole Oceanogr. Inst., Woods Hole, MA 02543), and Stephen Wolf (Acoust. Div. Naval Res. Lab., Washington, DC 20375)

Nonlinear internal waves along the ocean thermocline are known to affect modal compositions and other properties of acoustic signals. Such nonlinear waves were observed during the SWARM 95 Orthogonal Propagation Experiment, where extensive data were collected from two broadband sources. One source was a 30-s linear frequency modulated sweep signal, and the other was a 0.1-s airgun shot. Both sources were repeated every minute and received by two vertical linear arrays several kilometers away. Sound speed profiles were measured near the source and the receiver arrays simultaneously. The influences of internal waves on the broadband frequency behavior of received and simulated signals are described. Features of energy spectral density and of transmission loss versus frequency plots are used to detect resonant frequencies where anomalous loss occurs. A time-frequency analysis of received signals at several ranges is performed to determine modal propagation characteristics, including modal transitions of the signals passing through the internal waves. An objective of this study is to link acoustic signal behavior to internal wave properties such as wavelength, packet length, and amplitude. [Work supported by ONR.]

1:30

2pUW3. Temporal coherence of matched-field, matched-mode, and matched-beam correlation as a function of frequency and range in the Sicily Strait. T. C. Yang (Naval Res. Lab., Washington, DC 20375), Martin Siderius (SACLANTCEN, La Spezia, Italy), and Kwang Yoo (Naval Res. Lab., Washington, DC 20375)

Ocean medium fluctuations reduce the temporal coherence of acoustic signals. Temporal coherence limits the time duration a replica field can be used for matched-field, matched-mode, and matched-beam source localization in a random ocean. During the Advent 99 experiment, which took place in the Sicily Strait in May 1999, narrow-band signals (12 tones in the band 200–1500 Hz) were transmitted from a bottom-mounted tower and received on a 64-phone line array covering the water column; the water depth was 80 m. The data were collected at three ranges: 2, 5, and 10 km. Using a reference signal as the replica, temporal coherence was measured for the vertical array in the phone domain (matched-field correlation), mode domain (matched-mode correlation), and in the beam domain (matched-beam correlation). For the latter two cases, coherence is a function of the modes and beams used. This paper presents the preliminary results of the data analysis and comparison with acoustic simulations based on an internal wave model. Implications for signal processing will be discussed. [Work was supported by ONR and SACLANTCEN.]

1:45

2pUW4. On the invariance of intensity striation patterns: Effect of shallow-water internal waves. Daniel Rouseff (Appl. Phys. Lab., College of Ocean and Fishery Sci., Univ. of Washington, Seattle, WA 98105, rouseff@apl.washington.edu)

In the second edition of their book, Brekhovskikh and Lysanov [*Fundamentals of Ocean Acoustics* (Springer-Verlag, Berlin, 1991)] introduced the concept of intensity invariance to a larger audience. They showed how contour plots of acoustic intensity, mapped in range and frequency, would show definite striation patterns. A single scalar parameter called beta could characterize these patterns. Referencing a large body of Russian literature, they claimed that beta was invariant to the details of the environment, and that in shallow water it equaled one. The published Russian work was primarily analytical; consequently, it concentrated on relatively simple environments. In the present paper, the concept of intensity invariance is examined by numerical simulation for more complicated scenarios. Realizations of time-evolving shallow-water internal wave fields are generated. Acoustic propagation through the internal waves is simulated using the parabolic equation. Both random background internal waves and more eventlike solitary waves are considered. Beta is estimated from images of intensity using radon transforms and tracked as the internal wave field evolves. The effect of varying bottom attenuation is quantified. Receiver depths both above and below the thermocline are considered. [Work supported by ONR.]

2:00

2pUW5. Analysis of internal wave interactions with broadband acoustic signals propagating along the shelf during the SWARM'95 experiment. Yongke Mu, Mohsen Badiy (College of Marine Studies, Univ. of Delaware, Newark, DE 19716, yongkee@udel.edu), William L. Siegmann, Scott Frank (Rensselaer Polytechnic Inst., Troy, NY 12180), James F. Lynch (Woods Hole Oceanographic Inst., Woods Hole, MA 02543), and Stephen N. Wolf (Naval Res. Lab., Washington, DC 20375)

A multi-institution experiment was conducted in 1995 on the New Jersey continental shelf to assess the Shallow Water Acoustic propagation in Random Media (SWARM'95). Two major source/receiver geometries were established. One was parallel to the continental shelf edge and the other was perpendicular. The orthogonal propagation consisted of two broadband sources transmitting signals perpendicular to the direction of internal wave field. The first signal was a transient airgun source (0.1-s duration), while the second was a linear frequency modulated sweep (30-s duration). Placed above and below the thermocline, these sources were transmitting signals every minute for a few hours at one location. Here we present a 2-h segment of these observations during 4 August 1995 when nonlinear internal waves were present. The transmissions were received by two vertical line arrays at different ranges and angles from the source. The dependence on the azimuth between the internal waves and the acoustic transmissions was examined. It is found that sound speed fluctuations induced by internal waves cause a 10–12-min temporal variations in the intensity of received signals. These temporal variations are azimuthally dependent since the environment is anisotropic due to internal solitary waves. [Work supported by ONR.]

2pUW6. Effect of transverse solitons on acoustic transmission. Gareth M. Houk, Stephen N. Wolf (Naval Res. Lab., Washington, DC 20375), and Mohsen Badiey (Univ. of Delaware, Newark, DE 19716)

Effects of solitary internal waves (solitons) on acoustic signals propagating transversely to the solitons were studied by transmitting 60- to 200-Hz LFM pulses along a 15-km path in shallow ocean water during periods of high and low internal wave activity. The data were collected as part of the SWARM95 experiment on the New Jersey continental shelf. Acoustic normal-mode analysis reveals amplitude fluctuations in the received acoustic modes in the presence of a series of transverse internal waves. Contemporaneous measurements without transverse solitons lack such fluctuations. The acoustic environment and internal wave field were modeled using data from thermistors, CTDs, a high-frequency acoustic flow visualization system, and ship-borne radar images. Theoretical studies [J. C. Preisig and T. F. Duda, *IEEE J. Ocean Eng.* **22**(2), 256–269 (1997)] predict that 2D range- and depth-dependent models are insufficient to describe the mode amplitudes for acoustic scattering by highly transverse waves. In order to test this prediction, the data are compared to 2D and 3D parabolic equation solutions in data-driven simulated environments. [Work supported by ONR.]

2pUW7. Shallow-water acoustic propagation experiments in the Florida Straits—Sensor installation and environmental data. Neil J. Williams, Harry A. DeFerrari, Hien B. Nguyen (Univ. of Miami, Rosenstiel School of Marine and Atmospheric Sci., 4600 Rickenbacker Cswy., Miami, FL 33149), and William A. Venezia (Naval Surface Warfare Ctr., Dania, FL 33004)

Measurements of shallow-water acoustic transmissions from a moored acoustic projector to a fixed triaxial array of acoustic hydrophones have begun off the southern coast of Florida in 145 m of water. The projector transmits *M*-sequence coded signals at frequencies ranging from 100 to 3200 Hz. Source-to-receiver ranges from 10 to 100 km in 10-km steps are planned. The receiving array is powered from shore and linked via a fiber-optic cable. It contains two 500-m-long sections deployed on the bottom and a third moored vertically with a pressure sensor to monitor hydrophone depth. Each section contains 32 hydrophones spaced in a nonlinear distribution to facilitate acoustic coherence measurements. Linux computers control the projectors and process data from the arrays. The computers can be accessed via TCP/IP connections and programmed from shore. Two environmental monitor moorings were deployed along the path from source to receiver. They contain arrays of temperature, conductivity-temperature, and temperature-pressure recorders. An extensive set of CTD casts was also made. Detailed geophysical measurements are planned for the spring—including cores and high-resolution bottom and sub-bottom profiling. A companion paper will discuss the acoustic data collected and preliminary results. [Work supported by ONR.]

2pUW8. Shallow-water acoustic propagation experiments in the Florida Straits—preliminary results. Harry DeFerrari, Neil Williams, Hien Nguyen (Univ. of Miami, Rosenstiel School of Marine and Atmospheric Sci., 4600 Rickenbacker Cswy., Miami, FL 33149), and William Venezia (Naval Surface Warfare Ctr., Dania, FL 33004)

Acoustic propagation experiments have begun on a newly developed shallow-water range off the south Florida coast. The measurements are concerned with fluctuations and coherence of signal and reverberation fields for ranges out to 100 km and for frequencies from 100 to 3200 Hz. The centerpiece of the experimental site is a set of three receiver arrays, two horizontal, 500 m in length, and one vertical array that is suspended in 148 m of water. A total of 96 hydrophones are processed *in situ* and data are transmitted to a shore station via fiber optics cable. The arrays are used to receive transmissions from a multifrequency moored source that transmits broadband *M*-sequences at each of six center frequencies from 100 to 3200 Hz. A first source deployment at a range of 10 km is complete and data were collected continuously for 28 days along with detailed measurements of sound-speed profile, bathymetry, and geoacoustic properties of the bottom and sub-bottom along the propagation path. Subsequent source moorings at longer ranges are underway. Results from the 32-element vertical array are presented here and compared with model predictions. Coherence and fluctuations of the signal and reverberation fields are computed and discussed. [Work supported by ONR.]

2pUW9. Acoustic propagation experiment in the East China Sea. Ma Li, Gao Tianfu, Wang Lan, Xiao Tian, and Xu Wugang (Inst. of Acoust., Chinese Acad. of Sci., Beijing, PROC)

An acoustic propagation experiment has been conducted recently in the shallow water of the East China Sea. The water depth of the experiment area is about 50 m. A receiving vertical array of 20 equal spacing hydrophones spanning the whole water column was suspended from one ship. The broadband explosive sources (7 and 40 m) were deployed by another moving ship. The transmission losses along three directions are recorded and the propagation range is from 1 to 20 km measured by GPS. The temperature chain was used to obtain the sound-speed profile data at the receiving ship over the duration of the experiment time. The temperature data show clearly that there is a warm water mass around the mid-depth ~25 m, and consequently caused a “double channel” in this shallow water. This warm water mass is likely related to the warm current of the Kuroshio. Significant impacts on acoustic propagation are observed: (1) a sound propagation shadow can be found in the mid-depth; (2) anomaly transmission loss can be found in a certain frequency band. Numerical simulation has been done by running PE with Pade approximation to make a comparison and to illustrate the impact.

Session 3aAA**Architectural Acoustics: Acoustical Design of Learning Spaces**

Brandon D. Tinianow, Chair

*Acoustical Laboratory, Johns Manville, 10100 West Ute Avenue, Littleton, Colorado 80162****Invited Papers*****9:00****3aAA1. Acoustical design of learning spaces.** Paul Tan (Pelton Marsh Kinsella, 1420 W. Mockingbird Ln., #400, Dallas, TX 75247, tanpl@c-b.com)

Learning spaces in the past 100 years have evolved from chalkboards and printed media to real-time digital distant learning centers equipped with the latest in communications and multimedia technologies. As a result, the modern learning space bears little semblance with its archetypal shoe-box classroom most are accustomed to. What are the criteria for designing a fully integrated learning environment, wherein human factors and technological systems function symbiotically within a carefully optimized space? This paper will seek to explore the impact of room acoustics, sound isolation, building systems, and equipment noise on the functionality and success of the modern learning environment through selected design cases.

9:20**3aAA2. Classroom acoustics: The effects of background noise and room-finish materials on speech intelligibility.** Gary Siebein (Dept. of Architecture, Univ. of Florida, P.O. Box 115702, Gainesville, FL 32611-5702)

This paper will review recent research conducted at the University of Florida defining conditions in actual school classrooms that contribute to speech intelligibility. Many classroom settings were observed to determine how communication paths among teachers and students occurred in modern classrooms. A survey of classrooms was conducted with measurements of background noise levels, STI, and reverberation time made in the rooms at locations corresponding to those found in actual rooms. A computer model and a physical model of a typical classroom were constructed to further study classroom acoustic situations. A second-order curve was found relating RASTI to background noise levels (as a result of air-conditioning system noise) and distance from the teacher. The effects of room-finish materials played a secondary role in increasing RASTI once background noise levels of NC 32 or less were achieved.

9:40**3aAA3. Designing and building for quiet in a school for deaf children.** John Guenther, Marcus Adrian (Mackey Mitchell and Assoc., St. Louis, MO), J. T. Weissenburger (Eng. Dynam. Intl., Inc., St. Louis, MO), and William Clark (Central Inst. for the Deaf, 818 S. Euclid, St. Louis, MO 63110)

Central Institute for the Deaf (CID) recently completed construction of a new 42 000-sq. ft. school for deaf children. High priority was placed upon designing and building a facility that would provide ideal acoustic environments that fostered learning and auditory/oral communication for students wearing powerful hearing aids or cochlear implants and teachers. A team composed of scientists, architects, and acoustical engineers was assigned the task of designing and building a school that would provide classroom environmental levels at or below the NC 20 contour, interclassroom attenuation exceeding 50 dB, reverberation times on the order of 0.4 s, and sound reinforcement for teachers' voices when facing the blackboard. In group spaces and in the hallways, higher noise levels and longer reverberation times were sought to provide students with experiences more like those faced in the real world. Challenges included a site bounded by a busy interstate highway and a medical center heliport. The team developed and implemented numerous unique acoustic treatments for the facility which are reviewed in the presentation. Although designed as a school for the deaf, the approaches are useful for designing any educational classroom environment. The school opened on 10 January 2000 and met all acoustic criteria.

10:00**3aAA4. Acoustical modeling and auralization as a design tool for a university concert hall renovation.** Bennett M. Brooks (Brooks Acoust. Corp., 27 Hartford Turnpike, Vernon, CT 06066, bbrooks@brooks-acoustics.com)

A 400-seat university concert hall was notorious for its bad acoustics. Music faculty identified the worst problem with this educational facility as poor stage communication, or the difficulty that performers had hearing one another. The hall was modeled using the CATT-Acoustic software system. Good agreement was found between objective measurements and model predictions for the original hall. Also, the subjective quality of the auralized model was judged to be an accurate representation of the original hall

by the faculty. These results provided confidence in the model as a design tool for hall renovations. Design decisions were made based on objective parameters, such as early reflection density, and also on the subjective quality of the sound, for a series of model configurations. The final design was largely realized and the renovated hall has been well received by students, faculty, and community. Sound samples illustrating auralized design options will be presented.

Contributed Papers

10:20

3aAA5. Seeking improved speech intelligibility in a university classroom. Dean R. Heerwagen and Paul D. Sampson (Univ. of Washington, Box 355720, Seattle, WA 98195-5720)

Speech intelligibility has been studied while systematically modifying an existing classroom. The "original" classroom had acceptable background noise levels, but excessive reverberation. Room modifications included installing a lowered ceiling and successive additions of absorption to the classroom walls. The unoccupied rhyme tests (RTs) at 1 kHz progressed from an "original" 1.10 s to a "final" 0.53 s. Modified rhyme tests (MRTs) were also conducted with a volunteer group of faculty and staff. For each room condition, MRTs were administered across a range of signal-to-noise ratio (SNR) conditions. The principal parameters of this study were the physical conditions of the room (as indicated by RT and U50 measures), the test scores, the SNRs for the tests, distances between a loudspeaker and each volunteer, whether the volunteer was a native English speaker, and whether the volunteer was normal hearing or hearing impaired. Statistical analyses of the data indicate that the most important determinant of test performance was the SNRs. The correlation between test scores and RTs is also significant, but less so. Additionally, correlations between test scores and native and non-native speakers and between test scores and those with and without hearing impairment were also significant.

10:35

3aAA6. Further acoustical analysis of infant/toddler rooms in daycare centers. Matthew V. Golden (Grad. Prog. in Acoust., Penn State Univ., P.O. Box 30, State College, PA 16804, golden@sabine.acs.psu.edu) and Tom Frank (Penn State Univ., University Park, PA 16802)

At the last ASA meeting, preliminary data was shared on the background noise in daycare centers. At this meeting, further research in the acoustical analysis of infant/toddler rooms in seven daycare centers will be discussed in relation to speech communication. Information concerning

the collection and analysis of ambient noise, signal-to-noise ratios (SNRs), reverberation times (RT), and early decay time (EDT) will be shared. Comparisons will be made between the measurements and data collection techniques of ambient noise from dosimeters, sound level meters, and other long-term data acquisition techniques. Comparisons of the overall and octave band RTs and EDTs from each of the rooms in the study will be shown. Finally, the calculation of SNRs will be shown for both steady-state and dynamic signals. The steady-state SNR uses several different speech spectra compared to the occupied 1/3-octave band noise levels. The dynamic SNRs are calculated using microphones placed at both the caregiver and the infant or toddler. [Work supported by PHS/NIH (1-R01-HD31540-01A2) Otitis Media, Behavior and Attention in Daycare.]

10:50

3aAA7. A classroom acoustic model to evaluate prescriptive options to meet a performance standard. Richard D. Godfrey (Integrex, Bldg. 75, 2790 Granville Rd., Granville, OH 43023, dick.godfrey@owenscorning.com)

A standard to prescribe the acoustical performance of classroom spaces is now under development. As drafted, the standard will contain performance requirements, and many members of the working group would like to include prescriptive requirements as well. In order to make these two approaches consistent, an acoustical model of the classroom space is needed to predict the effects of component performance on the overall acoustic performance of the space. A model based on classical acoustics has been developed which allows the designer to select components performance characteristics from menus of measured performance. These input data are entered into an energy balance which predicts the classroom sound pressure level as a function of position in the room and the reverberation time. These performance characteristics are then compared to various metrics being considered by the working group. In this paper the formulation of the model will be described, and proposed prescriptive options will be evaluated for consistence with the performance metrics.

3a THU. AM

THURSDAY MORNING, 1 JUNE 2000

GEORGIAN ROOM, 8:30 A.M. TO 12:05 P.M.

Session 3aAB

Animal Bioacoustics, Psychological and Physiological Acoustics and Speech Communication: Cognitive Aspects of Complex Sound Perception in Animals

Robert J. Dooling, Chair
Psychology Department, University of Maryland, College Park, Maryland 20742

Chair's Introduction—8:30

Invited Papers

8:35

3aAB1. Complex sound source determination: Behavioral studies on goldfish. R. R. Fay (Parmly Hearing Inst. and Dept. of Psych., Loyola Univ. of Chicago, Chicago, IL 60626)

A series of experiments is reported on complex sound perception in goldfish using classical respiratory conditioning and a stimulus generalization paradigm. In general, animals are initially conditioned to a simple or complex sound, and then tested for generalization to novel sounds that differ systematically from the conditioning sound along one or more stimulus dimensions. Generalization

gradients to the novel sound set indicate what sounds are perceived as similar or dissimilar, and what features of sounds are analyzed or attended to during conditioning and generalization testing. So far, these kinds of experiments have demonstrated perceptual dimensions in goldfish that resemble the dimensions of pure tone pitch, complex pitch, timbre, and roughness as perceived by humans. In addition, several experiments have revealed auditory stream segregation in goldfish that resembles that demonstrated for human listeners. It is concluded that goldfish share with humans a sense of hearing that is primitive for vertebrates, and is likely shared with most other vertebrate species. At a qualitative level, at least, goldfish know what we know about sounds and their sources. [Work supported by the NIH, NIDCD.]

9:00

3aAB2. Target tracking by echolocation in a dynamic auditory scene. Cynthia F. Moss (Dept. of Psych., Prog. in Neurosci. and Cognit. Sci., Univ. of Maryland, College Park, MD 20742) and Annemarie Surlykke (Odense Univ., DK-5230 Denmark)

The echolocating bat transmits ultrasonic vocalizations as it flies, and it perceives a dynamic, 3-D representation of the world from echoes of its sonar signals. We hypothesize that the bat's auditory system organizes dynamic echo input to segregate and track multiple reflecting sources in the environment, and these perceptual processes depend on the features of the animal's sonar signals. Our research here focuses on the bat's timing of sonar cries to enhance spatial tracking of objects in a dynamic auditory scene. Using two distinct behavioral methods, we quantified features of the bat's vocal signals in target tracking tasks. In a psychophysical study, bats were trained to track changing echo delays of phantom targets. The bat's sonar cries were digitized, electronically delayed, and played back to the animal, simulating targets that approached or disappeared at different rates. Features of the animal's sonar vocalizations were measured for each trial and summarized across conditions. In another study, high-speed synchronized video and audio recordings provided measures of the bat's sonar behavior as it tracked and intercepted tethered insects that were either stationary, moving smoothly, moving erratically, displaced abruptly, or positioned near obstacles. The data illustrate acoustic control for tracking in a dynamic auditory scene.

9:25

3aAB3. The evasive responses of praying mantids and tiger beetles to complex ultrasonic pulse trains produced by echolocating bats. David D. Yager (Dept. of Psych., Univ. of Maryland, College Park, MD 20742, dy5@umail.umd.edu)

Two highly visual diurnal insect predators, praying mantids and tiger beetles, become the prey of echolocating bats at night. Although very distant phylogenetically, the two insects have evolved similar ultrasound-triggered behavioral responses to help them evade capture. The tiger beetle ears have a conventional tympanate design with sharp tuning at 30–35 kHz and directional capabilities. The mantis peripheral auditory system is highly unconventional with a single cyclopean ear in the metathorax that apparently sacrifices directionality for increased sensitivity. There are six to eight pairs of ascending auditory interneurons in the mantis thorax. Even though it slows response time, auditory processing in the brain is required for the evasive behavior, i.e., it is not a simple reflex. The tiger beetles behavioral response to ultrasound is similar, but there appears to be no system of large ascending interneurons. In both insects, the nocturnal ultrasound-triggered evasive behaviors probably evolved from diurnal defensive displays triggered by visual stimuli. Thus, the evolutionary innovation may not have been extensive new circuitry, but, rather, a context-dependent switch that links the behavior to vision during the day and to hearing at night.

9:50

3aAB4. Discrimination of temporal features in complex sounds by the bullfrog, *Rana catesbeiana*. Andrea M. Simmons (Dept. of Psych., Brown Univ., Providence, RI 02912)

The advertisement calls of the sympatric species the bullfrog (*Rana catesbeiana*) and the green frog (*Rana clamitans*) differ in duration and first harmonic periodicity. Sensitivity of male bullfrogs to complex sounds differing in these acoustic attributes was examined in the field using the evoked-calling technique. In experiment 1 (duration series), males were presented with a series of synthetic signals, all with the spectral structure of the bullfrog's advertisement call, but varying in duration along a continuum from bullfrog-like to green-frog-like. In experiment 2 (pitch series), signals were of bullfrog-like duration, but varied in first harmonic periodicity in steps from bullfrog-like to green-frog-like. The number and latencies of the males' evoked vocal responses differed significantly to the exemplars within both series. The forms of the discrimination functions suggest that males perceived these acoustic features as distinct categories. Discrimination was invariant with sound-pressure level over a 30-dB range. These data show that bullfrogs use temporal features to distinguish between conspecific and heterospecific vocalizations, and may achieve this by a process similar to categorical perception as defined in other animals.

10:15–10:30 Break

10:30

3aAB5. Why macaque screams differ. Harold Gouzoules and Sarah Gouzoules (Dept. of Psych. and Yerkes Primate Ctr., Emory Univ., Atlanta, GA 30322, psyhg@emory.edu)

Screams of four species of macaques (*Macaca mulatta*, *M. nemestrina*, *M. nigra*, *M. arctoides*) were compared for similarities and differences with respect to predictions of Morton's motivation-structural rules [Morton, Am. Nat. **111**, 855–869 (1977)]. Screams from victims of attack that involved contact aggression (pulling, pushing, slapping, grappling, and biting) from a higher-ranking opponent were examined. For each species, 100 screams from females three years of age or older were digitized and acoustic features

of each call measured. Discriminant function analysis was used to determine whether or not the 400 vocalizations could be assigned to the correct caller species on the basis of their acoustic structure. Calls were assigned to the correct species at a significantly higher rate (93.5%) than expected by chance. Each of the four macaque species used acoustically distinct screams in a shared context. While the differences in the species' vocalizations suggest no simple correlation between immediate context and the acoustic forms of screams, there was general correspondence between the structure predicted by motivation-structural rules and inferences about the internal state of the vocalizer derived from the typical intensity of aggressive patterns that characterize each of the four species.

10:55

3aAB6. Contributions of nonhuman animal models to understanding human speech perception. Keith R. Kluender (Dept. of Psych., Univ. of Wisconsin, 1202 West Johnson St., Madison, WI 53706)

Broadly speaking, nonhuman animal models contribute to understanding speech perception by humans in two ways—by analogy and by homology. The former is generally easier and examples are more abundant. Because demonstrating homology requires deeper explication of underlying mechanisms, claims can be more precarious but carry potentially greater explanatory payoff. When studying nonhuman organisms as an analogy, the emphasis is typically upon how animal physiological or behavioral processes have adapted to fulfill requirements of particular ecological niches. By contrast, study of animals as homology often violates ecology in search of common underlying processes, and the animal becomes a method more than an object of study. Examples of findings for animal analogies and homologies will be reviewed. Data will be presented from experiments in which nonhuman subjects play the role of homology in revealing both foundational sensory processes and more plastic processes of perceptual development. Animal models provide advantages for describing sensory representation of speech unadulterated by effects of experience, and animal models allow control over experience permitting better characterization of processes through which experience shapes perception. Nonhuman animal models are providing important insights into general processes of audition and learning essential to human speech perception. [Work supported by NIH and NSF.]

Contributed Papers

11:20

3aAB7. 3D models of lateral line excitation patterns during prey acquisition by Lake Michigan mottled sculpin. Sheryl Coombs (Parmlly Hearing Inst., Loyola Univ. of Chicago, IL 60626) and James Finneran (SPAWAR Systems Ctr., San Diego, CA 92152-6505)

One of the biggest impediments to understanding how fish acquire information through their lateral line sensory system is the inherent difficulty in adequately specifying and measuring hydrodynamic stimuli, which can vary in frequency, amplitude, and phase along the sensory surface of the fish. Even theoretically simple stimuli, like dipole sources, can create rather spatially complex water motions. The peripheral excitation pattern after the stimulus field has been transduced can be even more complex and depends on a number of factors, including the density of the fish and the amplitude and direction of its body motion in the field, the number, type, and spatial distribution of lateral line end organs on the fish, the axis of source vibration, and the distance and orientation of the fish with respect to the source. We have developed a 3D computational model that takes these factors into account and is used to (1) compare the pre-transduction distribution of the stimulus (flow) field along the sensory surface of the fish to the post-transduction excitation pattern of primary afferent fibers and (2) describe how excitation patterns change during actual approach pathways taken by sculpin when responding to preylike (dipole) sources.

11:35

3aAB8. Cortical responses in rats to periodic frequency-modulated sounds. Itzel Orduna, Eduardo Mercado III, Daphna Shohamy, Mark A. Gluck (CMBN, Rutgers Univ., 197 University Ave., Newark, NJ 07102), and Michael M. Merzenich (Coleman Lab., UCSF, San Francisco, CA 94143)

Spectral variations over time are a common characteristic of naturally occurring sounds. They constitute a prevalent feature in communication signals in several species. Studies of mammals have shown that auditory cortex neurons respond to single frequency-modulated (FM) sweeps and

that most responses are selective for sweep direction and/or rate. Researchers have also used trains of FM sweeps to estimate spectrotemporal receptive fields in auditory cortex. In the present study, microelectrode recordings were used to explore how the auditory cortex responds to trains of FM sweeps in anesthetized rats. Maps of 20–60 penetrations were made for each subject. Sweep frequencies ranged from 1–16 kHz with FM rates ranging from 4–24 octaves/s and repetition rates from 2–24 sweeps/s. Both down-sweeps and up-sweeps were presented. Neuronal responses were analyzed in terms of onset, offset, oscillatory and directionally selective properties. Several types of responses predominated. Most units responded to sound onset. In contrast, responses to signal offsets were rare. Oscillatory responses were typically evoked only for repetition rates less than 12 sweep/s. Directionally selective responses were limited to these rates. These results indicate that oscillatory responses in the auditory cortex are limited to low repetition rate stimuli, and that directional selectivity depends on repetition rate. [Work supported by NSF.]

11:50

3aAB9. Plasticity of spectrotemporal sensitivities in auditory cortex. Eduardo Mercado III, Daphna Shohamy, Itzel Orduna, Mark A. Gluck (CMBN, Rutgers Univ., 197 University Ave., Newark, NJ 07102), and Michael M. Merzenich (Coleman Lab., UCSF, San Francisco, CA 94143)

Response characteristics of auditory cortex can be altered by repeatedly pairing sounds with basal forebrain stimulation [M. P. Kilgard and M. M. Merzenich, *Science* **279**, 1714–1718 (1998)]. Although many neurons in auditory cortex respond most strongly to time-varying sounds, most studies of stimulation-induced plasticity have focused on changes in responses to tone pips. We examined stimulation-induced changes in neuronal sensitivities to frequency-modulated sweep trains (bandwidth=2–16 kHz, duration=1 s, sweep rates=4–24 octaves/s, repetition rates=2–24 sweeps/s). Adult rats received electrical stimulation of basal forebrain paired with 1–10 varieties of sweep trains, 300–500 times per day, for 9–16 days. Some sounds were presented in combination with bandlimited

Gaussian noise. After stimulation, neuronal responses were recorded from 20–80 sites in the auditory cortex of each rat. The spectrotemporal sensitivities of auditory cortical neurons were dramatically altered in stimulated rats. Changes in response characteristics were not straightforwardly related to features of the sounds that had been paired with stimulation.

Certain sounds that normally evoked responses correlated with sweep repetition rate in control rats tended to evoke either aperiodic or uncorrelated periodic responses from neurons of stimulated rats. Current theories of auditory cortical plasticity do not account for these results. [Work supported by MBRS, NIMH, NSF.]

THURSDAY MORNING, 1 JUNE 2000

ENGLISH ROOM, 9:00 TO 11:00 A.M.

Session 3aAO

Acoustical Oceanography: Comparisons of Acoustical and Conventional Measurements of Ocean Thermohaline Structures

Brian D. Dushaw, Chair

Applied Physics Laboratories, University of Washington, 1013 NE 42nd Street, Seattle, Washington 98105-6698

Chair's Introduction—9:00

Invited Paper

9:05

3aAO1. Tomographic maps of the New England Shelfbreak Front. Ching-Sang Chiu (Code OC/Ci, Dept. of Oceanogr., Naval Postgrad. School, Monterey, CA 93943 chiu@nps.navy.mil), James F. Lynch, and Glen Gawarkiewicz (Woods Hole Oceanogr. Inst., Woods Hole, MA 02543)

In July 1996, an integrated acoustic-oceanographic experiment was carried out in the Mid Atlantic Bight south of New England to study the oceanographic structure of the shelfbreak front and its effects on sound propagation. The experiment employed a suite of acoustic and oceanographic sensors including sound sources on the slope, vertical hydrophone arrays on the shelf, and a SeaSoar (towed CTD) that provided several volumetric surveys of the frontal zone. In order to obtain accurate tomographic maps for this complex shelf-slope environment, an inverse scheme that can properly handle strong mode coupling was developed. In the formulation, a variation in the modal travel time is expressed as a function of the evolution history and coupling history of the modes associated with the “background” ocean and changes in these histories due to sound speed changes. The inversion of the observed modal travel times is accomplished using nonlinear least-squares estimations. In this presentation, the quality of these tomographic maps is discussed using results from a resolution-variance analysis. Additionally, the data and resolution kernels of the tomographic measurements are compared to those of the SeaSoar measurements to illustrate that the two independent data sets are highly complementary to each other.

Contributed Papers

9:35

3aAO2. Acoustic monitoring of oceanographic features and processes in the Arctic Ocean. Konstantin Sabinin (N. N. Andreev Acoust. Inst., Shvernik St., Moscow, 117036, Russian Federation) and James Lynch (Woods Hole Oceanogr. Inst., Woods Hole, MA 02543)

Acoustic transmissions can be used to monitor two of the most important features of the Arctic Ocean, namely, the heat content of the Atlantic water and the upper ocean salinity. In this presentation, we will first discuss the optimal configuration of an array that might be used for heat content studies. We will then look at the newer concept of monitoring the upper ocean salinity in the Arctic Ocean using acoustics, which we have dubbed “acoustic halinometry.” Both the basic concepts of halinometry and the configuration of a possible experimental array will be discussed. In particular, we will show the results of numerically simulating an acoustic halinometry experiment using data from large-scale Arctic oceanographic surveys conducted during 1973–1979. We will also look at the integration of the acoustics with conventional measurements in a “combined monitoring” array concept. This combined monitoring concept will be applied via simulation to one of the primary connections to the Arctic Ocean, the Bering Strait.

9:50

3aAO3. A comparison of acoustic thermometry, XBT, TOPEX, and HOT observations of ocean temperature in the northeast Pacific. Brian D. Dushaw and the ATOC Group^{a)} (Appl. Phys. Lab., Univ. of Washington, 1013 NE 40th St., Seattle, WA 98105-6698)

Time series of temperature have been measured acoustically in the northeast Pacific as part of the Acoustic Thermometry of Ocean Climate (ATOC) project. These time series are compared with other available data types. The acoustic time series of transmissions from the California and Kauai acoustic sources were obtained during 1996–1999. As a result of marine mammal protocols, the time series are intermittent; the California source was turned off in Fall 1998. Assuming that variations in sea-surface height observed by TOPEX/POSEIDON are caused by thermal expansion, the amplitude of the annual cycle of heat content derived from altimetry is larger than that found by the acoustic data, Levitus climatology, and monthly maps of ocean temperature from XBTs of opportunity. The heat content “anomalies” determined by the XBT maps are comparable in size to the differences between the XBT and acoustically derived heat content. A variety of problems with the XBT sampling may account for these

differences. The 12-year time series of temperature derived from the Hawaiian Ocean Time series (HOT) data highlights the mesoscale noise in single-point sampling. However, thermal variability at 100-day time scales is observed in the acoustic data obtained between Hawaii and California using the Kauai source. Acoustic thermometry is complementary to altimetry and hydrography. ^{a)}The Acoustic Thermometry of Ocean Climate

(ATOC) Group is: A. B. Baggeroer and C. Wunsch (MIT); D. Menemenlis (JPL); T. G. Birdsall, K. Metzger (Univ. of Mich.); C. Clark (Cornell Univ.); J. A. Colosi (WHOI); B. D. Cornuelle, M. Dzieciuch, W. Munk, P. F. Worcester (SIO); D. Costa (Univ. of Calif., Santa Cruz); B. D. Dushaw, B. M. Howe, J. A. Mercer, R. C. Spindel (APL-Univ. of Wash.); A. M. G. Forbes (CSIRO, Hobart).

10:05–11:00

Panel Discussion

THURSDAY MORNING, 1 JUNE 2000

FRENCH ROOM, 8:30 TO 10:45 A.M.

Session 3aBB

Biomedical Ultrasound/Bioresponse to Vibration: Lithotripsy

Robin O. Cleveland, Chair

Aerospace and Mechanical Engineering, Boston University, 110 Cummings Street, Boston, Massachusetts 02215

Contributed Papers

8:30

3aBB1. Study of weak underwater shock waves generated by Ho:YAG pulse laser beam. S. Hamid, R. Hosseini (Shock Wave Res. Ctr., Inst. of Fluid Sci., Tohoku Univ., 2-1-1 Katahira, Aoba, Sendai 980-8577, Japan, hosseini@ceres.ifs.tohoku.ac.jp), Takayuki Hirano, Osamu Onodera, and Kazuyoshi Takayama (Tohoku Univ., Sendai, Japan)

For medical application of underwater shock waves as a less-invasive approach, a reliable micro shock wave source is required. The present paper reports progress in production of underwater micro shock waves by direct irradiation of pulse laser beam. Energy source was a Q-switched Ho:YAG laser (Nippon Infrared Industries Co. Ltd.) with 91 mJ/pulse energy measured at the end of a 0.60-mm diameter glass optical fiber, pulse duration of 200 ns, and wavelength of 2.1 μm . The laser beam was focused and transmitted through the optical fiber. The generation and propagation of underwater shock waves from the roughened end of the fiber were quantitatively visualized by double-exposure holographic interferometry. Sequential flow visualization revealed that plasma generated by the laser beam drove spherical shock waves. The diameter of plasma region increased, then reduced and detached from the fiber end. Heat-induced flow in front of the fiber vanished after 100 ms. Peak overpressures were measured at various stand-off distances. The weak shock waves produced by this method have potential to be applied for precise medical procedures such as revascularization in neurosurgery.

8:45

3aBB2. The impact of high-dose lithotripsy on renal structure and function. Lynn R. Willis (Dept. of Pharmacology and Toxicology, Indiana Univ. School of Medicine, 635 Barnhill Dr., Indianapolis, IN 46202, willisl@iupui.edu), Andrew P. Evan, Bret A. Connors, Philip Blomgren, and James E. Lingeman (Indiana Univ. School of Medicine, Indianapolis, IN 46202)

These studies characterized the effects of normal and high “doses” of shock waves on renal structure and function in healthy and compromised kidneys of young, anesthetized pigs. A “normal” dose of 2000 shock waves (24 kV, unmodified HM3 lithotripter) to one kidney produced lesions comprising about 8% of functional renal mass, transiently reduced renal blood flow and glomerular filtration rate (GFR) in both kidneys, and reduced regional perfusion in shocked kidneys. High doses of shock waves to one kidney (8000 shocks at 24 kV) produced larger lesions

(about 14%), intensified the initial vasoconstriction in both kidneys, and sustained the reduction of GFR in the shocked kidneys for at least 24 h after treatment. The administration of 2000 shock waves (24 kV) to pyelonephritic (compromised) kidneys exaggerated the structural/functional impairment by inducing large lesions and intense vasoconstriction normally seen only after 8000 shock waves in healthy kidneys. We conclude that both shock wave number and preexisting renal disease increase the severity of tissue injury and functional impairment produced by shock wave lithotripsy. Since renal ischemia and inflammation occur in association with such injury, subsequent renal scarring and permanent loss of functional renal mass may also be related to shock wave dosage.

9:00

3aBB3. Effect of shock waves on cytoskeleton of human renal cell carcinoma. S. Moosavi Nejad, Makota Satoh, Naomasa Ioritani, and Seiichi Orikasa (Dept. of Urology, School of Medicine, Tohoku Univ., Seiryomachi, Aoba, Sendai, 980-8574 Japan)

The first clinical extracorporeal shock wave lithotripsy (ESWL) in early 1980 revolutionized the surgical management of urolithiasis. Since then, studies have been under way to discover additional uses of ESWL in other areas of medicine including treatment of cancer. However, the shock wave tissue interactions have not yet completely been understood. To analyze *in vitro* subsequent alteration of cell cytoskeleton, cultured human renal carcinoma cells were exposed to the underwater shock waves, generated by an experimental setup with 20 MPa overpressure at the focal region. The cultured cells provided a three-dimensional lattice work of structural proteins (actin, tubulin, and vimentin) analogous to those observed in the solid tumor. Scanning electron microscopic examination revealed morphological changes of the treated cells located in the focal region. Cells were detached from the substratum and changed from the normal fibroblast-like shape into a spherical shape. Disorganization and disassembly of the fibrillar cytoskeletal proteins were observed in the treated cells by immunofluorescence microscopy. The detached cells were collected and recultured again. After 6 h of spreading on a culture dish, reorganization of those fibers in the treated cells was compared with control cells and no significant difference was found.

3a THU. AM

3aBB4. SWL stone fragmentation *in vitro* is improved by slowing the SW delivery rate. David A. Lifshitz, James C. Williams, Jr., Andrew P. Evan, Drew L. Rietjens, James A. McAteer (Dept. of Anatomy and Cell Biol., Indiana Univ. School of Medicine, 635 Barnhill Dr., Indianapolis, IN 46202, mcaateer@anatomy.iupui.edu), Michael R. Bailey, Lawrence A. Crum (Univ. of Washington, Seattle, WA 98105), and Oleg A. Sapozhnikov (Moscow State Univ., Moscow 119899, Russia)

Fast shock wave (SW) rates in lithotripsy (SWL) generate enhanced cavitation that could promote stone fragmentation. We tested the idea that SWL at the high end of clinical SW rate (2 Hz) acts to improve stone comminution. Model stones (Ultracal-30 cement) were exposed to SWs (20 kV, 400 SWs) at 0.2, 0.5, 1, and 2 Hz in a research electrohydraulic lithotripter. Fragmentation was assessed by measuring number, size, and projected surface area of the fragments. Stones treated at 0.2 Hz exhibited significantly greater fragmentation ($p < 0.01$) than stones at 1 or 2 Hz, while fragmentation between 0.2 and 0.5 Hz was similar. Mean \pm SEM for fragment area increase was $370 \pm 53\%$ at 0.2 Hz ($n = 10$ stones), 280 ± 34 at 0.5 Hz (8), 130 ± 31 at 1 Hz (5), and 101 ± 16 at 2 Hz (20). This pronounced enhancement of fragmentation at very slow SW rate was unexpected. High-speed camera images of cavitation at solid objects show an increased bubble cloud at faster SW rates. The bubble cloud may interfere with transmission of acoustic energy to the stone surface. These *in vitro* data suggest the possibility that patient treatment at fast SW delivery rates may decrease the efficiency of stone comminution. [Work supported by NIH P01-DK43881.]

9:30–9:45 Break

9:45

3aBB5. Dynamics of lithotripter shock-wave-induced bubble oscillation in constrained media. Pei Zhong, Yufeng Zhou, and Songlin Zhu (Dept. of Mech. Eng. and Mater. Sci., Duke Univ., Box 90300, Durham, NC 27708, pzhong@acpub.duke.edu)

Rupture of small blood vessels is often observed *in vivo* following shock wave lithotripsy (SWL), and cavitation has been implicated as a potential mechanism for the injury. To understand more precisely the underlying mechanical process of the injury, the dynamics of SWL-induced bubble oscillation in constrained media have been investigated. Silicone tubes (0.3–1.5-mm inner diameter), filled with circulating saline solution containing 0.1% Alunex contrast agent, were immersed in castor oil and placed in the acoustic field of a Dornier XL-1 lithotripter. Bubble dynamics induced inside the silicone tubes were characterized using high-speed shadowgraph and passive cavitation detection via a 20-MHz focused hydrophone. The result shows that the symmetric bubble oscillation, typical for SWL-induced cavitation in water, is largely disrupted due to the constraint of the silicone tube on bubble expansion, leading to an asymmetric elongation of the bubble along the tube axis. The subsequent collapse of the bubbles inside the silicone tube (or blood vessels) is therefore significantly weakened, and thus the potential to cause tissue injury. Additional experiments using hollow cellulose fibers suggest that the damage to small blood vessels may be caused by the rapid, large intraluminal expansion of the bubble. [Work supported by NIH.]

10:00

3aBB6. Bubble translation due to radiation force in SWL. Dahlia L. Sokolov, Michael R. Bailey, Lawrence A. Crum (Appl. Phys. Lab., Univ. of Washington, 1013 NE 40th St., Seattle, WA 98105, dsokolov@apl.washington.edu), and Oleg A. Sapozhnikov (Moscow State Univ., Moscow 119899, Russia)

The clustering of cavitation bubbles may lead to enhanced stone comminution and influence the extent of tissue damage during shock wave lithotripsy (SWL) treatment. Recent research has focused on changing the SWL pulse, or timing between pulses, to intensify or mitigate collapse or

localize these clusters. Such research has targeted radial, not translational motion. We investigate whether bubble translation due to radiation force is sufficiently large to influence cluster formation. The translational dynamics of a single spherical bubble were modeled according to the formulation proposed by Watanabe and Kukita [Phys. Fluids **5**(11) (1993)]. After radius-time data were obtained using the Gilmore equation, translational motion was calculated by numerical integration of the Watanabe equation. Calculations were performed for a range of bubble sizes ($R_0 = 2\text{--}20\ \mu\text{m}$) and pressure rise times ($10^{-9}\text{--}10^{-7}$ s). The results show that, during bubble growth and collapse induced by a single pulse or two pulses with microsecond delays, bubble translations are ~ 0.1 mm. Although bubble translation from a single pulse may not have a noticeable effect on bubble distribution, the effect may be cumulative for the 1000+ shots fired during clinical SWL treatment. [Work supported by NIH DK43881, NSF, CRDF, and FIRCA.]

10:15

3aBB7. Simultaneous detection of acoustic and light emissions from cavitation bubbles in SWL. Thomas J. Matula, Michael R. Bailey, Paul R. Hilmo, and Lawrence A. Crum (Appl. Phys. Lab., Univ. of Washington, 1013 NE 40th St., Seattle, WA 98105, matula@apl.washington.edu)

A typical pulse in electrohydraulic shock wave lithotripsy (SWL) consists of an intense positive pressure pulse, followed by a longer negative-pressure tail. Computer models of the bubble dynamics associated with such a pulse suggest that the positive pressure pulse compresses the bubble ($R_{[0]} = 3\text{--}10\ \mu\text{m}$) to a submicron size. The negative-pressure tail then causes the bubble to undergo a dramatic expansion, followed by an inertially dominated (presumably spherical) collapse hundreds of microseconds later. Acoustic and light emissions are generated at both collapses. We have examined the simultaneous acoustic and optical emission from a cavitation field generated by SWL in order to determine whether the sonoluminescence is principally due to the initial compression of the bubble, or the final inertial collapse. Using two confocal 1-MHz, piezoelectric hydrophones and a PMT mounted on a light-tight water-filled container, we have observed acoustic and light emission corresponding to both the compression and inertial collapse of the bubble field. Our initial results suggest that the light emission occurs most frequently during the initial bubble compression. These results may have implications for understanding the sphericity of the bubble dynamics produced in SWL. [Work supported by NIH and DARPA.]

10:30

3aBB8. Dynamic photoelastic study of the transient stress fields in solids during shock wave lithotripsy. Xufeng Xi and Pei Zhong (Dept. of Mech. Eng. and Mater. Sci., Duke Univ., Box 90300, Durham, NC 27708)

Photoelastic and shadowgraph imaging techniques were used to visualize shock wave propagation, evolution, and the resultant transient stress fields in solids during shock wave lithotripsy. In parallel, theoretical analysis of the wavefront evolution inside the solids was performed using the ray-tracing method. Excellent agreement between the theoretical prediction and experimental results was observed. Moreover, the effects of sample size and shape on wave evolution and associated stress fields induced inside the solids were evaluated, both theoretically and experimentally. Finally, stone fragmentation tests were carried out using stone phantoms of different geometry and size, and the characteristics of the damage patterns were documented. By correlating the stone fragmentation patterns with the behavior of different wave components induced in the target stones, it was found that while the damage cracks near the posterior surface of the sample were initiated by the reflected longitudinal tensile wave, the transmitted shear wave plays a critical role in crack extension. [Work supported by NIH.]

Session 3aEA

Engineering Acoustics: Transducers and Arrays

Sung-Hwan Ko, Chair

College of Engineering, Seoul National University, Seoul 151-742, Korea

Contributed Papers

9:00

3aEA1. Characterization of a cosine-shaped horn. Jorge Arenas and Malcolm Crocker (Dept. of Mech. Eng., Auburn Univ., 201 Ross Hall, Auburn, AL 36849-5341, arenajo@eng.auburn.edu)

Several approaches to the problem of the acoustic wave propagation in a duct with a varying cross section have been described in the literature. It has been shown that a simple one-dimensional analysis gives accurate predictions of the sound propagation. In this research, the impedance of a cosine-shaped horn is obtained using the Wentzel-Kramers-Brillouin (WKB) approximation to solve the Webster wave equation and the results are compared with those obtained from a numerical solution of the equivalent system of nonlinear equations. A comparison of the throat impedance ratio is presented for similar exponential, conical, parabolic, and catenoidal horns with the same overall dimensions. The mouth of the horn is assumed to be terminated in an infinite baffle. In addition, since the device may be used as an impedance transformer, the results for the impedance ratio for two pipes of different areas joined by a cosine-shaped connector are compared with several well-known couplings. Such a cosine-shaped connector will act as a simple discontinuity when its length is short compared with a wavelength and as a transformer for acoustic impedance for higher frequencies.

9:15

3aEA2. Directivity patterns of rectangular pistons on prolate spheroids. J. E. Boisvert and A. L. Van Buren (Naval Undersea Warfare Ctr. Div. Newport, Newport, RI 02841-1708, boisvert@lego.npt.nuwc.navy.mil)

A general expression for the directivity of a rectangular piston arbitrarily located upon a rigid prolate spheroidal baffle is formulated. The piston is assumed to vibrate with uniform normal velocity, and Neumann boundary conditions are imposed in the solution of the problem. The formal solution is expressed in terms of a modal series representation in spheroidal wave functions, valid at any distance from the spheroid. The prolate spheroidal wave functions are obtained using computer programs that have been recently modified and extended to provide accurate values of the wave functions at high frequencies. Substitution of the limiting form of the radial wave functions (as r approaches infinity) into the solution yields the piston far-field single element pattern (SEP). The SEPs for several different piston sizes, orientations, and locations on the spheroidal baffle are presented at various frequencies. [Work supported by ONR Code 321 and Naval Undersea Warfare Center In-house Laboratory Independent Research (ILIR) program.]

9:30

3aEA3. Recent transmit-mode experimental results on thicker injection molded 1-3 piezocomposite. Kim Benjamin (NAVSEA Newport, Undersea Warfare Ctr., Div. Code 2162, Newport, RI 02841)

Three transducers which utilize a new 25.4-mm (1.0-in.)-thick injection molded 1-3 piezocomposite as the active material have been fabricated and acoustically calibrated in the transmit mode. The study, which included a single, double, and quadruple layer assembly of 25.4-mm (1.0-in.)-thick composite, indicates that considerable transmit source levels are possible over a very broad frequency range using this newly developed

thicker material. Furthermore, the element design can easily be packaged into various two-dimensional array configurations. For the multilayer cases, the nulls present in the untuned response were mitigated by applying appropriate phase shifts for each individual layer. Measured results compared well with model predictions for all three cases, indicating that the transducer layering scheme was properly implemented. [This work was funded by NAVSEA NUWC Code 82 and Code 21.]

9:45

3aEA4. Structure-borne noise reduction for an infinite cylindrical shell (theory of elasticity). Sunghwan Ko, Woojae Seong, and Sangwoo Pyo (Dept. of Naval Architecture and Ocean Eng., College of Eng., Seoul National Univ., Seoul 151-742, South Korea)

A theoretical model was developed to evaluate the reduction of structure-borne noise generated by an axially symmetric ring force which is applied on the interior of the cylindrical shell. The vibrating cylindrical shell is coated with a microvoided elastomer that is acoustically soft material designed for the reduction of the generated noise. The analytical model is a two-layer shell structure comprised of an outer layer of coating that is perfectly bonded to a cylindrical shell. The outer surface of the coating and the inner surface of the shell are in contact with water and air, respectively. The analysis for this problem is based on the theory of elasticity and pertinent boundary conditions. Effects of various parameters such as coating thickness and material properties on the noise reductions are presented.

10:00-10:15 Break

10:15

3aEA5. An optimization process for estimating two-port parameters. M. R. Serbyn and A. P. Kumar (Phys. Dept., Morgan State Univ., 1700 E. Cold Spring Ln., Baltimore, MD 21251, rserbyn@morgan.edu)

One process of parameter identification for a linear time-invariant two port consists of solving the inverse problem, given a set of values of the external variables. Such parameter estimates are subject to two sources of error: the measurement uncertainty and the propagation of error. The latter, in the worst case, can greatly exceed the former [M. R. Serbyn, J. Acoust. Soc. Am. **106**, 2232 (1999)], but can be reduced by a judicious choice of the external variables, as implied by previously reported results, for example, [Li-Feng Ge, J. Acoust. Soc. Am. **97**, 324-330 (1995)]. For this investigation a simple electrical circuit, characterized by a Moebius transformation, $w = (a11z + a12)/(a21z + a22)$, with known parameter values, was chosen. It provided a baseline for comparing the parameter values estimated from measurements of z and w made at the output and input ports, respectively. Each estimate was based on a set of three (z, w) values and the reciprocity condition. By varying the nature and magnitude of the load impedance, z , and the corresponding input impedance, w , it was possible to minimize the errors in the computed values of the transformation matrix. Both simulated and experimental results were compared with the theoretical standard.

10:30

3aEA6. An acoustic volumetric array system. Jonathan M. Rigelsford and Alan Tennant (School of Eng., Univ. of Hull, Hull HU6 7RX, UK, J.M.Rigelsford@eng.hull.ac.uk)

A microphone array system based on a three-dimensional random geometry is introduced. Three-dimensional microwave volumetric array antennas have recently been investigated as potential systems for radar and communications applications [A. Tennant and A. F. Fray, "A 64 element broad band volumetric array antenna," IEEE International Antenna and Propagation Symposium, Atlanta, 1998]. In a volumetric array, the elemental receivers occupy a solid region of three-dimensional space. A spherical array geometry is preferred. This provides the potential for full scan coverage. The projected aperture of a spherical array is direction independent, with the beam width and gain of the directional response being constant for all scan angles. The system developed consists of 64 omnidirectional microphones arranged at pseudorandom locations within a spherical volume. The nonperiodic element arrangement has two advantages. First, it eliminates the possibility of large grating lobes, and therefore enables wide-angle beam steering over a very large frequency bandwidth. Second, this geometry results in an array reception pattern with an average side-lobe level that is inversely proportional to the number of array elements. The potential of acoustic volumetric arrays is examined and the results of theoretical and experimental investigations presented. Applications areas include covert surveillance, imaging, and auditorium characterization.

10:45

3aEA7. Combined sound-pressure and pressure-gradient hydrophone design. Boris Aronov, Lawrence Reinhart, and David A. Brown (Acoust. Res. Lab., Dept. of Elec. and Comput. Eng. and the Ctr. for Marine Sci. and Tech., Univ. of Massachusetts, North Dartmouth, MA 02747)

A common design of a sound-pressure hydrophone relies on two bi-morph piezoelectric circular plates symmetrically attached to a common base in the form of a ring or cylinder. This may be considered as having

two inertially coupled partial systems vibrating in symmetrical and anti-symmetrical modes. The symmetrical mode corresponds to the in-phase electrical connection of plates (hydrophone) and provides an output proportional to the omnidirectional sound pressure. The antisymmetrical mode corresponds to antiphase electrical connection of the plates and provides the figure-eight output proportional to the gradient of the pressure field. In order to make the electromechanical frequency responses of partial systems identical, the coupling between these systems must be made as small as possible. This can be achieved by proper design of the common base and achieving immunity to unwanted actions. The sound-pressure hydrophone should be insensitive to acceleration caused by the pressure gradient and the pressure-gradient hydrophone should be insensitive to sound pressure. The properties of a combined hydrophone and the procedure of plate equalizing are investigated. The experimental data obtained are in a good agreement with theoretical expectations.

11:00

3aEA8. Review of ultrasonic transducer calibration techniques at Penn State University Applied Research Laboratory Aaron M. Foulk, W. Jack Hughes, and David Van Tol (Appl. Res. Lab., Penn State Univ., State College, PA 16804, aaronfoulk@hotmail.com)

Recent emphasis on development of higher-frequency transducers for use in medical ultrasound and sonar applications has necessitated the creation of a reliable means of evaluation of these transducers. This work provides an overview of recent progress in the development of a facility for this purpose at the Pennsylvania State University Applied Research Laboratory. [Research funded by ONR.]

THURSDAY MORNING, 1 JUNE 2000

PEACHTREE BALLROOM, 8:30 A.M. TO 12:05 P.M.

Session 3aED

Education in Acoustics: Publishing Excellence in JASA

Allan D. Pierce, Cochair

Acoustical Society of America, P.O. Box 323, East Sandwich, Massachusetts 02537

P. K. Raju, Cochair

Department of Mechanical Engineering, Auburn University, Ross 201, Auburn, Alabama 36849-5341

Chair's Introduction—8:30

Invited Papers

8:35

3aED1. Asking the right questions and following through: Obstacles encountered by prospective authors. Allan D. Pierce (Acoustical Society of America, P.O. Box 323, East Sandwich, MA 02537, adp@bu.edu)

Much of the best work in acoustics is presented at the Society's meetings. The *Journal* benefits if more presenters of such work submit archival-quality manuscripts. Obstacles include writer's block, counter-pressures from employers and others, inadequate support services, poor time organization, distrust (not necessarily unjustified) of the fairness and value of the JASA peer reviewing process, and a lack of appreciation of the long-term value of JASA publication. The Editor advocates expending the effort, gives suggestions on how authors can overcome the obstacles, and presents ideas (and pleas for suggestions from authors) as to how manuscript processing can be improved. The continued increase in quality of those papers eventually published remains the *Journal's*

foremost priority, and authors can help by giving more attention to the selection of questions addressed in their research and by justifying the significance of such questions in their manuscripts. Recognition of the “right question” is an art that can be acquired, but which requires continued effort. Stimulation comes from many sources; the Editor especially recommends extensive reading of current literature, careful listening to talks at the “right” meetings, and a rapport with researchers who have a knack for asking the “right questions.”

9:05

3aED2. Frequent criticisms of submitted manuscripts. Sid P. Bacon (Dept. of Speech and Hearing Sci., Arizona State Univ., Box 871908, Tempe, AZ 85287-1908, spb@asu.edu)

Few manuscripts submitted to the *Journal* for consideration sail through without some critical comments by the reviewers and/or the associate editor. The manuscripts that the present editor sees are in the area of psychological acoustics, and many of the types of criticisms are undoubtedly similar to those made for manuscripts in other areas. A brief list (with sanitized examples) is given of the more common types of criticisms, and a discussion is given as to which of these lead to rejection of the article and as to which can be overcome by a revision. It is advised that authors view criticisms dispassionately and try to understand the substance of each criticism. In some cases, the criticisms are unjustified and the authors should endeavor to give a persuasive rebuttal rather than merely placating the reviewers. A good strategy for a successful author who desires rapid publication is to anticipate possible criticisms and to write the paper so that they will be avoided.

9:35

3aED3. Journal issues of peer review, contribution, and quality. William M. Carey (Dept. of Aersp. and Mech. Eng., College of Eng., Boston Univ., 110 Cummington St., Boston, MA 02215, wcarey@bu.edu)

Allan Pierce, Editor-in-Chief, has recently written concise overviews of the *Journal's* history, publication criteria, and issues. To stimulate discussion, this paper follows his lead and addresses several resolvable issues that could improve the quality of published articles. Concerning the peer review and the reviewer, when should reviewers recuse themselves or disclose to the editor a close or adverse relationship with the author? What is the advisory role of the reviewer? What review criteria are necessary to ensure a timely, thorough, and fair review? Since articles are intended to be contributions to the field of knowledge, what constitutes a scholarly review of past work and faithful referencing? Should formal dissertations, laboratory reports, and edited conference proceedings be referenced if available from national information services or libraries? Should articles published in edited proceedings or formally archived reports be published as journal articles? Should research, numerical or experimental, be reported in incremental steps or as a completed investigation with significant results? Is the conduct of a large experiment sufficient for publication? Could prompt publication of letters to the editor about articles be used as a peer-feedback-quality control?

10:05

3aED4. The experience of an Associate Editor with the review process of papers in the *Journal*. Courtney B. Burroughs (Appl. Res. Lab., The Penn State Univ., P.O. Box 30, State College, PA 16804)

One of the most important parts of the Society is the *Journal*, in which information on new research findings are disseminated. Before a paper is published in the *Journal*, it must be approved by an Associate Editor who typically seeks the advice of two experts on the subject of the paper. Based on 6 years experience as an Associate Editor for the *Journal*, an overview of several aspects involved in the review process will be given. What constitutes a good review and what reviewers are looking for in an acceptable paper will be presented. Although the decision of acceptance rests with the Associate Editor, it is rare that this Associate Editor does not adhere to the recommendations made by reviewers. However, this Associate Editor does read all of the papers prior to deciding either to accept a paper after revisions are made in response to reviewer comments, reject the paper, or defer a final decision until after the paper has been revised and submitted for further review. An overview of this Associate Editor's experience with making the acceptance/rejection decision and what is felt to be the ingredients required for an acceptable *Journal* paper will be given.

10:35

3aED5. Anatomy of a readable publication. John C. Burgess (Dept. of Mech. Eng., Univ. of Hawaii, 2540 Dole St., Honolulu, HI 96822)

Richard Borden, some time ago, came up with a four-point formula that applies to any type of communication. The colloquial expressing of this formula consists of the phrases (1) Ho hum! (2) Why bring that up? (3) For instance! (4) So what? The present paper rephrases this formula so that it can be used by authors of research articles for JASA. The phrase “Ho hum,” for example, translates to the requirement that the opening paragraph of an article should attract the reader's interest and motivate the reader to continue reading. Examples are taken from various JASA papers that indicate that those papers which are manifestly readable are in actuality (even if unconsciously) adhering to Borden's formula. It is suggested that authors who want their papers to be read and understood give serious consideration to making use of Borden's formula.

11:05

3aED6. The importance of critical book reviews. Philip L. Marston (Dept. of Phys., Washington State Univ., Pullman, WA 99164-2814, marston@wsu.edu)

According to C. S. Lewis, “Literature exists to teach what is useful, to honor what deserves honor, to appreciate what is delightful” [C. S. Lewis, *The Discarded Image* (Cambridge U.P., Cambridge, 1967), p. 214]. While Lewis's context was remote from scientific and technical literature, these considerations are relevant to publication of timely book reviews in the *Journal of the Acoustical Society of America*. In addition to promoting a general awareness of books and the subject matter considered therein, book

reviews can recognize and promote quality and the archiving of appropriate materials in libraries. The recognition of the real or apparent faults of a book is a valuable service along with the recognition of a book's positive attributes; however, it is appropriate to remember that aspects of the publication process may not be controlled by a book's author or authors.

11:35

3aED7. Publishing speech papers in JASA. Anders Lofqvist (Dept. of Logoped. and Phoniatics, University Hospital, SE-221 85 Lund, Sweden, Anders.Lofqvist@logopedi.lu.se)

There are three sections for speech communication papers in JASA: production, perception, and processing and communication systems. In contrast to some other branches of acoustics, authors of speech-related papers have a wide variety of journals to choose from, covering several different disciplines. Among these are journals in psychology, cognitive science, linguistics, phonetics, speech and hearing science, otorhinolaryngology, neurophysiology, human movement science, and signal processing. Speech papers also cover production and perception in different human languages. In addition, papers in speech deal with both normal and clinical aspects, e.g., voice and speech disorders, and the effects of hearing impairment on speech. Papers in speech production are also concerned with analyzing articulatory movements as well as the mathematical modeling of phonation and articulation. The presentation will focus on why JASA should be a top choice for authors of papers in speech communication. Criteria for the selection of speech papers will be discussed as well as reviewing policy.

THURSDAY MORNING, 1 JUNE 2000

TOWER 1205, 8:30 TO 11:30 A.M.

Session 3aMU

Musical Acoustics: Visualization Methods for Musical Instrument Acoustics

Uwe J. Hansen, Chair

Department of Physics, Indiana State University, Terre Haute, Indiana 47809

Chair's Introduction—8:30

Invited Papers

8:35

3aMU1. Holographic analysis of musical instruments. Thomas D. Rossing (Phys. Dept., Northern Illinois Univ., DeKalb, IL 60115)

Musical instruments, when excited by bowing, plucking, blowing, or striking, vibrate in rather complicated ways. These complicated vibrations can be described in terms of normal modes of vibration, and understanding the normal modes or eigenmodes is important to understanding sound production in the instrument. Holographic interferometry offers by far the best spatial resolution of operating deflection shapes (and hence of normal modes), since it looks at an almost infinite number of points. Holographic interferograms may be recorded on photographic film or electronically by means of TV holography. We discuss various methods for holographic interferometry and its applications to several different types of musical instruments.

9:00

3aMU2. Extending modal analysis techniques: Representing the sound field. Uwe J. Hansen (Dept. of Phys., Indiana State Univ., Terre Haute, IN 47809) and Ingolf Bork (Physikalisch-Technische Bundesanstalt, Braunschweig, Germany)

Modal analysis enables a visual representation of structural normal-mode vibrations. This is accomplished by obtaining a series of transfer functions over a predetermined point grid on the structure, and fitting them to a model of coupled harmonic oscillators. These data can subsequently be visualized in slow-motion animation on the computer screen, representing the motion of the structure at the normal-mode frequencies. The key element in this procedure is the phase coherence between the excitation and the response signal. The technique can be extended to represent a sound field by retaining the phase coherence between a drive signal and a response signal detected by a microphone placed successively at points on a spatial grid in the sound field. The moving displacement on the computer screen represents pressure variations in their phase relation to the source oscillations. The technique will be illustrated with measurements in the air column of a flute, and the sound field above and below the sound board of a grand piano, as well as in a plane in front of the instrument.

9:25

3aMU3. Visualization of violin sound-radiation results from multiplanar near-field acoustic holography. Lily M. Wang and Courtney B. Burroughs (Grad. Prog. in Acoust., Penn State Univ., P.O. Box 30, State College, PA 16804)

The application of multiplanar near-field acoustic holography (NAH) to violins has produced an enormous amount of radiated sound-pressure and intensity data in three dimensions surrounding the instruments. The data have been analyzed to identify areas of significant sound energy radiation on the violin surface and the subsequent paths of that energy. The source regions have then been correlated to modal analysis results, as well as compared between structurally different violins. To complete such analyses has

required powerful visualization tools. This presentation reviews the visualization software called application visualization system (AVS) and the techniques used for the investigation, including animating results on several sequential planes and viewing data as three-dimensional bricks that may be cut to view the interior behavior. [Work supported by NSF Graduate Research Fellowship, Lucent Technologies/Bell Laboratories GRPW Grant, and AAUW Selected Professions Dissertation Fellowship.]

9:50

3aMU4. The use of color in representing complex number data. Gabriel Weinreich (Phys. Dept., Randall Lab., Univ. of Michigan, Ann Arbor, MI 48109-1120)

With the general availability of color printers, the application of color for representing physical quantities has taken on a new interest. Since hue is by its nature a cyclic variable, it is ideally suited for representing the phase of a complex number; at the same time, color saturation can be used to specify amplitude. In this way, a complex number is conveniently represented by a single color; and a complex number that depends on, say, frequency, such as the input admittance of a wind instrument or the radiativity of a violin, can be mapped into a one-dimensional colored band. The great advantage of such a method is that these one-dimensional bands can be "stacked" in a perpendicular direction to indicate how the whole function behaves when another independent variable is introduced; for example, one might exhibit a radiativity as a function of both frequency and direction, or an input impedance as a function of both frequency and the number of open tone holes. A detailed application of the method to the "Swiss Cheese Violin" of Carleen Hutchins will be presented. [Work supported by NSF.]

10:15

3aMU5. Acoustic vortex appearing at the pre-steady state and disappearing at the steady state in organ pipes. Shigeru Yoshikawa (Dept. of Acoust. Design, Kyushu Inst. of Design, 4-9-1 Shiobara, Minami-ku, Fukuoka, 815-8540 Japan, shig@kyushu-id.ac.jp)

In visualizing the motion of the smoked jet during the attack transient, an acoustic (acoustically induced) vortex was observed alternately above and beneath the pipe edge at the pre-steady state just prior to the steady state. Interestingly enough, this acoustic vortex disappeared at the steady state. This finding is contrary to Fabre's, in which the acoustic vortex appears at the steady state [Acust. Acta Acust. **82**, 863-877 (1996)]. The rotation of the acoustic vortex is the reverse of that of the hydrodynamic vortex produced at the initial phase of the attack transient. According to Howe's theory [J. Sound Vib. **70**, 407-411 (1980)], our acoustic vortex seems to absorb the final excess in acoustic energy generation occurring at the pre-steady state and to lead the finally saturated amplification of the jet stability wave. The acoustic vortex may then be convected by the jet flow into regions where the vorticity can no longer continue to interact with the acoustic field. On the other hand, Fabre's acoustic vortex was interpreted as some sound source corresponding to higher harmonics. Slow-motion pictures captured with a high-speed digital video camera will be demonstrated.

10:40

3aMU6. Alternative mouthpiece design for viewing the lip reed in motion. R. Dean Ayers and Michael S. Lodin (Dept. of Phys. and Astron., California State Univ., Long Beach, 1250 Bellflower Blvd., Long Beach, CA 90840, rdayers@csulb.edu)

Mouthpieces for stroboscopic viewing of the lip reed have been designed for the full range of brass instruments, from trumpet to tuba. In order to provide an unobstructed frontal view of the lips, the axis of the backbore has been rotated 90° from the axis of the rim, so that it emerges from the side of the cup. This is similar to the original design of Daniel W. Martin [J. Acoust. Soc. Am. **13**, 305-308 (1942)]. Flat end windows are desirable for avoiding optical distortion, but they limit the range of viewing angles. Placement of the lips on the rim and the angle of the air channel between the lips can vary noticeably from performer to performer, so a flat window that works well for one embouchure may be quite inadequate for another. The large cup volumes of the low brass mouthpieces allow for greater flexibility in design without getting into severe distortion. A cylindrical end window with its axis parallel to that of the rotated backbore provides a wide range of viewing angles. A preliminary survey of embouchures used by beginning and advanced trombone players will be presented. [Work supported in part by the Scholarly and Creative Activities Committee at CSULB.]

11:05

3aMU7. A real-time/non-real-time spectrum analyzer for musical sounds. James W. Beauchamp and Timothy J. Madden (School of Music and Dept. of Elec. and Computer Eng., Univ. of Illinois at Urbana-Champaign, Urbana, IL 61801)

A color spectrum analyzer specialized for visualization of musical sounds has been developed for the Macintosh Power PC computer. Input samples (monaural or stereo) can be obtained from microphone or CD-ROM input in real time or from a sound file in nonreal time. Spectral analysis can be tuned specifically to the fundamental frequency of the input signal using sample rate conversion (tuned case), or it can be based on a specified window size at the input sample rate (nontuned case). In the tuned case each spectral component corresponds to a harmonic of the input. Several different displays can occur simultaneously: an oscilloscope waveform display; a 1D spectrum, where an amplitude-vs-frequency graph varies as time progresses; a spectral waterfall display; a 2D spectrum showing frequency vs time with amplitude registered as degree of display brightness; a 3D display showing amplitude vs time vs frequency at an arbitrary orientation. In addition to time-varying graphs which are available for both real and nonreal time, static plots are available for preanalyzed file input. For stereo signals it is possible to plot output/input transfer functions. Spectral time averaging is also available for cases where the instantaneous spectrum is unstable.

Session 3aPAa

Physical Acoustics: Acoustics of Multiphase Flow I

Andrea Prosperetti, Cochair

Mechanical Engineering Department, Johns Hopkins University, 119 Latrobe Hall, 3400 North Charles Street, Baltimore, Maryland 21218

Lev A. Ostrovsky, Cochair

University of Colorado, NOAA/ETL, ET1, 325 Broadway, Boulder, Colorado 80303

Chair's Introduction—8:00

Invited Papers

8:05

3aPAa1. Acoustics of two-phase fluids. Robert I. Nigmatulin (Ufa Branch of Russian Acad. of Sci., 6 K. Marx Str., Ufa 450025, Russia)

Wave propagation in two-phase fluids is accompanied by a great variety of inter- and intraphase hydromechanical and thermo-physical processes. These processes are initiated by a wave and, in turn, influence wave propagation strongly. Specific and anomalous features of the waves are the following: (1) Relaxation nonequilibrium processes need some relaxation time and space to transfer one equilibrium state in front of the wave to another equilibrium state behind. That is why the thickness of the wave or its relaxation zone may be large. (2) Attenuation of the wave is governed by nonlinear dispersion of disturbances and by dissipation because of the viscous friction between the carrier and dispersed phases and thermal dissipation, which is strongly dependent on phase transition possibility. (3) Bubbly liquid may show not only the attenuation, but even amplification of the shock wave. (4) For gas-drop mixtures with two-component carrier gas (noncondensable gas and vapor) it is possible to see an anomalous nonmonotonous effect of the mass drop content on the attenuation of the harmonic forced oscillation wave. (5) The breakdown of drops and bubbles has a strong influence on the wave propagation. It intensifies the interface force, heat, and mass transfer interactions. Sometimes it may appear as a "vapor explosion."

8:35

3aPAa2. The Nakoryakov–Pokusaev contribution to the acoustics of multiphase flow. Isaac Schreiber (Dept. of Chemical Eng., Ben-Gurion Univ. of the Negev, Beer-Sheva 84105, Israel)

Under the guidance of academician Nakoryakov and Professor Pokusaev, a cycle of theoretical and experimental studies of wave propagation in the gas–vapor bubbly liquid has been accomplished. It established the basis for the gas dynamics of gas–liquid multiphase flow. Here the main achievements will be listed. (1) The bridge was built between the modern theory of nonlinear waves and the progress in heat transfer. (2) The comprehensive picture of the dynamics and structure of the wave propagation in gas–vapor bubbly liquid was found. The main bright results are the following: discovery of oscillating and monotonic shock waves and derivation of criteria for their existence; models for sound propagation in bubbly liquid and in boiling water (Korteweg–Burgers and Klein–Gordon approximations); acoustics of the slug-plug regime; two-wave structure of shock waves with shape fronts. These results strongly influenced the modern theory of the acoustics of foams and solids saturated by bubbly liquid. This lecture is dedicated to the jubilees of Nakoryakov (65th birthday) and Pokusaev (65th birthday).

9:05

3aPAa3. Nonlinear wave interactions in water with bubbles: Solved and unsolved problems. Lev A. Ostrovsky (Univ. of Colorado, CIRES/NOAA ETL, 325 Broadway, R/E/ET1, Boulder, CO 80303)

This lecture is an outline of some nontrivial aspects of nonlinear acoustics of bubbly liquids. It is common knowledge now that a very small volume content of bubbles can increase the nonlinearity parameter of the gas–liquid mixture by several orders. Basic problems of nonlinear wave propagation in such media with shock wave formation, harmonic generation, etc., have been thoroughly discussed during the last decades, especially for a less realistic case of equal-size bubbles. Not as well understood are waves in bubbles widely distributed in sizes, especially nonlinear scattering problems when the standard "continual" approximation is inapplicable. Also, collective behavior of bubbles interacting with each other (not just via the macroscopic acoustic field) can radically change the nonlinear properties on a medium. Here, some relevant theoretical and experimental results are discussed, including nonlinear scattering of sound on sea bubbles; the use of resonance bubble layers to enhance nonlinear transformation; and some effects of bubble interaction due to period-averaged forces such as radiation pressure. The relationship between coherent and incoherent parts of acoustic field and some ideas of an "acoustic masker" are also discussed.

3aPAa4. The coupling of airborne sound into air-filled poro-elastic soils. James M. Sabatier (Natl. Ctr. for Physical Acoust., Univ. of Mississippi, University, MS 38677, sabatier@olemiss.edu)

The coupling of airborne sound into near-surface soils will be reviewed. Particular emphasis will be given to research conducted at the National Center for Physical Acoustics over the past 10 years, including the works of Arnott, Frederickson, and Hickey. For the purpose of acoustically describing pore properties, the ground is modeled as an air-filled, porous, rigid-framed material. Measurements using both transmission and reflection of sound have been used in conjunction with rigid framed acoustic propagation models to determine air flow resistivity, porosity, and tortuosity. In other research, the ground is modeled as a layered poro-elastic medium following the work of Biot. Measurements of the motion of the elastic frame are made using geophones, buried microphones, and laser Doppler vibrometers. Responses of these sensors to airborne or acoustically induced seismic waves will be described. A recent application of this understanding of the ground is the detection of buried land mines. Some brief comments will be made with regard to this application.

10:00–10:20 Break

10:20

3aPAa5. Forward and inverse modeling in porous media. Joseph F. Lingeitch, Michael D. Collins (Naval Res. Lab., Washington, DC 20375), Andrew J. Fredricks, and William L. Siegmann (Rensselaer Polytechnic Inst., Troy, NY 12180)

Techniques for solving forward and inverse wave propagation problems involving poroelastic layers will be discussed. Parabolic equation techniques are efficient for solving problems in laterally varying media. A parabolic equation for poroelastic media has been developed and applied to problems in ocean acoustics [J. Acoust. Soc. Am. **98**, 1645–1656 (1995)]. This approach has been generalized to the anisotropic case. The smallness of shear speeds in many ocean sediments has motivated the study of poroacoustic media [J. Acoust. Soc. Am. **104**, 783–790 (1998)], which is a limiting case of Biot theory in which the rigidity vanishes. Although parabolic equation techniques have not been fully generalized from acoustics to poroelasticity, studying the intermediate case of poroacoustics has helped to bridge the gap. The parabolic equation techniques have been used as tools for solving inverse problems. This approach is presently being applied to field data. One of the issues that arises in solving the inverse problem is the mapping between the coefficients of the wave equation and the wave speeds. The inverse of this mapping can be used to define problems in terms of natural parameters rather than moduli. [Work supported by ONR.]

10:45

3aPAa6. Methods for measurement of the attenuation coefficient and phase velocity of Optison and their relevance to *in vivo* efficacy. Michael S. Hughes, Alexander L. Klibanov, John H. Wible, and Gary H. Brandenburger (Mallinckrodt, Inc., 675 McDonnell Blvd., Hazelwood, MO 63042, mxhughe@mkg.com)

We report methods and results of measurements of contrast media attenuation coefficient and phase velocity for concentrations ranging from 1.7×10^5 to 21×10^6 particles/ml. These concentrations cover the range likely to occur during *in vivo* application of ultrasound contrast media such as Optison. For the concentrations cited in this study, the attenuation coefficient exhibits a single peak between 1 to 2 MHz, while the phase velocity rises sharply and then levels off, typically near the same frequency at which the attenuation peak occurs. Previous studies from our laboratory have shown that these acoustic properties of ultrasound contrast media can change significantly with changing experimental conditions. For instance, the frequency at which the peak attenuation of Albunex occurs can shift by as much as 1 to 1.5 MHz as the measurement temperature is increased from 22° to 37 °C. Similarly, the attenuation peak width (full width at half maximum) can narrow by more than a factor of 2 (the exact amount depends on concentration). The peak attenuation values can shift by as much as 50% with increasing acoustic pressure. The primary aim of this study is to measure acoustic properties which are as close as possible to those occurring in clinical practice.

Contributed Papers

11:10

3aPAa7. Design and characterization of a system for acoustophoresis. Todd L. Brooks and Robert E. Apfel (Dept. of Mech. Eng., Yale Univ., New Haven, CT 06520-8286, todd.brooks@yale.edu)

The manipulation of particles using acoustic forces is a basis for a potentially valuable continuous separation method. The method is unique in that particles can be distinguished based on their relative compressibilities, as well as on mass density and size. An acoustic standing wave is oriented perpendicular to the fluid flow direction which enables an incoming particle mixture to be separated into different streams at the fluid output. Effects which cause mixing such as heat convection and acoustic streaming must be minimized through careful design. An imaging system has been implemented to record the trajectories of individual particles as they flow through the separation cell. Trajectories of various types of particles ranging from 5 to 55 microns are compared to theoretical esti-

mates of separation effectiveness, with a particular emphasis on the role of secondary inter-particle acoustic forces. [Work supported by NASA through Grant No. NAG8-1351.]

11:25

3aPAa8. Just a box of rain: Making rain with sound. Gretchen Snoeyenbos, Giovanni De Santi, and R. Glynn Holt (Dept. of Aerosp. and Mech. Eng., Boston Univ., 110 Cummington St., Boston, MA 02215)

Clouds don't always produce rain. One of the reasons is that microdroplet coalescence does not occur fast enough to produce drops large enough to fall. A brief review of the literature on particle and aerosol agglomeration will be presented. The results of an experiment in which ultrasonic standing waves are employed to enhance microdroplet coalescence in a laboratory cloud will be reported. A force analysis will be

presented to allow prediction of the efficiency of the coalescence process for arbitrary acoustic field frequency and droplet radius. If possible, a comparison will be made between experimental results and model predictions.

11:40

3aPAa9. New models and measurements concerning sound attenuation in concentrated airborne suspensions. Keith Attenborough, Qiang Wang (Univ. of Hull, Hull HU6 7RX, UK), and Steve Woodhead (Univ. of Greenwich, Chatham Maritime ME4 4AW, UK)

In the inertial regime of frequency-radius space, irregularity and aggregation of particles can result in values of acoustic attenuation that are significantly different from those predicted by assuming separated smooth spherical particles. Data obtained previously from suspensions of alumina

particles and olivine sand in air at audio frequencies, together with new data obtained at low ultrasonic frequencies in suspensions of glass beads and silica flour, are compared with predictions. It is shown that neither a coupled-phase theory modified to allow for nonspherical shapes nor effective radius theories are able to account for these data. Qian [Phys. Rev. E **53**(3), 2304–2306 (1996)] has suggested that suspensions may be treated as fractal media and used the acoustic Reynolds number as the fractal dimension in modifying scattering theory. A new fractal modification of multiple scattering theory for acoustic attenuation is derived. The theory uses $\omega\tau_v$ (ω is the angular acoustic frequency, τ_v is the dynamic relaxation time of the particles) as a fractal scale. Fitted values of fractal dimension obtained at a single frequency are found to enable fits with data at other frequencies. Moreover, the fractal approach is found to enable discrimination between the effects of particle irregularity and aggregation. [Work supported by EPSRC (UK).]

THURSDAY MORNING, 1 JUNE 2000

PEACH ROOM, 8:15 A.M. TO 12:00 NOON

Session 3aPAb

Physical Acoustics: Scattering and Radiation

Christine Valle, Chair

Mechanical Engineering Department, University of Maine, Orono, Maine 04473

Contributed Papers

8:15

3aPAb1. Amplitude variation of resonant frequencies and Q-factor in impulse acoustic resonant spectroscopy. Alexander M. Sutin (Stevens Inst. of Technol., 711 Hudson St., Hoboken, NJ 07030, asutin@stevens-tech.edu), Robert A. Guyer (Univ. of Massachusetts, Amherst, MA 01003), and Paul A. Johnson (Los Alamos Natl. Lab., Los Alamos, NM 87545)

One of the oldest and simplest nondestructive tests of the quality of an object is to tap it and listen to the radiated sound. The translation of this test into a modern instrumentation system leads to impulse acoustic resonance spectroscopy, IARS. The processing of the acoustic time train after impulse excitation (tap) allows one to follow Q and resonance frequencies as a function of time. Examination of the time train can reveal important properties of this elastic state and of the state of the object. IARS will be illustrated with data on a suite of automobile brake drums made of powdered aluminum. The time evolution of Q and a resonance frequency will be described. Typically Q is of order one-half of its late time (equilibrium) value at short times whereas a resonance frequency shifts by of order 0.2% from early (0.2 s) to late (5 s) time. These qualitative properties have two possible explanations: (1) amplitude-dependent internal friction or (2) the “slow dynamics” that is characteristic of elastic systems inhabited by hysteretic elastic elements. [Work supported by Stevens and by the Department of Energy: Office of Basic Energy Sciences.]

8:30

3aPAb2. On a novel application of the Helmholtz integral in the development of a noiseless sonar. Anthony J. Romano, Joseph A. Bucaro, Brian H. Houston, and Earl G. Williams (Naval Res. Lab., Washington, DC 20375-5350)

A novel application of the Helmholtz integral is presented in the development of a noiseless sonar. Given total pressure and normal velocity information over the surface of a structure which has been excited both by an incident field as well as interior noise sources, when the field point

resulting from the Helmholtz integral is evaluated inside the surface of the structure, only the incident field remains. Therefore, it is possible to integrate out the effects of both self-noise (due to interior noise sources) and scattering (due to incident excitation) throughout a “virtual” volumetric sensor array projected within the structure. This approach will be discussed and demonstrated using numerical results based on finite element simulations. [Work supported by ONR.]

8:45

3aPAb3. Inverse scattering of three-dimensional obstacles. Dilip N. Ghosh-Roy (SFA, Inc., 1401 McCormick Dr., Landover, MD 20785), Luise S. Couchman, and Jeremy A. Warner (Naval Res. Lab., Washington, DC 20375-5350)

The results of three-dimensional reconstructions of obstacles in an infinite, homogeneous ambience from the far-field scattering of plane acoustic waves in a resonance frequency region are presented. The reconstructions include penetrable and impenetrable objects. Also discussed are reconstructions of orientations (that is, Euler angles), surfaces, and interiors of penetrable axisymmetric scatter. In addition, inverse obstacle scattering using a highly restricted angular aperture in monostatic and quasi-monostatic data collection configurations is also discussed.

9:00

3aPAb4. Backscattering enhancements of ultrasound by tilted solid plastic cylinders in water due to the caustic merging transition: High angular resolution scans. Florian J. Blonigen and Philip L. Marston (Dept. of Phys., Washington State Univ., Pullman, WA 99164-2814, blonigen@physalpha.physics.wsu.edu)

Bulk shear and longitudinal waves can make important contributions to the scattering of ultrasound by tilted finite plastic and rubber cylinders in water when the phase velocity of the wave is less than the speed of sound of the surrounding fluid. At a certain critical tilt angle, a back-

scattering enhancement is observed as a result of merging rainbow caustics associated with bulk transmitted rays internally reflected off the cylinder truncation [F. J. Blonigen and P. L. Marston, *J. Acoust. Soc. Am.* **107** (2000)]. The backscattering away from this angle is otherwise weak. The critical angle is calculated using the Bravais effective refractive index for the projections of rays on the base plane of the cylinder. New ultrasonic observations are reported for a polystyrene cylinder over a wide range of tilt angles at high angular resolution. The backscattering amplitude angle scan takes on the shape of a Pearcey function with the peak amplitude offset from the critical angle. Backscattering amplitudes at the critical angle are also compared with a ray theory prediction as a function of frequency for both the polystyrene and a silicone rubber cylinder. [Work supported by ONR.]

9:15

3aPAb5. Backscattering enhancements associated with mode conversion in thin tilted circular plates in water: Observations and Gaussian beam model. Brian T. Hefner and Philip L. Marston (Phys. Dept., Washington State Univ., Pullman, WA 99164, bhefner@mail.wsu.edu)

When a thin tilted circular plate is illuminated with high-frequency sound, a significant enhancement occurs when the angle of incidence corresponds to the extensional wave coupling angle. It has been found that there are two mechanisms responsible for this enhancement. The largest response is due to the extensional leaky wave which travels along the diameter and reflects from the plate edge. This can be modeled in the thin-plate limit using a Gaussian beam model which was first developed for the reflection of leaky Rayleigh waves on tilted truncated cylinders [K. Gipson and P. L. Marston, *J. Acoust. Soc. Am.* **107**, 112–117 (2000)]. The second response is due to extensional waves traveling along off-diameter paths. When these waves strike the plate edge, they mode-convert into shear waves. These shear waves cross the diameter to strike the opposite edge and mode-convert into extensional waves which then reradiate back towards the source. Since the shear waves do not leak energy into the water, this enhancement has a relatively slow decay as the waves undergo multiple reflections around the plate edge. The Gaussian beam model has been extended to quantitatively model this enhancement in the thin-plate limit. [Work supported by ONR.]

9:30

3aPAb6. Raman-Ray whispering gallery experiment and enhanced backscattering from merged caustics: Optical analogies of acoustical wavefields. Philip L. Marston, Charles C. Barnes, Yibing B. Zhang, and David B. Thiessen (Dept. of Phys., Washington State Univ., Pullman, WA 99164-2814)

Two optical experiments for viewing wavefields associated with acoustical processes are demonstrated. The first is a modern version of a little-known whispering gallery experiment by Raman and his associate Ray [C. V. Raman and G. A. Sutherland, *Proc. R. Soc. London, Ser. A* **100**, 424–428 (1922); B. Ray, *Bull. Calcutta Math. Soc.* **12**, 225–230 (1922)], which simulate wavefields for a source on a curved reflecting surface. Our simulation uses a long, narrow front-surface mirror bent uniformly by attached weights. The light source is the diffraction of laser light by a knife edge placed adjacent to the mirror. The caustic associated with once-reflected rays (a cardioid curve) is easily distinguished as are other caustics associated with rays having only a few reflections. Amplitude contributions near each caustic are approximated by Airy functions. The second experiment is analogous to an acoustical backscattering enhancement by bluntly truncated tilted penetrable circular cylinders associ-

ated with the merging of end-reflected rainbow caustics [F. J. Blonigen and P. L. Marston, *J. Acoust. Soc. Am.* (in press)]. Light backscattered by a tilted blunt glass fiber displays the merged caustics for the appropriate tilt. [Work supported by ONR and NSF-CRDC.]

9:45

3aPAb7. Numerical results for acoustic scattering from a coated wedge. R. Hughes, J. Niemiec, S. Solomon (Naval Surface Warfare Ctr., Carderock Div., Bethesda, MD 20817-5700), and Herbert Überall (Catholic Univ. of America, Washington, DC 20064)

Earlier work on the acoustic wedge problem mainly dealt with an impenetrable wedge, including the work of Sommerfeld (1896), Bromwich (1915), Carslaw (1920), and Oberhettinger (1954, 1958). The present investigation assumes a line source parallel to the edge of the wedge, and extends Oberhettinger's approach for the scattering of a sound wave from the edge of a perfectly reflecting wedge, to the case that the field does not satisfy a Dirichlet or Neumann boundary condition on the wedge surface, but an impedance boundary condition. Such a condition is capable of describing the effect of a lossy coating on the wedge surface as shown for the electromagnetic case (Überall, 1964, 1966), and the corresponding impedance for the acoustic case has been obtained here for one, two, and multiple layers of coating. Numerical results for the wedge-scattered sound intensity will be presented and discussed.

10:00–10:15 Break

10:15

3aPAb8. Radiation impedance matrix of a rectangular aperture, with arbitrary vibration, in a rigid baffle. Allan D. Pierce and Robin O. Cleveland (Dept. of Aerosp. and Mech. Eng., Boston Univ., Boston, MA 02215, adp@bu.edu)

The present paper shows that multidimensional integrals associated with radiation from a rectangular surface in a plane rigid baffle can be reduced to one-dimensional integrals. If the vibrational excitation is expanded into a summation over a set of basis functions, the surface pressure and velocity may be related by an infinite matrix. Each element of the matrix is expressed as a quadruple integral. The numerical evaluation of such integrals, although possible, is prohibitively time consuming for many practical applications. However, for some choices of the basis functions, each quadruple integral may be reduced analytically to single integrals. Explicit expressions are given for square surfaces and with the basis functions taken as products (or sums of products) of trigonometric functions. Numerical results agree well with previous results published in the literature for the pressure distribution over a surface when the vibration is uniform.

10:30

3aPAb9. Dispersion of axially symmetric waves in fluid-filled cylindrical shells. X. L. Bao, H. Überall (Phys. Dept., Catholic Univ. of America, Washington, DC 20064), P. K. Raju, A. C. Ahji (Auburn Univ., Auburn, AL 36849-5341), I. K. Bjørnø, and L. Bjørnø (Tech. Univ. of Denmark, DK-2800 Lyngby, Denmark)

Acoustic waves normally incident on an elastic cylindrical shell can cause the excitation of circumferential elastic waves on the shell. These shells may be empty and fluid immersed, or fluid filled in an ambient

medium of air, or doubly fluid loaded inside and out. Circumferential waves on such shells have been investigated for the case of aluminum shells, and their phase-velocity dispersion curves have been obtained for double fluid loading [Bao, Raju, and Überall, *J. Acoust. Soc. Am.* **105**, 2704 (1999)]. Similar results were obtained for empty or fluid-filled brass shells [Kumar, *Acustica* **27**, 317 (1972)]. We have extended the work of Kumar to the case of fluid-filled aluminum shells and steel shells imbedded in air. These cases demonstrate the existence of circumferential waves traveling in the filler fluid, exhibiting a certain simplicity of the dispersion curves of these waves. This is in striking contrast to the results for double (outside and inside) loading by two fluids of comparable density, where circumferential waves in both external and internal fluids were found, their interaction causing segmentation and repulsion phenomena of their dispersion curves. The condition of standing circumferential waves determines the eigenfrequency spectrum of the shell.

10:45

3aPAb10. A general acoustic background for evacuated and water-filled submerged shells. Michael F. Werby (NRL Code 7181, Stennis Space Center, MS 39571, werby@nrlssc.navy.mil) and Herbert Überall (Catholic Univ. of America, Washington, DC)

Historically, the acoustic background for elastic solids dates back to the early 1980s to the work of Flax, Überall, and Dragonette. Werby extended this concept in 1991 to evacuated elastic shells by employing the concept of entrained mass. Here we extend the concept to one valid for a water-filled shell. The new background is equally valid even for the limit of an evacuated shell. We derive the equations used in the method and illustrate its use for a variety of problems.

11:00

3aPAb11. A comparative study of the scattering cross sections from incident plane wave acoustic waves on submerged evacuated and water-filled elastic shells. Herbert Überall (Dept. of Phys., Catholic Univ. of America, Washington, DC, huberall@aol.com) and Michael F. Werby (NRL, Stennis Space Center, MS)

Employing the acoustic background for the case of elastic solids and elastic shells has facilitated the study of elastic resonances excited by incident acoustic signals. Once an adequate acoustic background became available it became possible to properly isolate and characterize such resonances. The recent development of an acoustic background for water-filled elastic shells allows for the characterization of water-filled shells. We present a comparative study for the elastic shell case for the evacuated and water-filled cases and thereby seek to elucidate new resonance mechanisms for the fluid-filled case.

11:15

3aPAb12. Multiple scattering attenuation and anisotropy of ultrasonic surface waves. John A. Scales and Kasper Van Wijk (Ctr. for Wave Phenomena, Colorado School of Mines, Golden, CO 80401, kasper@dix.mines.edu)

Multiple scattering of waves induces bulk effects such as attenuation and anisotropy that are important in seismology, optics, medical imaging, and other fields involving propagation in disordered media. Measurements are reported of ultrasonic surface wave propagation in a strong-scattering, quasiperiodic medium consisting of a grooved surface of aluminum. Using

noncontacting optical methods we have tracked the evolution of surface wave pulses within the scattering medium. Waves propagating parallel to the grooves propagate nearly attenuation and dispersion free, whereas waves propagating normal to the grooves are dispersed and exponentially attenuated with distance as energy is transferred from the direct pulse into the multiple-scattering coda. We measure this attenuation length and show that there is, in addition, a scattering induced anisotropy in the phase velocity. [This work was partially supported by the U.S. Army Research Office under Grant No. DAAG55-98-1-0070.]

11:30

3aPAb13. Damage detection in thin composites using guided waves. Christine Valle (212 Boardman Hall, Univ. of Maine, Orono, ME 04473, valle@umeme.maine.edu)

Lightweight, high-temperature composite materials have the potential to provide significant economic and performance benefits to aerospace structures. However, one problem inherent in using composites within complex structures in harsh environments is the variety of damage mechanisms experienced by the composite itself. Therefore, conventional NDE techniques that are sensitive to only one kind of damage can give an overly optimistic assessment of the health of the composite. A global measure that takes into consideration all of the different damage mechanisms better addresses the nondestructive evaluation needs of composites used in critical applications. This study investigates the relationship between the composite's dispersion relationship—generated with a 2D fast Fourier transform representation of transient ultrasonic signals, created with a finite-element (FE) model—and the energy content of the structure, as a function of damage amount. The total amount of energy contained in the structure is measured, and when normalized with respect to a virgin state, illustrates the loss due to a given amount of damage. This method provides a quantitative assessment of the amount of damage in the material, and establishes the numerical and analytical foundations for the application of FE models to damage monitoring in composites.

11:45

3aPAb14. New techniques for the characterization of materials *in situ* using ultrasound. Konstantinos Christidis and G. P. P. Gunarathne (Dept. of Electron. and Elec. Eng., The Robert Gordon Univ., Aberdeen AB11FR, UK)

Characterization of materials *in situ* using ultrasound is an attractive concept but has two main problems due to: (1) arbitrary geometrical shape of the target, and (2) surface roughness of the material. Some work has been previously reported in measuring the surface roughness using ultrasound [J. Stor-Pellinen and M. Luukkala, *Sensors and Actuators* **A49**, 37–40 (1995)] and also material characterization using ultrasonic feature extraction [G. P. P. Gunarathne and R. W. Keach, *Ultrasonics* **34**, 411–419 (1996)]. However, there is no integrated approach developed to characterize materials *in situ*, taking into account the effects due to shape and roughness of the target. In this work a comprehensive study of the effects of target shape and roughness on ultrasonic time- and frequency-domain measurements were carried out. A mathematical model which compensates for these variables was then developed, so that the compensated signals could be used to evaluate the material properties. Using these mathematical models and practical measurements, a formula which compensates for the combined effect of curvature and target roughness was developed. This model was then used to characterize materials *in situ*, using stored data and artificial neural networks. Close agreement between theoretical and practical measurements has been demonstrated.

Session 3aPP

Psychological and Physiological Acoustics: Localization and Precedence (Poster Session)

Ruth Y. Litovsky, Chair

Hearing Research Center, Department of Biomedical Engineering, Boston University, 44 Cummington Street,
Boston, Massachusetts 02215

Contributed Papers

All posters will be on display from 8:30 a.m. to 12:00 noon. To allow contributors an opportunity to see other posters, contributors of odd-numbered papers will be at their posters from 8:30 a.m. to 10:15 a.m. and contributors of even-numbered papers will be at their posters from 10:15 a.m. to 12:00 noon.

3aPP1. Spatial unmasking of nearby pure-tone sources in a simulated anechoic environment. Norbert Kopčo (Dept. of Cognit. and Neural Systems, Boston Univ., Boston, MA 02215, kopco@cns.bu.edu) and Barbara G. Shinn-Cunningham (Boston Univ., Boston, MA 02215)

Spatial unmasking of nearby pure-tone sources masked by broadband noise was studied in a simulated anechoic environment. The predicted amount of binaural unmasking was calculated using interaural differences from a spherical-head model [B. G. Shinn-Cunningham, S. G. Santarelli, and N. Kopčo, *J. Acoust. Soc. Am.* (in press)] and the Colburn model of binaural processing [H. S. Colburn, *J. Acoust. Soc. Am.* **61**, 525–533 (1977)]. Predictions were generated for several spatial configurations of signal and masker in the horizontal plane (varying target and masker azimuthal positions and distances) and for multiple tone frequencies. Predictions were compared with behaviorally measured detection thresholds obtained with headphone simulations using individually measured HRTFs. Results show that changes in detection threshold for free-field signals are dominated by monaural effects due to large interaural level differences. Overall, spatial unmasking effects are very large for nearby sources due to simple energy effects at the better ear. Binaural unmasking is observable only for a subset of spatial configurations where the energy of signal and/or noise at both ears is sufficient to allow binaural processing, and even for these configurations the effect is generally weaker than that observed with distant sources. [Work supported in part by AFOSR Grant F49620-98-1-0108.]

3aPP2. Spatial unmasking for nearby speech sources in a simulated anechoic environment. Jason Schickler (Hearing Res. Ctr., Dept. of Biomed. Eng., Boston Univ., Boston, MA 02215), Norbert Kopčo, Barbara G. Shinn-Cunningham, and Ruth Y. Litovsky (Boston Univ., Boston, MA 02215, shinn@cns.bu.edu)

The “cocktail party effect” is traditionally studied with sources varying in azimuth, where binaural and level cues are large. In these conditions, speech intelligibility improves dramatically when the target speech and maskers are spatially separated. In the present study, we investigated spatial unmasking of nearby sources for several configurations of target and sources varying in distance and azimuth. Targets were sentences from the IEEE corpus and maskers were speech-shaped noise. Locations were simulated over headphones using anechoic head-related transfer functions (both individually measured and using a spherical-head model). Speech reception thresholds were measured adaptively, varying target level while keeping masker level constant. For nearby sources, overall energy effects

(due to changes in target and masker level with distance) are large. Therefore, target level was normalized to equate the signal-to-noise ratio at the better ear in all conditions. This approach allowed direct measurement of binaural unmasking in the various conditions. Overall energy effects (removed by the normalization) are also reported. Results are compared with detection threshold measurements from similar experiments. [Work supported in part by AFOSR Grant F49620-98-1-0108 to BGSC and NIDCD Grant DC02696 to RYL.]

3aPP3. Localization of ripple-spectrum noise. Ewan A. Macpherson (Kresge Hearing Res. Inst., Univ. of Michigan, Ann Arbor, MI 48109)

Ripple-spectrum stimuli were used to investigate the level of spectral detail important for use of spectral cues for vertical-plane localization. Free-field localization judgments were obtained for 250-ms, 0.5–16 kHz noise bursts with various log-ripple spectra. Ripple density was varied between 0.25 and 8 cycle/octave at a depth of ± 20 dB. Depth was varied between ± 5 and ± 20 dB at a density of 1 cycle/oct. Multiple ripple phases were employed at each combination of density and depth. ± 20 -dB rippled spectra substantially increased errors in vertical-plane localization in the range 0.5–2 cycle/oct. When systematic errors occurred, judgments were biased toward the location at which the listener’s directional transfer function (DTF) best matched the rippled source spectrum filtered by the target-location DTF. At 1 cycle/oct, localization accuracy was degraded only for ripple depths at or above ± 15 dB. Although the upper limit for ripple discrimination is 10 cycle/oct [Supin *et al.*, *J. Acoust. Soc. Am.* **106**, 2800–2804 (1999)], these results indicate that detail finer than 2 ripples/oct does not strongly influence spectral cue processing. This has significance for localization modeling and for virtual-stimulus synthesis. [Work supported by NIH Grants R01DC00420 and T32DC00011.]

3aPP4. Free-field and virtual studies of the precedence effect in the median-sagittal plane: Duration effects. Gerald Ng, Roberto Dizon, Ruth Litovsky, and H. Steven Colburn (Hearing Res. Ctr., Boston Univ., 44 Cummington St., Boston, MA 02215)

The study of precedence in the median-sagittal plane is made difficult by poor localization ability in this plane, as well as by front-back confusions. We report on both free-field and virtual localization studies in which

a source and simulated reflection are presented from two speakers chosen from an array of six, all of which are in the frontal median–sagittal plane. Subjects perform a forced-choice identification of which speaker emitted the combined stimulus. Stimuli consisted of lead–lag pairs of 1-, 10-, 25-, or 50-ms broadband noise bursts, where the onset of the reflection lags that of the source by 2 ms. Single-source cases were presented intermixed with the lead–lag paired stimuli. In the free-field condition, identification accuracy increased with duration for the single-source stimuli. For the precedence stimuli, a majority of responses is biased towards the elevation of the leading stimulus. However, details of individual performance show some evidence of response bias towards certain elevations, bimodal distributions, and perceptual averaging of lead and lag locations. Virtually presented stimuli using individualized HRTFs yielded localization judgments comparable to free-field stimuli for single-source conditions, but not for lead–lag pairs. [Work supported by NIH (DC02696 and DC00100) and ONR-MURI (Z883401).]

3aPP5. Investigations of the precedence effect with narrow-band continuous noise. Roberto Dizon and H. Steven Colburn (Hearing Res. Ctr. and Biomed. Eng. Dept., Boston Univ., Boston, MA 02215)

The precedence effect has been shown to be strong for stimuli with onset transients, e.g., clicks. Zurek [J. Acoust. Soc. Am. **67**, 952–964 (1980)] showed that a continuous, broadband noise whose fine structure leads that of a simulated reflection dominates lateralization even when onset gating cues have been removed. We report on experiments similar to those of Zurek, but using narrow-band noises of varying bandwidths and center frequencies. In a 2IFC headphone task, the leading and lagging stimuli are placed on either side of midline ($\pm 300\text{-}\mu\text{s}$ ITD) in the first interval, and the signals to the two ears are switched in the second interval. Subjects report whether the stimulus in the second interval is to the right or left of the first. Results are consistent with a precedence effect for bandwidths as narrow as a critical band. For narrower bandwidths, onset dominance is still observed, but with oscillations in lateralization consistent with the interaural differences expected for narrow-band stimuli, given the ITDs of the lead and lag stimuli and the lead–lag delay. As bandwidths approach zero, lateralization is dominated by these interaural differences, and onset dominance is not evident. [Work supported by NIH (DC00100) and ONR-MURI (Z883401).]

3aPP6. Optimal integration of directional information across a population of simple MSO neurons. Sergei Lubensky (Dept. of Elec. Eng. and Computer Sci. Res. Lab. of Electron., MIT, 77 Massachusetts Ave., Cambridge, MA 02139), Barbara G. Shinn-Cunningham, H. Steven Colburn (Boston Univ., Boston, MA 02215), and Nathaniel I. Durlach (MIT, Cambridge, MA 02139)

A simple computational model of source laterality was developed that optimally combines information across a population of neurons in the medial superior olive. A simple point-neuron MSO model was used to simulate the output of each of the interaural-phase sensitive neurons in the population. Each neuron has a best frequency and best interaural phase delay (IPD), generating an intermediate display similar to that proposed by Colburn [J. Acoust. Soc. Am. **61**, 525–533 (1977)]. Information in this population was combined (i.e., across best frequency and best IPD) to determine the *a posteriori* probability of a source from each possible angle relative to the interaural axis. From this output, source laterality was predicted from the maximum likelihood estimate of source direction. Similar to previous models [R. M. Stern and C. Trahiotis, Proceedings 11th International Symposium on Hearing (1997), pp. 336–345], this approach looks for consistency in the neural responses across frequency. However, the current model explicitly weights information corresponding to different best frequencies and IPDs based on the uncertainty in the responses of

the neurons in the population. Lateralization predictions were generated for a variety of sources and compared with results from the literature. [Work supported in part by AFOSR Grant No. F49620-98-1-0108.]

3aPP7. Auditory search asymmetry with pure tone, narrow-band noise, amplitude-modulated tone, and frequency-modulated tone. Noriaki Asemi (R. I. E. C./ G. S. I. S., Tohoku Univ., 2-1-2 Katahira, Aoba-Ku, Sendai 980-8577, Japan), Yoichi Sugita (Natl. Inst. of Bioscience and Human-Technol., Japan), and Yôiti Suzuki (Tohoku Univ., Sendai, Japan)

To examine basic features of sounds used in the perception of sound environment, auditory search tasks seem promising to us. Thus the response time to detect a target sound among distracting sound(s) was measured by three experiments. In the first experiment, pure tones and narrow-band noises (1/12 oct) were used. Either of them was used as a target and the other was used as distractor(s). The number of the distractor was varied from 1 to 6. Each sound was presented from one of seven loudspeakers located on the horizontal plane. In the second and third experiments, pure tones and amplitude-modulated tones, pure tones and frequency-modulated tones, respectively, were used as either a target or distractor(s). Results show that the response time to detect a narrow-band noise, an amplitude-modulated tone, or a frequency-modulated tone among distracting pure tones was hardly affected by the number of the distractors. However, the time required to detect a pure tone increased with the number of distractors. These results indicated that our auditory system utilizes temporal changes in amplitude and frequency of sound as a basic feature for the detection of a sound in a sound environment. Precise experimental results as well as further consideration will be given in the presentation.

3aPP8. Perception of a secondary auditory image with three sound sources. Bernard T. G. Tan, Sing Hai Tang, and Gongqiang Yu (Phys. Dept. of NUS, Low Kent Ridge Rd., Singapore 117543, osa-head@nus.edu.sg)

The precedence effect with three sound sources, using loudspeakers placed in a horizontal plane at equal distances facing the subject, was investigated. In this three-source sound system, one of which is undelayed and the other two delayed, a listener facing these sources perceives an unexpected secondary auditory image in addition to the expected primary auditory image. The two delayed sources have time delays relative to the undelayed source within the range in which the precedence effect operates. The secondary auditory image is believed to be due to summing localization taking place between the two sound sources, which have been delayed, when the relative time delay between them is within the range for summing localization to occur.

3aPP9. The use of off-frequency information in a high-frequency binaural discrimination task. Steven van de Par (Philips Res. Labs. Eindhoven, Prof. Holstlaan 4, NL-5656 AA Eindhoven, The Netherlands and IPO-Ctr. for User-System Interaction, Eindhoven, The Netherlands), Constantine Trahiotis, and Leslie R. Bernstein (Univ. of Connecticut Health Ctr., Farmington, CT 06030)

The ability of listeners to discriminate between N_0S_π and N_0S_0 stimuli was measured as a function of signal-to-noise ratio. Three types of Gaussian-noise maskers were employed: (1) broadband (300 to 6000 Hz) (2) high-frequency (2900 to 6000 Hz), or (3) low-frequency (300 to 3100

Hz). The signal was a 3-kHz sinusoid. The overall patterning of the data supports the hypothesis that listeners can and do utilize “off-frequency” information to enhance binaural performance. The data also suggest that the off-frequency information within auditory filters centered above the frequency of the signal is especially salient.

3aPP10. Perceptual segregation of individual components from a complex tone via changes in interaural time delay (ITD). John F. Culling (School of Psych., Cardiff Univ., P.O. Box 901, Cardiff CF10 3YG, UK)

When all the components of a complex tone share a common interaural time delay (ITD), a single component can be made to perceptually segregate from the complex by shifting its ITD to a unique value. Segregation of different components in succession can evoke the perception of a melody [Kubovy *et al.*, *Science* **186**, 272–274]. A series of five experiments demonstrated that (1) the ITD transition is more responsible for the segregation effect than the uniqueness of the ITD of the segregated component, (2) independent change in ITD of one component relative to coherent change among the other components is not sufficient to cause segregation, (3) cyclic modulation of ITD causes segregation which is independent of modulation rate, (4) temporal gaps in the stimulus of the order of 100 ms, and centered on the ITD transitions, can prevent segregation. These findings are consistent with an explanation based on sensitivity to within-channel reductions in interaural correlation within a sliding temporal window of around 100-ms duration; when different ITDs are temporally juxtaposed within a time frame of around 100 ms, these ITDs are encompassed by the window's span and the wave forms within it do not correlate perfectly at either delay.

3aPP11. Perception of acoustic occlusion using body-scaled judgments. Michael S. Gordon and Lawrence D. Rosenblum (Dept. of Psych., Univ. of California, Riverside, Riverside, CA 92521, rosenblu@citrus.ucr.edu)

Occlusion has been found to compel visual perception of object position and scene analysis [G. A. Kaplan, *Percept. Psychophys.* **6**, 193–198 (1969); W. H. Warren and S. Wang, *J. Exp. Psychol.* **13**, 371–383 (1987)]. Auditory occlusion may also provide salient information about the locations of both sound sources and surfaces in an acoustic scene. In the current experiment, blindfolded participants were asked to listen to sound occluded by an aperture and judge whether unobstructed passage was afforded for various aperture sizes (e.g., W. H. Warren and S. Wang, 1987). Judgments were found to be highly accurate, suggesting the salience of occlusion information for auditory perception. Information from the relative loudness between aperture sizes may have contributed to this ability.

3aPP12. Head-slaved tracking of an acoustic target by human listeners. Robert S. Bolia and W. Todd Nelson (Air Force Res. Lab., 2255 H. St., Wright–Patterson AFB, OH 45433-7022)

Numerous researchers have investigated the extent to which humans are able to track a visual target simply by following its motion with corresponding motions of the head. The results of these investigations have many practical applications, including vehicle guidance and the tracking and designation of moving targets. In spite of the amount of work done in

this area, few, if any, studies have addressed the question of how well humans are able to track an acoustic target. This is surprising, given that, in addition to its potential advantages for multi-sensory tracking displays, head-slaved tracking of acoustic targets might prove utile as a means of navigation for the visually impaired. In the present study, listeners performed a head-slaved tracking task of an acoustic target the motion of which was constrained to the horizontal plane, under factorial combinations of field-of-regard ($\pm 30, 45$, or 90 degrees), spatial resolution of the audio display ($3, 7$, or 10 degrees), and the presence or absence of visual feedback (a color change in the visual display). Results will be discussed in terms of potential applications, and compared with performance on visual tracking tasks. [Work supported by the Air Force Office of Scientific Research.]

3aPP13. Velocity DLs for a simulated sound source: Effect of intensity. Sarah Hassett and Lawrence Feth (Dept. of Speech and Hearing Sci., 110 Pressey Hall, Ohio State Univ., Columbus, OH 43210)

The Doppler effect refers to the frequency change observed for a moving sound source. The listener will observe a higher than emitted frequency as the moving source approaches, and a lower frequency as it recedes. Previous work [J. Acous. Soc. Am. **106**, 2209 (1999)] determined the difference limen (DL) for the velocity of a computer-simulated moving sound source. The simulated source moved along a straight path that passed within 5 m of the listener. Subjects were tested at 500, 1000, and 4000 Hz for reference velocities of 1, 2, 4, and 8 m/s. Signal duration was 500 ms. A 2Q,2AFC adaptive procedure determined the listener's velocity DL. The results indicated that DLs for both velocity and inferred frequency increase as reference velocity decreased, indicating that Weber's law does not hold for either parameter. The influence of source intensity variations for the 5-m path were minimal, thus the current experiment repeated the conditions of the prior experiment except that the path passed 0.5 m from the location of the observer. The effects of simultaneous frequency and intensity changes were investigated and compared with model predictions. [Work supported by a grant from the Ohio State University College of Social and Behavioral Sciences.]

3aPP14. Influence of visual cues on the perception of surround sound. Andrzej Czyzewski, Artur Kornacki, Bozena Kostek, Piotr Ody, and Slawomir Zielinski (Sound Eng. Dept., Tech. Univ. of Gdansk, 80-952 Gdansk, Poland, kid@sound.eti.pg.gda.pl)

Contemporary digital video, film, or multimedia presentations are often accompanied by the surround sound. Techniques and standards involved in digital video processing are much more developed than concepts underlying creating recording and mixing of the multichannel sound. The main challenge in the sound processing in the multichannel system is to create an appropriate basis for connecting multimodal context of visual and sound domains. Therefore one of the purposes of experiments is to study in which way and how the surround sound interferes or is associated with the visual context. This kind of study was hitherto carried out when two-channel sound technique was associated with a stereo TV. However, there are not many studies done yet that associate surround sound and digital video presented at the TV screen. The main issue in such experiments is the analysis of the influence of visual cues on perception of the surround sound. This problem will be addressed in the paper.

Session 3aSA

Structural Acoustics and Vibration and Noise: Combining Active and Passive Control of Vibration and Noise II

Alain C. Berry, Cochair

Department of Mechanical Engineering, University of Sherbrooke, Sherbrooke, Quebec J1K 2R1, Canada

Thomas J. Royston, Cochair

Department of Mechanical Engineering, University of Illinois, 842 West Taylor, Chicago, Illinois 60607-7022

Invited Papers

8:30

3aSA1. Active stabilization of liquid capillary bridges using optically sensed modal amplitudes. David B. Thiessen, Mark J. Marr-Lyon, and Philip L. Marston (Dept. of Phys., Washington State Univ., Pullman, WA 99164-2814)

Cylindrical capillary bridges consisting of liquid between two circular supports naturally become unstable and break when the length of the bridge exceeds its circumference for the situation where the weight or buoyancy of the bridge can be neglected. This is the Rayleigh–Plateau (RP) slenderness limit which is relevant to the management of liquids and to the formation of liquid drops. We have demonstrated methods of suppressing this instability based on the optical sensing of the instantaneous modal amplitude and the rapid adjustment of the axial distribution of applied radial stress. This applied stress may be the result of ultrasonic radiation pressure [M. J. Marr-Lyon *et al.*, *J. Fluid Mech.* **351**, 345–357 (1997)] or electrostatic stresses from an array of electrodes [M. J. Marr-Lyon *et al.*, *Phys. Fluids* (accepted)]. We have stabilized bridges as much as 42% beyond the RP limit by phasing the stress so that the effective modal spring constant becomes positive; however, the modal damping is then decreased. The closed-loop response includes terms associated with boundary-layer damping and the sensor's response time. The analysis indicates that damping of short bridges may be enhanced by reversing the feedback phase. [Work supported by NASA.]

8:50

3aSA2. Active structural intensity control using strain sensing. Pascal Audrain (G.T.V., I.A.M., Université du Maine, 72085 Le Mans Cedex 9, France, audrain@laum.univ-lemans.fr), Patrice Masson, and Alain Berry (Université de Sherbrooke, Sherbrooke, PQ J1K 2R1, Canada, Patrice.Masson@gme.usherb.ca)

An investigation of structural intensity control using strain sensing is presented in this paper. As opposed to previous work, the approach taken here does not assume waves propagating predominantly in one direction. Moreover, the instantaneous intensity is completely taken into account in the control algorithm, i.e., all the terms are considered in the real-time control process and, in particular, the evanescent waves are considered in this approach. A finite-difference approach using discrete PVDF strain sensors is used as the sensing scheme. A feedforward filtered-X LMS algorithm is adapted to this energy-based control problem, involving a nonpositive definite quadratic form in general. In this respect, the approach is limited to cases where the geometry is such that the intensity component will have the same sign for the control source and the primary disturbance. Experimental validation of the approach is conducted on a free–free beam covered with viscoelastic material. A comparison is made between classical acceleration control and structural intensity control and the performance of both approaches is presented. These results tend to indicate that using intensity control allows the error sensors to be placed closer to the control source and the primary disturbance, while preserving a good control performance.

Contributed Papers

9:10

3aSA3. Progress toward a “smart acoustic blanket.” Joseph A. Bucaro, Brian H. Houston, Thomas Howarth, Robert Corsaro, James Tressler, and Nicholas Lagakos (Naval Res. Lab., Washington, DC 20375-5350)

This paper presents recent progress we have made toward the goal of developing what we call “smart acoustic blankets.” Such technology would have application in many acoustic areas including aeroacoustics, architectural acoustics, and traffic sound barriers in addition to our principal focus which is the reduction of launch noise sound transmission through satellite payload fairings. These smart blankets would be configured with sound actuation layers, acoustic pressure and velocity sensors, and electronic controller modules. The status of the development of each of these components as well as the overall concept will be discussed. [Work supported by ONR.]

9:25

3aSA4. Hybrid tool for quick characterization of multi-layered panels. Ibrahima Sow, Olivier Beslin, and Jean Nicolas (GAUS, Dept. Genie Meca., Univ. of Sherbrooke, Sherbrooke, PQ J1K 2R1, Canada, isow@vulcain.gme.usherb.ca)

With the advancement of science and technology, materials with high damping capabilities and high modulus of elasticity are increasingly popular as they reduce vibration and noise. However, most of the time, there is a lack of tools that allow quick and easy characterization of the effective materials properties when used in 2-D structures. In this paper, a method is proposed for the characterization and analysis of viscoelastic composite material consisting of a three-ply sandwich panel. The method is based on a quick PC-based numerical code (involving hierarchical finite element method coupled with Beranek's formulation of sandwich panels) and a dedicated simple experimental setup. This approach ensures (i) proper extraction of the equivalent properties such as the modulus of elasticity,

the loss factor according to temperature, or frequency of the entire sandwich composite, (ii) characterization of any core viscoelastic resin, and (iii) evaluation of vibroacoustics indicators. The method is validated using commercially available 3M ISD112 and good agreement is found between experimental and numerical results. The method is then used to carry out a parametric study for different panel configurations according to temperature and frequency in order to characterize a viscoelastic resin.

9:40

3aSA5. Electro-elastic laminate theory for discrete piezoelectric patches on laminated plates. Senthil Gopinathan, Vasundara Varadan, and Vijay Varadan (Penn State Univ., 149 Hammond Bldg., University Park, PA 16802)

Classical laminated plate theory (CLT) has been applied successfully in the past to laminates with discrete piezoelectric patches bonded to the surface or embedded within the layers (1–7). The basic assumptions made in the earlier models were that the strains inside the patches are assumed to be constant and hence the presence of the sensor and actuator patches were neglected while modeling the dynamic properties of the laminate. The validity of these assumptions, the effect of the size of the patches and these assumptions on the solutions obtained, has not been studied. In this paper, the CLT is applied to a laminate with surface-bonded piezoelectric patches without the above-mentioned assumptions. A detailed modeling of the patches is developed by expressing the electric potential inside the patch as a quadratic function of thickness coordinate. The equations of motion are derived for a generally orthotropic laminate and solution method for these equations. Analytical solutions are obtained for a plate bonded with one and five collocated piezoelectric actuator/sensor patches. The effect of the passive and active stiffness of the surface bonded actuator and sensor patches on the dynamic characteristics of host plate structure is studied.

9:55

3aSA6. Structural intensity in a line-excited infinite elastic plate in the high-frequency range. Sabih I. Hayek and Jungyun Won (Active Vib. Lab., Penn State Univ., 227 Hammond Bldg., University Park, PA 16802 sihesm@engr.psu.edu)

An infinite elastic plate is excited by a mechanical line force, which generates an active structural intensity field in the plate. In an earlier paper [Hayek *et al.*, *J. Acoust. Soc. Am.* **105**, 1299 (1999)], the structural intensity was evaluated analytically for a plate using the Bernoulli–Euler theory, which is valid in the low-frequency range. The paper then employed the same theory for the active control of total structural intensity by use of collocated point forces and moments. In this paper, the evaluation of the structural intensity field due to line forces and moments is modeled by use of the more exact Mindlin plate theory, which is valid in a much higher frequency range. The main objective is to achieve active control of the total structural intensity in infinite elastic plates using the Mindlin plate theory through a collocated line force and a line moment actuators at an arbitrary location on the plate.

10:10–10:30 Break

10:30

3aSA7. Impact of piezoceramic transducer hysteresis in the hybrid structural vibration control problem. Soon-Hong Lee and Thomas J. Royston (Dept. of Mech. Eng., MC251, Univ. of Illinois at Chicago, 842 W. Taylor St., Chicago, IL 60607)

Hybrid vibration control via passive electrical shunting and active actuation of a piezoceramic wafer bonded onto the surface of a simply supported beam is investigated theoretically and experimentally with emphasis on modeling and understanding the impact of the piezoceramic hysteresis. A nonlinear rate-independent hysteresis model is experimentally identified for the PZT wafer by itself and then integrated into the coupled dynamic equations of the overall system consisting of the beam and electrically shunted PZT wafer. Experimental studies of the system

validate the theoretical model. This model is then used to investigate the impact of PZT hysteresis on its vibration control performance in passive, active, and hybrid scenarios. [Work supported by grants from ONR and NSF.]

10:45

3aSA8. A new active structure acoustic control system. Dingguo Zou and Malcolm J. Crocker (Dept. of Mech. Eng., Auburn Univ., Auburn, AL 36849, zouding@eng.auburn.edu)

The acoustic power radiated from any vibrating structure can be written as a quadratic form of the structural modal velocities. The quadratic matrix is real, symmetric, and positive definite. All of the terms with even–odd and odd–even indices are zero. The congruent transformation instead of orthonormal transformation was used to diagonalize the matrix. This method takes advantage of the good properties of the quadratic matrix more efficiently than the orthonormal transformation does. Also, the method greatly simplifies the design of the PVDF sensors, which filter the combination of appropriately weighted vibration modes, providing a specific performance index in control strategy. Based on this method, an active control system was implemented. To minimize the radiated power from a vibrating structure, a procedure to optimize the distribution of the actuators was applied to the control system.

11:00

3aSA9. Sensing radiated acoustic power using smart sensors. Dingguo Zou and Malcolm J. Crocker (Dept. of Mech. Eng., Auburn Univ., Auburn, AL 36849, zouding@eng.auburn.edu)

This paper represents the radiated acoustic power from a vibrating structure as a quadratic form of the structural modal velocities. The weighting matrix is diagonalized by using the congruent transformation. A group of one-dimensional shaped smart sensors which can sense the acoustic power from the structure by filtering the combination of weighted vibration modes are designed. First, it was found that if the shaped function is chosen carefully, the measuring from a two-dimensional distributed parameter sensor using PVDF film is independent of the position. Then, the plate is vertically divided by some narrow strips. The narrow strips are assumed to be approximate one-dimensional plate. The one-dimensional shaped sensor equations can be used to measure a particular modal coordinate. The acoustic power can be obtained by weighting the measurements from multiple one-dimensional smart sensors. Finally, the approximate results are adjusted further by using the theoretical formulation.

11:15

3aSA10. Active vibration control of a structure using virtual sensors. Vivian F. Dias and Scott D. Sommerfeldt (Dept. of Phys., Brigham Young Univ., Provo, UT 84602, vdias@physics.byu.edu)

Presently, a large safety factor is used in determining routine maintenance schedules for critical machinery/parts, such as replacement of skins for aircraft wings, because any failure during operation could prove catastrophic. Assuming that the location of the structural damage can be identified at its inception, one could apply active vibration control on the structure as a means of reducing stresses in the damage zone, thus ensuring that the damage does not grow rapidly with time. However, the probability of having an error sensor at the damage location is small, which indicates that a different approach would be required for implementing the active control. Work will be reported aimed at developing an active vibration control system that allows the control to minimize the stress or vibration at any arbitrary point/zone on the vibrating structure, even where there is no error sensor present. This is accomplished by using what are referred to as “virtual sensors.” These virtual sensors use information from physical sensors at other locations on the structure to estimate the vibration response at the virtual sensor location. This approach offers the possibility of being able to control the structural vibration at any desired, arbitrary location on the structure.

3aSA11. Active/passive design for vibration control. John A. Rule and Robert L. Clark (Dept. of Mech. Eng. and Mater. Sci., Duke Univ., Durham, NC 27708, jarule@duke.edu)

A method has been developed for the control of panel vibration using a hybrid active/passive spatial optimization technique. The method relies on the rapid estimation of Hankel singular values (HSV's) to predict coupling of sensor/actuator pairs to particular vibration modes of interest. The two-step process first requires the selection of the best sensor/actuator pair for active control, using a design metric which is typically used to target the control of low-frequency, long-wavelength modes, while simultaneously minimizing higher frequency, or out-of-bandwidth, system response. With the active control system in place, potentially destabilizing out-of-bandwidth modes are identified for passive control. A similar design metric is then applied which emphasizes coupling of passive damping material to the high-frequency, short-wavelength structural modes. Numerical simulations were performed to demonstrate this concept. It was then implemented on an 18- \times 16-in steel plate. Plate response was measured first for the uncontrolled plate, then the plate with just the active control system in place, and finally the complete hybrid active/passive system. Performance of each system was analyzed, with benefits and tradeoffs considered.

3aSA12. Active vibration isolation using variable structure control strategy. Jiaqiang Pan and Shuwen Pan (Dept. of Instrumentation Sci. and Eng., Zhejiang Univ., Hangzhou 310027, PROC)

In this paper, the authors present a new type of control design methodology used for active vibration isolation: Variable Structure Control (VSC) with sliding mode. VSC is a special control strategy that is capable of making a control system very robust with respect to system parameter variations and external disturbances, and providing an easy way to design the control law for a plant, linear or nonlinear. It has developed into a general control design method and been applied to a wide variety of engineering systems since the 1990s. In this study, a vibration isolation system with both active and passive isolators is designed based on Ackermann's method. Using this method, designed control law can make system states move in a time-variant sliding mode surface in state space from initial moment. This ensures the stability and robustness of the closed-loop system. Designed feedback control law can be expressed explicitly by state variables and parameters of the nominal system without any uncertainty and is easily applied in practice. Simulation results of an active suspension of vehicles show that the control system using VSC strategy possesses good performance.

THURSDAY MORNING, 1 JUNE 2000

HENRY/GRADY ROOM, 8:30 TO 11:30 A.M.

Session 3aSC

Speech Communication: Phonetics, Prosody and the Lexicon (Poster Session)

Michael D. Hall, Chair

Psychology Department, University of Nevada, 4505 Maryland Parkway, Box 455030, Las Vegas, Nevada 89154-5030

Contributed Papers

All posters will be on display from 8:30 a.m. to 11:30 a.m. To allow contributors an opportunity to see other posters, contributors of odd-numbered papers will be at their posters from 8:30 a.m. to 10:00 a.m. and contributors of even-numbered papers will be at their posters from 10:00 a.m. to 11:30 a.m.

3aSC1. Speech perception by normal-hearing listeners under modulated noise masker. Bom Jun Kwon and Christopher W. Turner (Dept. of Speech Pathol. and Audiol., Univ. of Iowa, Iowa City, IA 52242, bomjun-kwon@uiowa.edu)

Previous psychoacoustic investigations on modulation detection interference (MDI) show that sensitivity to detect modulation in a target is reduced when the masker is modulated. The amount of reduction is primarily dependent upon modulation frequencies, but not as much upon carrier frequencies, supporting the existence of a modulation filterbank in the auditory system. In the current study, consonant identification by normal-hearing subjects was tested with noise maskers, either steady-state or modulated. Modulation interference (MI) was assessed by comparing recognition under the two masker types. In experiment I, speech with limited envelope modulations was tested with different rates of masker modulation. In experiment II, narrow-band speech and masker were tested under various carrier frequency conditions. Higher MI was observed when the rate of masker modulation lies within the range of speech modulations. Carrier frequency separation did not make a strong effect on MI except when the separation is zero. In many cases MI was small or negative, implying that a masking release also occurred. In conclusion, both modulation masking and masking release should be taken into consideration when addressing speech perception under modulated maskers. Psychoacoustic MDI results, therefore, only partially predict the effects of modulation interference on speech recognition. [Work supported by NIDCD.]

3aSC2. An EPG and acoustic study of fricative production and spectra. Marija Tabain^{a)} (Speech, Hearing and Lang. Res. Ctr., Macquarie Univ., Sydney, 2109, Australia, marija@srusna.shlrc.mq.edu.au)

In order to test the correspondence between variability in fricative production and variability in the fricative spectrum, the spectral center of gravity (COG) is compared with the EPG center of gravity measured along the horizontal dimension. The CV coarticulatory effects in particular are examined. Data are taken from CV tokens produced by four female speakers of Australian English. Tokens consisted of the coronal fricatives /T s S D z Z/ in seven monophthong vowel contexts. Rounded vowels are excluded so as to minimize spectral changes due to lip-rounding. The consistency between the articulatory and the acoustic data is particularly striking for the palato-alveolar fricatives, although results are also highly consistent for the other two fricative places of articulation examined. However, it is not always clear that the variability in fricative production is conditioned by the vowel context. Overall, the sibilant fricatives seem particularly resistant to coarticulation and to variability in production, while the nonsibilant dental shows a good deal of variability, including some conditioning by the vowel context. These results have implications for the hyper- and hypo-theory of speech production. It is suggested that sibilant fricatives do not lend themselves to the articulatory imprecision,

which characterizes other perceptually salient, and typologically common, speech sounds. ^{a)}Currently at Perceptual Science Laboratory, University of California at Santa Cruz.

3aSC3. Comparative perception of voicing contrasts. Joan M. Sinnott and Laura A. McArdle (Psych. Dept., Univ. of South Alabama, Mobile, AL 36699, jsinnott@jaguar1.usouthal.edu)

English, Spanish, and monkey listeners were compared in their perception of a synthetic labial VOT continuum (−60 to +70 ms). A go-right/go-left ID procedure was used, with food reinforcement for monkeys. In exp. 1, subjects first learned to identify the endpoints of the continuum as “BA” or “PA,” and then generalization responses were obtained for intermediate stimuli. The mean boundary for English listeners was at +25 ms, while boundaries for Spanish listeners ranged from −15 to +25 ms (mean=7.6 ms). The monkey boundary was near 0 ms and appeared to simply bisect the continuum, suggesting that the English boundary at +25 ms is not a salient feature of the “general auditory system” in the context of prevoiced stimuli. In exp. 2, subjects were trained to differentiate VOT contrasts that crossed either the “English” or the “Spanish” boundary. While subjects learned to reliably differentiate both contrasts, reaction times were fastest for the English contrast. These results indicate that although attentional differences between subject groups may influence their generalization responses along a VOT continuum (exp. 1), these groups have similar sensory capacities when explicitly trained to differentiate specific stimuli along a VOT continuum (exp. 2). [Work supported by NIDCD.]

3aSC4. Coproduction in VCV disyllables produced by children and adults. Carole E. Gelfer (Dept. of Comm. Disord. William Paterson Univ., Wayne, NJ 07470, gelferc@wpunj.edu) and Fredericka Bell-Berti (St. John’s Univ., Jamaica, NY 11439)

Previous studies of speech production of young children have suggested differences from adult productions. For example, developmental differences have been noted for VOT and the extent of coarticulation between vowels and fricatives. This study explores developmental trends for consonant-vowel interactions in VCVs disyllables as evidenced by spectral differences in V1 as a function of C and V2 identity, where V1 is schwa, C is one of the following: /p, t, k, b, d, g, s, sp, st, sk/, and V2 is /i/ or /u/. The children in this study are between the ages of 4 and 6. Measures of F_2 at both the offset and midpoint of schwa will be used to determine (1) the extent to which interactions show developmental trends, and (2) whether the extent of these interactions for either or both groups of speakers is influenced by consonant manner of articulation, place of articulation, and/or duration.

3aSC5. Anticipatory coarticulation in the speech of adults and children: A perceptual study. Sneha V. Bharadwaj and William F. Katz (Univ. of Texas at Dallas, Callier Ctr. for Commun. Disord., 1966 Inwood Rd., Dallas, TX 75235)

Two competing theories attempt to explain anticipatory coarticulation in the speech of adults and children: One suggests more extensive coarticulation in children’s speech than in that of adults, while the other suggests more extensive coarticulation in adult speech [see S. Nittrouer and D. Whalen, *J. Acoust. Soc. Am.* **86**, 1266–1276 (1989) for discussion]. To address this issue, a perceptual experiment was conducted in which adult listeners identified vowel information from gated CV syllables. The syllables /si/, /su/, /ji/, and /ju/ produced by children (5 and 7 years old) and adults were gated into four segments of varying lengths ($\frac{1}{2}$ fricative, $\frac{3}{4}$ fricative, full fricative, and fricative + $\frac{1}{2}$ vowel). The stimuli were presented to ten adult listeners in randomized order (experiment 1) and to another ten listeners, blocked by individual talkers (experiment 2). Syllable identification for key fricative gates was 6% better in blocked compared to random presentation. Both experiments revealed a trend of listen-

ers detecting more extensive coarticulation in children’s productions than in those of adults; however, this pattern resulted from heightened responses to a single 7-year-old talker. Results suggest children coarticulate with greater variability than adults, but not necessarily with a greater temporal extent.

3aSC6. Threshold comparisons for intensity-based duplex perception for speech. Michael D. Hall (Psych. Dept., Univ. of Nevada Las Vegas, 4505 Maryland Pkwy., Box 455030, Las Vegas, NV 89154-5030, hallm@nevada.edu)

Duplex perception (DP) occurs when a stimulus component simultaneously contributes to two percepts (Rand, 1974). A proposed DP variant involves substituting a tone glide for an F_3 -transition in a /da/ or /ga/ syllable. At full intensity the glide is perceived as a nonspeech chirp that simultaneously distinguishes the consonant. It also has been claimed that the glide contributes to speech (consonant) perception over a range of intensities that are insufficient for nonspeech (chirp) detection, raising the possibility that a phonetic module takes precedence over nonspeech perception (Whalen and Liberman, 1987). Others have failed to find a range of precedence when chirp detection thresholds were reevaluated using a 2AFC task (Bailey and Herrmann, 1993). Which conclusion is valid? The current investigation addressed this question by comparing individual thresholds for consonant perception and chirp detection using a variety of methods. Chirp detection thresholds were evaluated using the method of constant stimuli, 2AFC, and various reminder tasks (using a fixed standard). Thresholds for consonant perception were evaluated using identification and discrimination (same-different) tasks. Results analyzed according to both traditional threshold and detection theories support the claims for a range of precedence and suggest that discrepant chirp detection thresholds reflect a reliance on consonant perception.

3aSC7. Vowel reproduction exhibits an intrinsic dynamic organization. Betty Tuller and Gautam K. Vallabha (Ctr. for Complex Systems & Brain Sci., Florida Atlantic Univ., 777 Glades Rd., Boca Raton, FL 33431, vallabha@walt.ccs.fau.edu)

Speech perception and production are generally thought to be tuned by the speaker’s linguistic environment. If so, iterative reproduction of a vowel by a single speaker may amplify subtle biases introduced by the tuning. Four male monolingual speakers of American English were asked to reproduce 100 synthetic vowel-like stimuli, uniformly distributed in a two-formant acoustic space (F_1 , F_2 varying, F_0 =120 Hz, F_3 =2500 Hz, duration=200 ms). Fifteen of the subject’s own productions were chosen for serial reproduction, as follows: Each original production was presented to the subject through headphones, and the subject immediately mimicked it. After a 0.5-s delay, the “mimic” was played back to the subject as the target for his next production. This second mimic was the target for the following production, and so on, for ten iterations. Subjects made systematic errors in reproducing their own vowel sounds, and sequences of serially reproduced vowels behaved differently in different parts of the vowel space. These results suggest that some vowel regions are preferred over other regions, indicating that the cognitive processes for perceiving and producing vowels have complex intrinsic dynamics. [Work supported by NIMH.]

3aSC8. Consonants and vowels behave differently in silent center syllables. A. Min Kang (Haskins Labs., 270 Crown St., New Haven, CT 06511 and Dept. of Linguist., Yale Univ., New Haven, CT 06520, min.kang@yale.edu) and D. H. Whalen (Haskins Labs., New Haven, CT 06511)

Compared to consonants, steady-state vowels show no right ear advantage, activate different brain regions, and only trigger categorical perception when extremely short. Auditory explanations for this distinction assume different perceptual mechanisms for brief, dynamic consonants and

long, slow-changing vowels. Gestural views suggest steady-state vowels are perceived as only partially speechlike. The present study compares consonant and vowel identification and discrimination when based on the same brief, dynamic information in the form of silent center (SC) syllables. The CVCs simultaneously varying in equal-sized $F2$ steps along both a /b-d/ and an /e-/ / continuum were synthesized (full syllables). The SC syllables were produced by replacing the middle 60% with silence. Relative to full syllables, overall SC consonant discrimination improved from chance to 70%; vowel discrimination dropped from 85% to 70%. The SC vowel identification was affected by consonantal context but the reverse effect was weaker. Improved SC consonant discrimination may result from decreased masking or greater prominence of the consonant gesture. Decreased SC vowel discrimination could result from reduced acoustic information or weaker prominence of the inferred vowel gesture. While these results do not distinguish auditory and gestural theories, they do place the consonant/vowel distinction in a new perspective. [Work supported by NIH Grant No. HD-01994.]

3aSC9. The role of sentential prosody in learning voices. Lynne C. Nygaard and Jennifer S. Queen (Dept. of Psych., Emory Univ., Atlanta, GA 30322, lnygaard@emory.edu)

The present experiment was designed to investigate the contribution of sentential prosody to the perceptual learning of talker's voice. Listeners were trained over a three-day period to learn talker's voices from three types of utterances: (1) sentence-length utterances produced with natural sentential prosody, (2) utterances in which individual words were produced in list format and digitally assembled into complete sentences, and (3) words produced in list format and scrambled to remove semantic and syntactic coherence. Equal numbers of male and female listeners learned a set of ten talker's voices (five male and five female) and learning was evaluated at each training session. On the fourth day of testing, listeners were given a generalization test that consisted of novel sentence-length utterances produced with natural sentential prosody. Results indicated that learning curves were comparable across conditions, but the ability to generalize to new utterances depended on the utterance type used during training. Individual differences among both listeners and talkers were also observed. Male listeners identified male voices significantly better than female voices. Female listeners identified male and female voices equally well. These results suggest that the perception of talker identity is dependent both on listener- and task-specific characteristics.

3aSC10. Perceptual adjustment to foreign-accented English. Constance M. Clarke (Dept. of Psych., Univ. of Arizona, Tucson, AZ 85721)

A central goal of speech research is to understand how the speech perception system handles variability. Previous work on variability due to differences in speakers' voices has demonstrated that experience with a voice facilitates word intelligibility [L. C. Nygaard and D. B. Pisoni, *Percept. Psychophys.* **60**, 355–376 (1998)]. The present study investigated whether a more abstract feature of voices, such as foreign accent, can also be perceptually learned and utilized in perception. The following question was tested: Does experience with an accent generalize to the perception of a new voice with the same accent? Two groups were trained with English sentences produced by either Spanish- or Chinese-accented voices. Both groups were then given a word intelligibility test in which sentences were presented in noise. Test sentences included both familiar and new Spanish- and Chinese-accented voices. Training improved intelligibility of familiar voices but did not affect perception of new voices with the trained accents. Results indicate that experience with an accent does not improve

perception of a new voice with the same accent. Perceptual learning seems to be restricted to individual voices. Results are discussed with respect to current theories of speech perception. [Work supported by NSF.]

3aSC11. Imitation of phrases in conversational speech. Jennifer S. Pardo (Psych. Dept., Yale Univ., P.O. Box 208205, New Haven, CT 06520, jennifer.pardo@yale.edu) and Carol A. Fowler (Haskins Labs., New Haven, CT 06511)

Linguistically significant gestures of the vocal tract play an important role in speech perception and production. Imitation of speech across conversational partners is a natural consequence of the close connection between perception and production. Following research that finds imitation in single-word shadowing, this study paired unacquainted same-sex talkers in a conversational task designed to induce repetition of key phrases across partners. Immediately repeated items were assessed for imitative fidelity by asking a separate set of listeners to choose the more similar item in an AXB task that compared the sample item from one talker (X) with a read and a repeated item from the other talker (A/B). In general, listeners chose the repeated over the read item as more similar to X. Additionally, read items that were sampled immediately after the conversation hindered a listener's ability to choose the repeated item more so than items read long before the conversation. Other effects related to the timing of items in the conversation and the role of the participants in the conversation were also observed.

3aSC12. Age-related differences in regularity of speech rate. Richard J. Morris (Dept. of Commun. Disord., Florida State Univ., Tallahassee, FL 32306-1200, rmorris@mailers.fsu.edu)

Older adults speak more slowly and exhibit greater speaking rate variability than do young adults [i.e., Liss *et al.*, *J. Geron.* **45**, P35–45 (1990); Ramig, *J. Commun. Disord.* **16**, 217–226 (1983)]. These findings have been reported for rates of speech, articulation, and reading as well as for individual consonants and vowels. However, one recent work found that their group of older adults exhibited greater regularity in their repetition of syllables [Watson and Fozo, *ASLHA Conv.* (November, 1999)]. These authors hypothesized that the older speakers may have less flexibility in the timing of their speech. The purpose of the present study was to test the hypothesis of reduced timing flexibility in the speech of older people. Ten older men and ten younger men served as subjects for this study. All subjects were in good general health and had normal hearing for their age. The subjects were directed to repeat a series of single syllable words, 'pat', 'bat', 'sack' and 'shack', at slow, normal and fast rates. The speech was recorded onto a DAT recorder and analyzed using the CSpeech software. Results are discussed in terms of age differences in speech timing across the three speaking conditions.

3aSC13. The role of prosody in scope relations. Mary Baltazani (UCLA Linguist. Dept., Los Angeles, CA 90024, marybalt@ucla.edu)

It is well known that sentences with two quantifiers or operators can be ambiguous depending on their scope relation. For example, "She didn't kick him because he screamed" can mean "His screaming prevented her kicking him" (because > not) or "His screaming was not the reason for her kicking him" (not > because). This paper reports on a production and a comprehension experiment on the role of prosody in disambiguating scope relations in Greek sentences. In the production experiment, five participants read aloud seven sentences each, containing negation and another operator ("because" or a quantifier like "two"), embedded in disambiguating contexts. Prosodic analysis of the utterances showed that speakers produced different prosodic structures for the two interpretations: a big prosodic boundary after negation to deliver the quantifier > negation interpretation, and no boundary for the reverse interpretation. In the comprehension study, 40 participants listened to the 35 sentences out of context and performed a forced-choice paraphrase selection task. The results

show a significant correlation between prosodic cues and scope disambiguation in Greek. This suggests that both speakers and listeners unconsciously make use of these prosodic cues to interpret sentences.

3aSC14. Lexical access in a divided attention task: Evidence from a new verbal transformation methodology. Peter W. Lenz, Richard M. Warren, and James A. Bashford, Jr. (Dept. of Psych., Univ. of Wisconsin—Milwaukee, Milwaukee, WI 53201, bashford@uwm.edu)

A recorded repeating word undergoes illusory changes (Verbal Transformations, or VTs) to forms consisting of syllables and words occurring in the listener's lexicon. When the repeating word is delivered dichotically with a long interaural delay, the listener hears two distinct lateralized images, each changing independently [see Warren and Ackroff, *Nature* (London) **259**, 475–477 (1976)]. The present study introduces a new methodology minimizing vigilance and memory demands posed by the traditional VT paradigm. Rather than continuous monitoring and reporting of changes occurring on each side, listeners simply reported the forms being heard when cued by a light flash. In keeping with earlier findings based upon continuous bilateral monitoring, the simultaneous organizations of the same word were found to change independently. Contrasting with earlier studies, however, the presence of a second image of the same word did not decrease the transition rate, presumably due to reduced task demands. Interestingly, the transition rate decreased dramatically when two different repeating words were presented. Implications of these findings regarding mechanisms employed in processing speech will be discussed. [Work supported by NIH.]

3aSC15. Phoneme restoration in infants. Rochelle S. Newman (Dept. of Psych., E11 Seashore Hall, Univ. of Iowa, Iowa City, IA 52242, rochelle-newman@uiowa.edu)

In a classic study, Warren (1971) replaced the medial /s/ in legislatures with a cough, and found that adult listeners heard the word as being complete. This auditory illusion demonstrates effects of both bottom-up and top-down information: it occurs only when there is a possible masking noise (it does not occur if the /s/ is replaced with silence), and it occurs at least in part because of the listener's knowledge that legislatures is a real word. Thus, it involves an interaction between multiple sources of information. The present study investigates whether infants will also show such an effect, and thus whether this interaction between levels of language processing is itself something which develops over time. Infants are shown videos of a cat and a dog in a preferential-looking paradigm, and hear the words kitty and doggie both in the clear, and with the medial stop consonant replaced by either noise or silence. Results examine whether infants will treat the words as whole (that is, match them to the appropriate referent) when the middle sound is replaced with noise as compared to when it is replaced with silence. [Work supported by NIH Grant R03 HD37822-01 to the University of Iowa.]

3aSC16. Phonetic metamorphosis of white noise: A stop or a fricative. Valeriy Shafiro (City Univ. of New York Grad. Ctr., Dept. of Speech and Hearing, 365 5th Ave., New York, NY 10016)

The role of spectro-temporal coherence, lexical status, and word position in the perception of speech in acoustic signals containing a mixture of speech and nonspeech sounds was investigated. Stimuli consisted of nine words in which either white noise was inserted only into the silent interval preceding the onset of vocalic transitions ambiguous between /p/ and /f/, or in which white noise was mixed with the entire speech signal. A control condition contained no noise. Ten adults were asked to decide whether they heard a word with a /p/ or /f/ as the phoneme in the initial, medial, or final word position. The subjects perceived /f/'s 85% of the time when noise was inserted into the silent interval, 47% when noise was presented over the entire word, and 1% in the control condition. These results indicate the importance of spectro-temporal coherence for perceiving speech

in a mixture of sounds. The lexical status and position of the critical phoneme in the stimuli also appeared to affect subjects' responses, although to a smaller degree. Theoretically, the results support the auditory scene analysis view of perceiving speech in heterogeneous signals, while they cannot be adequately explained by the motor theory of speech perception.

3aSC17. A comparison of the vowel sequence illusion and the verbal transformation effect. Jeffrey M. Cooley, Richard M. Warren, and James A. Bashford, Jr. (Dept. of Psych., Univ. of Wisconsin—Milwaukee, Milwaukee, WI 53201, bashford@uwm.edu)

When repeated sequences of steady-state vowels have durations of their phonetic components below the 100-ms threshold for identification of order, the phonemes lose their identity. The sequences are heard as syllables occurring in the listener's lexicon (the Vowel-Sequence Illusion). When listening to repeating words, listeners also hear illusory syllables occurring in their lexicon (the Verbal-Transformation Illusion). Verbal transformations require the decay (satiation) of the initial veridical organization before the repetition-induced shifts in criteria permit nonveridical forms to be heard: The Vowel-Sequence Illusion does not require initial verbal satiation, and hence permits repetition-induced criteria shifts to tap directly into the lexicon-based syllabary. The intrinsic differences in the nature of the phonetic stimuli producing the two illusions result in characteristic distinctions between the organizations heard in the two illusions. Among these distinctions is the spectral splitting of the vowel sequences into two simultaneous organizations, one consisting of components above and the other below the 1500-Hz crossover frequency encountered in other contexts. This and other aspects of the two verbal illusions will be compared and contrasted, and demonstrations of both will be available. [Work supported by NIH.]

3aSC18. The relative time course of neighborhood and lexical effects. Rochelle S. Newman (Dept. of Psych., Univ. of Iowa, E11 Seashore Hall, Iowa City, IA 52242), James R. Sawusch, Paul A. Luce (Univ. at Buffalo, Buffalo, NY 14260), and Amanda Aubin (Univ. of Iowa, Iowa City, IA 52242)

We have previously demonstrated that neighborhood density can have effects similar to those of lexical status in phoneme identification tasks. In a typical task demonstrating a lexical effect, subjects might hear two series, one ranging from "bag" to "pag," while the other varies from "pal" to "bal." Subjects are more likely to classify ambiguous stimuli from each series as members of the category that makes a word (they would classify ambiguous items as "b" in the bag–pag series but as "p" in the pal–bal series). In a similar task where none of the four endpoints were words, subjects classified ambiguous items as members of the category that had a greater neighborhood density. The current study examines the relative time course of these effects. We created both lexical and neighborhood series using the identical initial contrast, and presented them to the same set of listeners. Neighborhood effects occurred primarily in the fastest responses, while lexical effects occurred in the medium and slower responses. That is, neighborhood effects appear earlier in processing, and dissipate quickly, whereas lexical effects take longer to build up. These results will be discussed in terms of their implications for models of word recognition.

3aSC19. Separable effects of neighborhood density and phonotactic probability on word recognition in speech. Nathan R. Large and Paul A. Luce (Dept. of Psych. and Ctr. for Cognit. Sci., Univ. at Buffalo, 245 Park Hall, Buffalo, NY 14260-4110, nathan@deuro.fss.buffalo.edu)

Previous research on the effects of probabilistic phonotactics on spoken word recognition has typically confounded phonotactics and similarity neighborhood density. In the present research, the inhibitory effects of density and facilitative effects of probabilistic phonotactics on spoken

word processing were manipulated simultaneously. Four sets of monosyllabic words and nonwords were created by orthogonally combining two levels of neighborhood density and two levels of probabilistic phonotactics. Speed and accuracy were examined in three different experimental paradigms: single-word shadowing, A-X same-different matching, and speeded neighbor generation. Words and nonwords were either blocked by

lexicity or intermixed. Blocking and mixing of stimuli were used to manipulate participants' focus on lexical and sublexical levels of processing. In part, our results demonstrate simultaneous effects (inhibitory and facilitative) of both density and phonotactics. The implications of our results for models of spoken word recognition and lexical access will be discussed.

THURSDAY MORNING, 1 JUNE 2000

AMERICAN ROOM, 8:00 TO 11:15 A.M.

Session 3aSP

Signal Processing in Acoustics and Underwater Acoustics: Algorithms and Performance I

Charles F. Gaumond, Chair

Naval Research Laboratory, 4555 Overlook Avenue S.W., Washington, D.C. 20375-5320

Contributed Papers

8:00

3aSP1. Assessing the relative detection performance of matched-field processing and conventional beamforming. Brian H. Maranda and David J. Thomson (Defence Res. Establishment Atlantic, P.O. Box 1012, Dartmouth, NS B2Y 3Z7, Canada)

The primary advantage of matched-field processing (MFP) over conventional plane-wave beamforming is the ability to localize an acoustic source in a single data snapshot; for example, MFP can help to decide whether or not the source is submerged. It is to be expected that MFP will also provide improved detection performance over the plane-wave beamformer when the acoustic pressure field at the array exhibits the complicated behavior that is typically observed in the oceanic waveguide. Accurately quantifying the performance gain is difficult, owing to analytical problems in determining the exact performance of a detector that thresholds many correlated search cells. However, on the basis of simple statistical arguments concerning the effective dimensionality of the search spaces, it is claimed that a rough comparison of detector performance is provided by the ratio of the quadratic forms that are used as the decision variables. This ratio is computed for several array and ocean scenarios, including both a vertical line array in shallow water and a horizontal line array in deep water. Two-dimensional plots (depth and range) are used to illustrate the detection gain that can be provided by matched-field processing.

8:15

3aSP2. Theoretical first-order bias and second-order variance for localization estimates from matched-field processing methods: Randomized signal. Aaron Thode, Eran Naftali, and Nicholas C. Makris (MIT, 77 Massachusetts Ave., Cambridge, MA 02139)

The bias and variance of range and depth estimates obtained from matched-field processing (MFP) methods are difficult to obtain analytically. Instead, the Cramer-Rao (CR) bound is typically computed to estimate the minimum theoretical variance of a localization estimate. Recently, it has been shown [Makris and Naftali, *J. Acoust. Soc. Am.* **106**, 2 (1999)] that the first-order bias and second-order variance of a general maximum-likelihood estimator can be computed analytically, using higher-order asymptotics in a Taylor series expansion. When applied to the specific case of a matched-field processor with a deterministic signal buried in correlated waveguide noise, it was found that the CR bound underestimates the true theoretical variance for low signal-to-noise ratio (SNR) signals and small sample sizes. Here, the case of a randomized signal in correlated ocean noise is addressed, which represents another facet of the generalized maximum-likelihood estimator problem. Similar

to the deterministic case, it is found that the second-order variance term is significant even at 20 dB SNR, and that the localization bias becomes significant as the SNR approaches zero.

8:30

3aSP3. Performance analysis for a short vertical array in a shallow-water waveguide. Tainfu Gao, Lan Wang (Inst. of Acoust., Chinese Acad. of Sci., Beijing, PROC), and E. C. Shang (Univ. of Colorado/NOAA/ETL, Boulder, CO 80303)

Vertical line array (VLA) has been widely used for field data collection in underwater acoustics. The performance of VLA strongly depends on: (1) the signal processing approach used, (2) the environmental knowledge, (3) the signal-to-noise ratio (S/N), and (4) the configuration of the array (length, position, curvature). In this paper, we concentrate our investigation on the length effect. Theoretical analysis has been done for a short array in a Pekeris shallow-water waveguide. For lower frequency in shallow water, the information provided by a VLA can be described by the capability of mode decomposition, which is determined by the spectrum of the eigenvalue of the "covariance matrix" of the field sampled by the VLA. The approximate decay behavior of the spectrum is calculated analytically for short array. It is shown that the estimated eigenvalues are very close to the exact eigenvalues. The indicator of the performance used in this paper is the number of "S," which represents the dimension of the "signal space," and can be easily determined by the spectrum decay law.

8:45

3aSP4. Weiss-Weinstein bound for matched-field parameter estimation. Wen Xu and Arthur B. Baggeroer (MIT, Cambridge, MA 02139, wenxu@mit.edu)

Matched-field parameter estimation, including source localization and environmental parameter estimation, is often implemented based on an ambiguity surface. The ambiguity surface is derived from the correlation between the observed signal field and the modeled signal field, and is often characterized by a multimodal structure since the typical signal field is a highly nonlinear function of the embedded parameters. In the presence of noise or other interferences, a high sidelobe level leads to a large error probability in parameter estimation. This sidelobe effect increases as the signal-to-noise ratio (SNR) decreases such that, below a threshold SNR, the estimation mean-square-error (MSE) is dominated by the error around the sidelobes, which is not predicted by the Cramer-Rao bound (CRB). To analyze this effect, the Weiss-Weinstein bound (WWB) is presented for the attainable performance of matched field methods. The bound is expressed in terms of the SNR and an ambiguity function defined based on the Green's function both at the receiver array. An evaluation example

demonstrates how the ambiguity behavior, particularly the sidelobe structure, affects the bound's evaluation in different SNR regions and leads to the well-known threshold phenomenon in nonlinear parameter estimation.

9:00

3aSP5. Matched-field processing in shallow water in the presence of soliton packets. Catherine Stamoulis (MIT Dept. of Ocean Eng., 77 Massachusetts Ave., Cambridge, MA 02139 and Naval Res. Lab., Washington, DC 20375, caterina@quake.mit.edu), Marshall Orr (Naval Res. Lab., Washington, DC 20375), and Ira Dyer (MIT, Cambridge, MA 02139), and SWARM Group (Naval Res. Lab, Univ. of Delaware, Naval Postgrad. School, Woods Hole Oceanogr. Inst.)

The performance of matched-field processing algorithms in shallow water, in the presence of soliton packets, has been investigated through analysis of data from the SWARM 95 experiment. In particular, acoustic data collected with a vertical array and a moving up-sweep source have been analyzed, in the frequency range 275–350 Hz. The source was placed below the surface mixed layer and its distance from the array varied between 2.5 and 27 km. Source localization using the conventional and minimum variance matched-field algorithms has been performed, in order to assess the robustness of these processors in the presence of environmental mismatch and their degradation as a function of range. The primary result of the analysis is that both processors fail to localize the source accurately, even at the shortest range (2.5 km). Although broadband processing and inclusion of corrections for the array tilt improved the results, the errors in source range (of the order of 1 km at short ranges and 5 to 10 km at long ranges) and depth remain significant. The performance of other matched-field processors is currently investigated.

9:15

3aSP6. Detection of impulsive signals in shallow water. Peter G. Cable (BBN Technologies, Union Station, New London, CT 06320)

The effect of multipath induced time spread on the detectability of impulsive signals transmitted in shallow water is considered. The effective signal duration and temporal shape of received signal energy, two related features of the pulse spread associated with propagation, were analyzed for the shallow-water coastal environments in the DARPA Area Characterization Tests (ACT I, II, and III). Measured multipath induced time spreads were between 30 and 40 ms and, for each environment, the integrated received energy was found to be well fitted by an exponential function of the integration time. Using the observed signal energy durations and shapes, a model of energy detection of impulsive signals in shallow water was formulated and will be described. An analysis of expected performance (probability of signal detection for fixed probability of false alarm) as a function of input signal-to-background energy ratio and integration time will be presented.

9:30–9:45 Break

9:45

3aSP7. First-order bias and second-order variance of maximum likelihood time delay and Doppler shift estimates. Eran Naftali and Nicholas C. Makris (MIT, 77 Massachusetts Ave., Cambridge, MA 02139)

Exact expressions for the bias and variance of a maximum likelihood time delay and Doppler shift estimate are difficult to obtain analytically. It has become popular in such nonlinear problems to compute limiting bounds on mean-square estimation error, such as the Cramer–Rao bound (CRB), since these are easier to obtain than the true variance. Recently it has been shown [Makris and Naftali, J. Acoust. Soc. Am. **106**, 2 (1999)] that the first-order bias and second-order variance of a general maximum-likelihood estimator (MLE) can be computed analytically, using higher-order asymptotics. This approach is applied to the classic radar/sonar problem of estimating the time delay and Doppler shift of a deterministic signal in additive white Gaussian noise. The MLE is shown to be unbiased to first order for this problem. By evaluating the asymptotic expansion for

the variance for three specific signal waveforms, Gaussian, LFM, and HFM, it is found that (a) the CRB yields an unrealistically optimistic variance estimate for SNR's lower than 20–30 dB, (b) both first- and second-order time-delay variance terms decrease with increasing signal bandwidth, and (c) both first and second Doppler-shift variance terms of LFM and HFM signals approach constant values as bandwidth increases.

10:00

3aSP8. Sensitivity analysis of an estimator-correlator. Roger W. Schwenke and Leon H. Sibul (Appl. Res. Lab., The Penn State Univ., P.O. Box 30, State College, PA 16804, rws143@psu.edu)

The estimator-correlator (EC) is a maximum likelihood-sense optimum detector for detection of objects that have multiple random highlights rather than a single deterministic highlight. The EC is also a model-based signal processor that uses target scattering function as an *a priori* statistical model. This EC has uses in sonar, medical ultrasound, and multipath communication channels. The receiver operator characteristics (ROCs) of this EC are derived. A sensitivity analysis is derived which computes the change in the ROC curve when there are errors in the *a priori* estimate of the scattering function. Examples of common scattering function errors and the resulting change in sensitivity are given. [This work has been supported by ONR, Code ONR333, Ms. Khine Latt Program Officer.]

10:15

3aSP9. Relating modem performance to the ocean channel properties. Michael Porter (Science Applications Intl. Corp., 888 Prospect St., La Jolla, CA 92037, michael.b.porter@saic.com), Vincent McDonald, Joseph Rice, and Paul Baxley (SPAWAR SSC, San Diego, CA 92152)

The ocean channel plays an important role in the performance of acoustic modems. However, the precise effects of such features as the sea surface, ocean bottom, bubble clouds, and internal waves is not well understood. In many shallow water areas the sound-speed profile varies from upward to downward refracting with the change of seasons enhancing the role of one or the other boundary. Furthermore, incoherent and coherent signaling schemes may be sensitive to different features of the propagation physics. To carefully study these issues, a series of experiments has been planned in diverse sites. The pilot experiment, ModemEx99, was conducted last May near San Diego and has revealed numerous interesting features. In particular, we see that there are strong transitions in the multipath structure with pronounced focal regions and shadow zones. These features correlate closely with areas of good and poor modem performance, respectively. This talk will interpret these results in light of the measured channel impulse response and predictions from broadband channel simulators.

10:30

3aSP10. Coherent underwater digital communication during LWAD 99-1 experiment: Performance and analysis. Azmi Al-Kurd (Naval Res. Lab., 4555 Overlook Ave. SW, Washington, DC 20375-5320, a.alkurd@ieee.org)

Underwater acoustic communications data using phase-modulated signals were collected during the Littoral Warfare Advanced Development (LWAD 99-1) experiment. The signals were projected from a drifting and a towed source. Data of several band rates were transmitted and used to study the temporal and spatial variation of the acoustic impulse response of the ocean and to evaluate the performance of a phase coherent digital communication algorithm. Postexperimental analysis showed that the experiment site is a harsh environment for coherent digital communication. The oceanic channel impulse response is very complex and represents a dynamic and extended multipath structure (changes in seconds). Also, the first arrival is weaker than the later arrivals (most of the time). In addition to the environmental limitations the receiver has to overcome system limitations and operate in a low signal-to-noise ratio conditions. Spatial and temporal diversity are implemented to improve the algorithm performance. The objective is to determine the characteristics of signal propa-

gation in littoral environment, and to determine whether and how these characteristics affect the bit-error-rate. The results from this experiment and future experiments will serve as guidance in the design of a reliable coherent digital communication system.

10:45

3aSP11. Quantifying data information content in matched-field inversion. Stan E. Dosso, Edward Chapin, and Michael J. Wimut (School of Earth and Ocean Sci., Univ. of Victoria, Victoria, BC V8W 3P6, Canada, sdosso@uvic.ca)

In matched-field inversion, geoacoustic properties and/or source position are determined by minimizing an objective function that represents the mismatch between measured and modeled acoustic fields. As this is a strongly nonlinear problem, linearized inversion and appraisal algorithms do not apply. A problem of practical importance is to quantify how the inversion results are influenced by data factors such as (i) signal-to-noise level, (ii) number of frequencies included in the objective function, (iii) whether multiple frequencies are summed coherently or incoherently, and (iv) sensor array configuration. This problem is examined here by applying a fully nonlinear analysis based on Gibbs sampling to construct *posteriori* probability distributions (PPDs) for the unknown model parameters. Constructing PPDs for different realizations of the above factors allows their effect on the solution of the matched-field inverse problem to be quantified and exploited.

11:00

3aSP12. Extraction of modal amplitudes for the analysis of two- and three-dimensional underwater sound propagation. Joseph M. Riley and Stewart A. L. Glegg (Florida Atlantic Univ., Ctr. for Acoust. and Vib., Dept. of Ocean Eng., Bldg. #36, Rm. 190, 777 W. Glades Rd., Boca Raton, FL 33431)

Experimental measurements have been conducted over a laboratory scale model of the Santa Lucia Escarpment for the purpose of analyzing two- and three-dimensional underwater sound propagation. Transmission loss (TL) was measured as a function of depth at a series of ranges in the downslope direction of the model (referred to as depth profile measurements) and TL was also measured over a finite area in the horizontal plane for fixed source and receiver depths (referred to as two-dimensional surveys). The two-dimensional surveys show across slope interference, which may be indicative of bathymetric refraction, and the modal structure of the depth profile measurements appears to change as a function of downslope range (a result which may be caused by mode coupling). In order to investigate the interference observed in the two-dimensional surveys and also determine if the range-dependent variation in the depth profile measurements is being caused by mode coupling, an inversion algorithm was applied to measurements of the complex pressure field for the purpose of extracting the modal amplitudes. Analysis of the experimentally extracted modal amplitudes indicates that mode coupling and bathymetric refraction contributed significantly to the acoustic pressure field. [Work supported by ONR.]

THURSDAY AFTERNOON, 1 JUNE 2000

TOWER 1206, 1:00 TO 3:00 P.M.

Session 3pAA

Architectural Acoustics: Classroom Acoustics (Poster Session)

Brandon D. Tinianow, Chair

Acoustical Laboratory, Johns Manville, 10100 West Ute Avenue, Littleton, Colorado 80162

Contributed Papers

All posters will be on display from 1:00 p.m. to 3:00 p.m. All contributors will be at their posters from 1:00 p.m. to 3:00 p.m.

3pAA1. Eliminating acoustical barriers to learning in classrooms—Case study of window ventilator noise. Bennett M. Brooks (Brooks Acoust. Corp., 27 Hartford Turnpike, Vernon, CT 06066, bbrooks@brooks-acoustics.com)

Room ventilation systems have long been a major cause of noise in classrooms. The recent drive for energy efficiency has motivated schools to partner with utility companies to replace aging central HVAC systems with individual room heat pump window ventilator units for space heating and cooling. An unfortunate consequence is that these window ventilators are significant noise sources. A typical window unit can produce 70 dB(A), or more, at 1 m. Clearly, this is unacceptable. Either ventilator manufacturers must commit to reduce unit noise output by at least 30 dB, or school designers must abandon the wall ventilator option in favor of quiet central HVAC installations.

3pAA2. Improving existing classroom sound isolation for advance media capabilities. Dana Hougland (A CODA Acoustic, LLC 9603 E. Orchard Dr., Englewood, CO 80111, acoda@aol.com)

A series of tests was conducted to assess the incremental improvement of the noise reduction and sound transmission loss between adjacent classrooms with various improvements to the demising construction. Classrooms were originally constructed without full height walls. Tests were conducted before modifications were implemented and after each modification was completed. The investigation was conducted as part of a larger classroom improvement project design to bring advanced media capabilities into 60 college classrooms campus wide. The results of the testing program are presented.

gation in littoral environment, and to determine whether and how these characteristics affect the bit-error-rate. The results from this experiment and future experiments will serve as guidance in the design of a reliable coherent digital communication system.

10:45

3aSP11. Quantifying data information content in matched-field inversion. Stan E. Dosso, Edward Chapin, and Michael J. Wimut (School of Earth and Ocean Sci., Univ. of Victoria, Victoria, BC V8W 3P6, Canada, sdosso@uvic.ca)

In matched-field inversion, geoacoustic properties and/or source position are determined by minimizing an objective function that represents the mismatch between measured and modeled acoustic fields. As this is a strongly nonlinear problem, linearized inversion and appraisal algorithms do not apply. A problem of practical importance is to quantify how the inversion results are influenced by data factors such as (i) signal-to-noise level, (ii) number of frequencies included in the objective function, (iii) whether multiple frequencies are summed coherently or incoherently, and (iv) sensor array configuration. This problem is examined here by applying a fully nonlinear analysis based on Gibbs sampling to construct *posteriori* probability distributions (PPDs) for the unknown model parameters. Constructing PPDs for different realizations of the above factors allows their effect on the solution of the matched-field inverse problem to be quantified and exploited.

11:00

3aSP12. Extraction of modal amplitudes for the analysis of two- and three-dimensional underwater sound propagation. Joseph M. Riley and Stewart A. L. Glegg (Florida Atlantic Univ., Ctr. for Acoust. and Vib., Dept. of Ocean Eng., Bldg. #36, Rm. 190, 777 W. Glades Rd., Boca Raton, FL 33431)

Experimental measurements have been conducted over a laboratory scale model of the Santa Lucia Escarpment for the purpose of analyzing two- and three-dimensional underwater sound propagation. Transmission loss (TL) was measured as a function of depth at a series of ranges in the downslope direction of the model (referred to as depth profile measurements) and TL was also measured over a finite area in the horizontal plane for fixed source and receiver depths (referred to as two-dimensional surveys). The two-dimensional surveys show across slope interference, which may be indicative of bathymetric refraction, and the modal structure of the depth profile measurements appears to change as a function of downslope range (a result which may be caused by mode coupling). In order to investigate the interference observed in the two-dimensional surveys and also determine if the range-dependent variation in the depth profile measurements is being caused by mode coupling, an inversion algorithm was applied to measurements of the complex pressure field for the purpose of extracting the modal amplitudes. Analysis of the experimentally extracted modal amplitudes indicates that mode coupling and bathymetric refraction contributed significantly to the acoustic pressure field. [Work supported by ONR.]

THURSDAY AFTERNOON, 1 JUNE 2000

TOWER 1206, 1:00 TO 3:00 P.M.

Session 3pAA

Architectural Acoustics: Classroom Acoustics (Poster Session)

Brandon D. Tinianow, Chair

Acoustical Laboratory, Johns Manville, 10100 West Ute Avenue, Littleton, Colorado 80162

Contributed Papers

All posters will be on display from 1:00 p.m. to 3:00 p.m. All contributors will be at their posters from 1:00 p.m. to 3:00 p.m.

3pAA1. Eliminating acoustical barriers to learning in classrooms—Case study of window ventilator noise. Bennett M. Brooks (Brooks Acoust. Corp., 27 Hartford Turnpike, Vernon, CT 06066, bbrooks@brooks-acoustics.com)

Room ventilation systems have long been a major cause of noise in classrooms. The recent drive for energy efficiency has motivated schools to partner with utility companies to replace aging central HVAC systems with individual room heat pump window ventilator units for space heating and cooling. An unfortunate consequence is that these window ventilators are significant noise sources. A typical window unit can produce 70 dB(A), or more, at 1 m. Clearly, this is unacceptable. Either ventilator manufacturers must commit to reduce unit noise output by at least 30 dB, or school designers must abandon the wall ventilator option in favor of quiet central HVAC installations.

3pAA2. Improving existing classroom sound isolation for advance media capabilities. Dana Hougland (A CODA Acoustic, LLC 9603 E. Orchard Dr., Englewood, CO 80111, acoda@aol.com)

A series of tests was conducted to assess the incremental improvement of the noise reduction and sound transmission loss between adjacent classrooms with various improvements to the demising construction. Classrooms were originally constructed without full height walls. Tests were conducted before modifications were implemented and after each modification was completed. The investigation was conducted as part of a larger classroom improvement project design to bring advanced media capabilities into 60 college classrooms campus wide. The results of the testing program are presented.

3pAA3. Eliminating acoustical barriers to learning in classrooms—Case study of reverberation reduction in elementary school gymnasiums. Bennett M. Brooks (Brooks Acoust. Corp., 27 Hartford Turnpike, Vernon, CT 06066, bbrooks@brooks-acoustics.com)

A gymnasium can be a highly reverberant space. This is a consequence of the large room volume coupled with an abundance of hard surfaces. School designers frequently overlook the need for reverberation control in gymnasiums, cafeterias, and other large school rooms. The high level of reverberation promotes a high noise level and interferes with speech intelligibility, degrading the primary functions of those spaces. Moreover, those rooms are often used for additional functions, such as student assemblies, community meetings, and school performances, which will also suffer from excess reverberation. A case study is presented for two elementary school gymnasiums, each with mid-frequency reverberation times of about 5 s. Renovation treatments to control reverberation were developed using simple computer models. Significant reductions in reverberation and noise level were achieved. As a result, student manageability was improved, and teacher sanity was restored.

3pAA4. Case studies of “cafetorium” spaces in schools. Gary W. Siebein and Martin A. Gold (Dept. of Architecture, 231 ARCH, Univ. of Florida, Gainesville, FL 32611, mgold@ufl.edu)

Case studies of “cafetorium” spaces in elementary and high schools were conducted to evaluate the performance of this hybrid room approach. Many schools, in an effort to reduce costs, attempt to create hybrid spaces that will accommodate multiple functions. One result of this effort is the cafetorium. This study includes issues raised in establishing design criteria with architects and clients; acoustic space planning, organization, and other preliminary design issues; budget, design criteria, and material selection issues; acoustic models and prediction tools; background noise and noise control issues; sound system integration; constructed spaces and follow-up acoustical measurements; and client and user comments regarding the performance of the spaces. Each specific issue is presented within the context of built projects ranging in scope from new elementary school projects to renovations of high school projects. The case studies present issues regarding the relationship between established acoustical design criteria, prediction tools, architectural integrity, and budgetary constraints as determiners of successful spaces for listening and learning.

3pAA5. Optimal classroom design applied in the OSU Fisher School of Business Graduate Studies Gerlach Hall. Angelo Campanella (Campanella Assoc., 3201 Ridgewood Dr., Columbus, OH 43026)

On endowing the recent expansion of the Ohio State University College of Business, Max M. Fisher said that his interest was not in buildings; they did not excite him. Rather, his interest was in value-added programs to improve the product, education in this case. We followed this philosophy in the architectural acoustics design. The graduate studies program included many classrooms of 40 to 100 seat capacity, as well as stock market and communications laboratories. The classrooms were to include distance learning that required teleconferencing capability. To these ends, all classrooms were designed for NC-25 background noise and $[1 + \log(V)]/10$ reverberation time. Measurements of the background noise of three classrooms and the reverberation time of one classroom are presented. In October 1998 associate professor David Greenberger remarked “The acoustics are awesome. Anyone can hear you if you talk in a soft voice. That’s great for faculty.”

3pAA6. Case studies of music facilities in schools. Gary W. Siebein and Martin A. Gold (Dept. of Architecture, 231 ARCH, Univ. of Florida, Gainesville, FL 32611, mgold@ufl.edu)

Case studies of new music facilities in elementary through junior college school levels are presented with a focus on the design process. The study includes issues raised in establishing design criteria with architects and clients; acoustic space planning, organization, and other preliminary design issues; budget, design criteria, and material selection issues; acoustic models and prediction tools; background noise and noise control issues; sound system integration; constructed spaces and follow-up acoustical measurements; and client and user comments regarding the performance of the spaces. Each specific issue is presented within the context of built projects ranging in scope from elementary school practice rooms to a high school auditorium for a magnet performing arts school. The case studies present issues regarding the relationship between established acoustical design criteria, prediction tools, architectural integrity, and budgetary constraints as determiners of successful spaces for listening.

THURSDAY AFTERNOON, 1 JUNE 2000

ENGLISH ROOM, 12:55 TO 3:00 P.M.

Session 3pAO

Acoustical Oceanography: Topics in Acoustical Oceanography

David M. Farmer, Chair

Institute of Ocean Sciences, 9860 West Saanich Road, Sidney, British Columbia V8L 4B2, Canada

Chair’s Introduction—12:55

Contributed Papers

1:00

3pAO1. Bioacoustic absorptivity and transmission loss in littoral seas. Orest Diachok (Naval Res. Lab., Washington, DC 20375, orest@wave.nrl.navy.mil)

Long-term, broadband measurements of transmission loss, TL, between a fixed source and a fixed receiver in littoral seas, permit isolation of the effects of bioacoustic absorptivity, α_b , from other environmental effects on TL. Concurrent echo sounder measurements and biological sampling facilitate classification of the causes of observed absorption

lines. Frequencies of maximum α_b at night, when pelagic fish are predominantly dispersed near the surface, are consistent with a modified form of Minnaert’s equation for the resonance frequency of bubbles, f_0 , where the inputs are the effective radius and eccentricity of swim bladders. Frequencies of maximum α_b during the day, when pelagic fish are predominantly in schools near the bottom, are downshifted, relative to Minnaert’s equation, due to the collective oscillations of schools (Diachok, 1999). The magnitude of α_b at f_0 (at night) is a function of the number density, n (number/m³), and other bioacoustic parameters. Nearly coincident absorption and echo sounder based estimates of n have been shown to be consistent (Diachok, Liorzou, and Scalabrin, 2000). The implications of

these results for the design of the next generation of long-term, broadband transmission loss experiments in littoral seas where n is high, such as the seas off California and China, will be reviewed.

1:15

3pAO2. Mapping the regional variability of the California Current acoustically using a waveform inversion method. Sang-Kyu Han, Curtis A. Collins, Christopher W. Miller, Ching-Sang Chiu (Naval Postgrad. School, 589 Dyer Rd., Rm. 200A, Monterey, CA 93943-5143, skhan@nps.navy.mil), and Peter Worcester (Scripps Inst. of Oceanogr., La Jolla, CA 92093-0225)

To measure the large-scale variability in the California Current System, tomography signals are transmitted once every four days by a HLF-5 sound source moored on top of Hoke Seamount, approximately 600 km offshore from California. After propagating through the California Current Front and then upslope through inshore waters, the acoustic signals are recorded by a decommissioned SOSUS array off Point Sur. The processed time series shows four stable groups of arrivals that are consistently present in all of the transmissions. The fine structure within each group, however, varies from time to time. Using a ray-based range-dependent propagation model, with input sound speed field and bathymetry derived from hydrographic and echo sounding measurements, the arrival structure was simulated. Model results indicate that each of the four groups was composed of overlapping individual ray arrivals and that the variation of the fine structure within each group was due to phase interference. A waveform inversion scheme accounting for the multipath interference within each ray group is formulated. Using least-squares techniques, the scheme estimates the coefficients of one- or two-dimensional empirical orthogonal functions analyzed from the CalCOFI data. The inverse technique and results are presented and discussed. [Research supported by NOPP.]

1:30

3pAO3. Breaking surf noise signatures: A comparison of plunging and spilling breakers. Steven L. Means and Richard M. Heitmeyer (Naval Res. Lab., Code 7120, 4555 Overlook Ave. SW, Washington, DC 20375, means@wave.nrl.navy.mil)

Range-time-frequency distributions of surf-generated noise were measured within the surf zone during the SandyDuck'97 experiment at Duck, NC. A 24-phone, 138-m, bottom-mounted, linear array located along a line perpendicular to the shore at a depth of 1 to 3 m recorded the surf-generated noise. Concurrent video measurements of the location, size, and time evolution of the individual breaking waves directly above the array were made from a nearby 43-m tower. The present paper presents time-averaged source level spectra per unit wavefront length and range-time signatures for both plunging and spilling breakers. The source level spectra are obtained by using a modified fast-field program to account for water column and geoacoustic propagation from the distributed source region to an individual hydrophone. The length and location of the leading edge of breakers are tracked in time from rectified video images. It is observed that the source levels from large spilling breakers are lower ($\sim 5-6$ dB) than those produced by large plunging breakers that occurred during the same general time period. Distinguishing characteristics between the range-time signatures of spilling and plunging breakers are also observed. [Work supported by ONR.]

1:45

3pAO4. Acoustic propagation measurements in the surf. Kerry W. Commander and John S. Stroud (Coastal Systems Station, Code R21, 6703 W. Hwy. 98, Panama City, FL 32407-7001, stroudjs@ncsc.navy.mil)

As part of the 1999 Near Shore Acoustic Network Experiment (NSANE), conducted off of the Scripps Institute of Oceanography pier in La Jolla, CA, measurements were conducted in a very shallow water (VSW)/surf zone (SZ) region to investigate acoustic propagation and coherence in coastal areas. These measurements utilized an acoustical array of 16 hydrophones (supplied by Bethnos, formerly Datasonics, Inc.) deployed over many tens of meters, and a broadband source. The data to be presented were acquired from hydrophones on two vertical arrays, separated by 240 m, which comprised half of the total array. The broadband projector was operated in the 3.5- to 28-kHz range using a variety of pulse types. The projector was aimed seaward, as it was located closer to shore than any of the receivers. Data were acquired in conjunction with environmental data by other NSANE team members. The data were analyzed using coherence techniques. Results of this acoustic measurement for select pulse types will be reported. [Work supported by ONR Code 321OE.]

2:00

3pAO5. Modal adiabaticity in shallow water. Part I: Volume interaction. E. C. Shang and Y. Y. Wang (CIRES, Univ. of Colorado/NOAA/Environ. Technol. Lab., Boulder, CO 80303)

To assess the modal adiabaticity of sound propagation in the ocean is crucial both for acoustic field prediction (forward problem) and tomographic retrieving (inverse problem). Most of the criteria in the literature is too restrictive (conservative). A new criterion of the modal adiabaticity for volume interaction is proposed in this paper. The adiabaticity of two strong volume interaction cases in shallow water has been estimated based on the new criterion: (1) the Polar front case, and (2) the solitary wave case. The results assessed by the criterion have been verified numerically by processing the modal decomposition of the PE field. It is found that the adiabatic modes can survive in a narrow belt along the "cutoff" line in the " $m-f$ " (mode number-frequency) plan for these two strong volume interaction cases. [Work supported by ONR and NOAA.]

2:15

3pAO6. Very-high-frequency normal mode modeling. Michael F. Werby (NRL, Code 7181, Stennis Space Center, MS 39571)

The relevant quantity that determines the complexity of modeling via the normal mode method is the number of modes required in the problem. That is only partly determined by the frequency. Other factors that enter the problem are the water column depth and the mean sound speed in the water column and the bottom. Crudely the number of modes is proportional to the product of the water column depth and the frequency times what we call the hardness parameter and as well as the inverse of the mean speed of sound in the water column. The hardness parameter is equal to the square root of the factor of 1 minus the square of the ratio of the mean sound speed in the water column to that in the bottom. We determine the actual equation for the number of modes in a refractive wave guide with a

layered bottom. We illustrate the modal spectrum for the general problem. Based on that we show some tricks useful to model high-frequency problems on the order of several thousands of modes.

2:30

3pAO7. Determining velocity profile characteristics from pulse signals in a wave-guide. Michael F. Werby (NRL, Code 7181, Stennis Space Center, MS 39571, werby@nrlssc.navy.mil) and Herbert Uberall (Dept. of Phys., Catholic Univ. of America, Washington, DC)

We use techniques similar to that of vertical seismic profiling to determine the velocity profile of a waveguide by examining contour plots from the emergence of pulse signals. The concept of acoustic adumbrations to determine features of the wave-guide channeling effect is discussed. It may be seen that the contour plots yield and outline adumbration that “shadows” the features of the profile.

2:45

3pAO8. A method for fast inversion of ocean waveguide parameters. Dejun Jiang (Inst. of Acoust., The Chinese Acad. of Sci., Beijing 100080, PROC, jdj@ocean.ia.ac.cn)

A method for fast inversion of waveguide parameters using a broadband pulse signal is presented. The input parameters are relative amplitudes and relative delays of normal modes which can be determined with TDE technique. The cost function consists of the sum of two parts; one is the relative amplitudes of normal modes which can be used to calculate the velocity and density of waveguide, the other is the relative delays of normal modes which can be used to calculate the attenuation coefficient. A genetic algorithm (GA) is used to estimate waveguide parameters. The result shows that the method not only increased the inversion speed significantly, but is also robust to ocean environment.

THURSDAY AFTERNOON, 1 JUNE 2000

PEACHTREE BALLROOM, 1:25 TO 2:30 P.M.

Session 3pID

Interdisciplinary: Hot Topics in Acoustics

Joseph Pope, Chair

Pope Engineering Company, P.O. Box 236, Newton, Massachusetts 02549

Chair's Introduction—1:25

Invited Papers

1:30

3pID1. Hot topics in engineering acoustics. Thomas R. Howarth (Naval Res. Lab., Washington, DC 20375-5350, howarth@nrl.navy.mil)

Underwater acoustic transduction has been undergoing great changes in modern device designs. After several decades of advancements in acoustic sensing technologies, current interests have included novel and application-specific designs in underwater acoustic projection. These changes are an outgrowth of advances in materials processing and manufacturing. The incorporation of newer materials is resulting in devices that are specifically tailored for individual applications as opposed to general designs that in the past have been used to cover a variety of applications. This presentation will introduce the concepts of the new materials advances, including piezocomposites, piezopolymers, electrostrictive ceramics, single crystals, hybrid piezoelectric ceramic/magnetostriction devices, and microflexensional drivers called cymbals. The state of development of these technologies will be introduced as they are being incorporated into new devices. The talk will show prototype designs and discuss future directions. [The continuing support of ONR 321 is gratefully acknowledged.]

1:50

3pID2. Hot topics in underwater acoustics. George V. Frisk (Dept. of Appl. Ocean Phys. and Eng., Woods Hole Oceanogr. Inst., Woods Hole, MA 02543)

The field of underwater acoustics has reached a level of maturity that enables it to lay claim to offspring such as acoustical oceanography and animal bioacoustics. This maturity should by no means be equated with stagnation, however, for the classical discipline of ocean acoustics remains a lively, active, and vibrant area of research. It continues to generate exciting results through its unique combination of wave propagation physical, signal processing, and at-sea experimentation. In recent years, the field has benefited enormously from scientific advances in oceanography and geophysics, as well as technological developments such as the Global Positioning System and microchip hardware. Current research activities in underwater acoustics embrace a wide band of frequencies, from tens of hertz to hundreds of kilohertz, and a broad range of oceanographic environments, from the surf zone to the deep ocean. Some of these exciting developments are reviewed in this paper.

3p THU. PM

3pID3. Hot topics in psychological and physiological acoustics: Compression. Sid P. Bacon (Psychoacoust. Lab., Dept. of Speech and Hearing Sci., Arizona State Univ., Tempe, AZ 85287-1908, spb@asu.edu)

The auditory system has an extraordinarily large dynamic range, evidenced, for example, by the fact that humans can detect an approximately 1-dB increment in the level of sounds that vary over a range of 120 dB. This enormous dynamic range is a consequence of the compression that is observed in the vibratory response of the basilar membrane in the cochlea. Damage to the cochlea, which occurs in many types of hearing loss, usually results in a reduction or loss of compression and, importantly, a reduction in the dynamic range of hearing. Physiological evidence for compression will be presented, along with results from recent psychophysical experiments that provide quantitative estimates of the degree of compression in humans. In addition, the effects of temporary and permanent hearing loss will be described, as will some of the signal processing strategies that are employed by many hearing aids in an effort to compensate for reduced compression.

THURSDAY AFTERNOON, 1 JUNE 2000

TOWER 1205, 1:00 TO 2:50 P.M.

Session 3pMU

Musical Acoustics: Choral and Solo Vocal Performance

Peter L. Hoekje, Chair

Department of Physics and Astronomy, Baldwin-Wallace College, 275 Eastland Road, Berea, Ohio 44017

Chair's Introduction—1:00

Invited Papers

1:05

3pMU1. The closed quotient as a factor in the singing voice. Donald G. Miller and Harm K. Schutte (Dept. of Biomed. Eng., Groningen Univ., Bloemsingel 10, 9712KZ Groningen, The Netherlands, d.g.miller@med.rug.nl)

The closed quotient (CQ), or percentage of the glottal cycle in which airflow is prevented by the contiguous vocal folds, has been a relatively neglected entity in recent literature on the singing voice. This descriptive paper considers the estimation of the closed quotient (CQ) through electroglottography, the audio signal, and videokymography; the factor of CQ in register, type of singing voice, and characteristic resonance strategies in singing; and phase relationships of glottal closing and opening with subglottal and supra-glottal pressure waves, and especially with the driving of the dominant standing wave of the vocal tract. The degree of acoustic coupling between the spaces above and below the glottis is examined on the basis of broadband signals from pressure transducers mounted on a catheter passed through the posterior commissure of the glottis.

1:35

3pMU2. Vocal warm-ups: What do they accomplish? Ingo R. Titze (Dept. of Speech Pathol. and Audiol., Natl. Ctr. for Voice and Speech, 330-SHC, Univ. of Iowa, Iowa City, IA 52242, ingo-titze@uiowa.edu)

In the past, vocal warm-ups have been little more than a few scales and arpeggios on a few selected vowels, whether for choir or for solo singing. In recent years, however, a greater understanding of the mechanism of the human voice, combined with a greater understanding of exercise physiology, has led to entire systems of carefully designed sequences of exercises. Specific muscles are targeted, even specific portions of the vocal fold. Of principal interest is the way in which the vocal tract is engaged, and ultimately disengaged, in providing acoustic loads for the vocal oscillator. A system of warm-up exercises will be presented here that have become the backbone of vocology, the science and practice of voice habilitation.

3pMU3. Characteristics of choral and solo singing, illustrated using singing-voice synthesis. Sten Ternstrom (Speech Music & Hearing, Royal Inst. of Technol., DKV 31, SE-100 44 Stockholm, Sweden)

Solo and choral voices typically differ acoustically in several ways. Some of these differences are motivated by the differences in the vocal tasks, i.e., “be heard” as opposed to “blend,” while some are due more to differences in vocal training. Solo and choir singers vary in their control strategies for sound level, fundamental frequency (pitch), formant frequencies (timbre), and voice-source properties (phonatory behavior). These differences are discussed and illustrated using recent real-time voice-synthesis techniques. The topics of the acoustical nature of the “chorus effect” and of child singing voice are also visited. [Work supported by the Swedish Natural Science Research Council, the Swedish Research Council for Engineering Sciences, and the Bank of Sweden Tercentenary Foundation.]

Contributed Paper

2:35

3pMU4. The acoustic characteristics of Korean traditional singing. Seung-Jae Moon (English Dept., Ajou Univ., Suwon, Korea 442-749, sjmoon@madang.ajou.ac.kr)

Korean traditional singing voice, called “Pansori,” is known for its peculiar sound quality compared to Western singing voice. It sounds very strained, yet hoarse. This study aims at investigating the acoustic characteristics of this singing voice. Singing voices of both male and female professional Pansori singers were recorded and analyzed. Also, regular conversation of one of the singers was recorded and analyzed to compare the singing vs conversational voice. From long-term average spectra of

suspended vowels and from harmonic structure analysis, several things were noted. Pansori singing voice showed more gradual downslopes in spectrum envelope than normal voice. And conventional methods of measuring formant frequencies were of no use for this particular singing voice. Instead of showing regular vowel formant patterns, Pansori seemed to be characterized with irregular formant patterns such as a single prominent peak throughout the whole frequency range, especially in a strained voice. Conversation data revealed similar characteristics of the singing voice, indicating that these characteristics were the results of the permanent change of the singer’s vocal source, not the results of the mode switch. [Work supported by a Research Fund provided by the Korea Research Foundation, Support for Faculty Research Abroad.]

THURSDAY AFTERNOON, 1 JUNE 2000

SPANISH ROOM, 1:00 TO 2:45 P.M.

Session 3pPA

Physical Acoustics: Acoustics of Multiphase Flow II

Isaac Shreiber, Chair

Department of Chemical Engineering, Ben-Gurion University of the Negev, Beer Sheva 84105, Israel

Contributed Papers

1:00

3pPA1. Characterization of acousto-electric drop arrays with application to evaporation studies. Yibing Zheng and Robert E. Apfel (Dept. of Mech. Eng., Yale Univ., New Haven, CT 06520, yibing.zheng@yale.edu)

An acousto-electric levitator has been developed to study liquid drop and solid particle clusters and arrays. This device uses both an acoustic standing wave and a high dc electric field simultaneously to form and levitate several charged drops in a two-dimensional array in air. A 2-D theoretical model is employed to study the profile of the acoustic field and the forces that hold the arrays at the center of the levitator. A study of the evaporation of alkane drops and arrays has been performed. It is shown that when a minimal acoustic field is used to levitate drops (by compensating with an electric field), the evaporation rate in the acousto-electric levitator is almost the same as that of drops suspended by a glass fiber. The experimental results for an isolated drop are compared with the theory of a quasi-steady model. The point source method is used to calculate theoretically the results for array evaporation. Both single and multiple component drops are studied. [Work supported by NASA (NAG3-2147) and by NSF (CTS-9870015).]

1:15

3pPA2. High-resolution longitudinal and shear wave spectroscopy for analysis of complex colloids. Cormac Smyth, Evgeny Kudryashov, and Vitaly Buckin (Dept. of Chemistry, Univ. College Dublin, Belfield, Dublin 4, Ireland)

In the present paper we describe the applications of the high-resolution ultrasonic longitudinal and shear wave resonator techniques for food and bio-colloids. Our longitudinal wave technique allows us to measure the velocity and attenuation of ultrasonic waves with a resolution of 0.0001% for ultrasonic velocity and 0.1% for the ultrasonic attenuation in small volumes (0.1 to 1 ml) in the frequency range 4–20 MHz. The shear wave technique provides information on viscoelastic parameters of liquids and gels in the frequency range 4–25 MHz. Both techniques allow continuous, temperature ramp and automatic titration measurements for analysis of various processes in complex colloids. Examples of the application of both techniques for studying heat-induced milk coagulation, phase transitions in milk fats, acid milk gelation, and the detection of food pathogens are described. Combination of both techniques shows a high sensitivity to various processes, e.g., pregelation and gelation, in food colloids and provides the unique ability to evaluate simultaneously the changes in volume

and shear viscoelastic moduli. Overall our ultrasonic measurements demonstrate their high potential for nondestructive analysis of complex food and bio-colloids.

1:30

3pPA3. Optical manipulation of bubbles in water using solid-state laser technology. Pey-Schuan Jian, William E. Torruellas, David B. Thiessen, and Philip L. Marston (Dept. of Phys., Washington State Univ., Pullman, WA 99164-2814)

For some experiments on the acoustics of bubbles in water, it may be desirable to counter buoyancy using optical radiation pressure. A previous demonstration of the optical trapping of small bubbles that would have otherwise been freely rising used a green continuous downward-directed beam from an Ar-ion gas laser having an irradiance minimum on the beam axis [B. T. Unger and P. L. Marston, *J. Acoust. Soc. Am.* **83**, 970–975 (1988)]. Some of the difficulties in using solid-state laser technology for this purpose are that the beam should be green (to minimize absorption) and have an axial irradiance minimum (for lateral stability). We have generated the required beam using a diode-pumped Nd:YVO₄ laser crystal with intracavity-second-harmonic-generation in a lithium-tri-borate crystal. The 5.5-W external green beam power levels achievable were more than sufficient for suspension of gas bubbles having 20- and 40- μm radii generated by electrolysis. The bubble size is measured using forward scattering. Occasionally, two vertically offset bubbles can be trapped along the axis of the downward directed beam. [Work supported by NASA.]

1:45

3pPA4. The effects of interfacial viscosity on the attenuation and phase speed of acoustic waves in bubbly liquids. Joseph C. Jankovsky, Ronald A. Roy, and William M. Carey (Dept. of Aersp. and Mech. Eng., Boston Univ., 110 Cummington St., Boston, MA 02215, jankov@bu.edu)

The attenuation and phase speed of bubbly fluids have been shown to be highly dispersive due in part to the resonance response of individual bubbles. Surface-active materials, or surfactants, are known to dramatically alter the viscoelastic dynamics of an interface without affecting the bulk fluid properties. In this theoretical model the effects of surface viscosity on the propagation of linear pressure waves through a bubbly fluid are examined. The attenuation and phase speed are calculated for varying radii, void fraction, and interfacial viscosity for air bubbles in water. The results show an increase in attenuation for frequencies far removed from the bubble resonance for radii less than 100 microns. There is no significant increase in attenuation for larger bubbles (order 1 mm). Interestingly, near resonance, the interfacial viscous case has a lower attenuation than the clean interface, most likely due to the reduced radiation damping associated with smaller radial excursions. Likewise, resonance effects on the phase speed dispersion curve are also significantly diminished due to the interfacial viscosity. These results could be important for understanding propagation near the surface of the ocean, where surface-active materials can dramatically affect the dynamics of micron-sized bubbles and bubble distributions. [Work supported by ONR.]

2:00

3pPA5. A reduced model of cavitation physics for use in sonochemistry. Brian D. Storey and Andrew J. Szeri (Dept. of Mech. Eng., Univ. of California, Berkeley, CA 94720-1740, bstorey@newton.berkeley.edu)

Sonochemistry involves focusing acoustic energy through tiny cavitation bubbles to increase chemical activity. The violent bubble collapses lead to temperatures of several thousand Kelvin which drive chemical

reactions. In previous work (B. D. Storey and A. J. Szeri, “Water vapor, sonoluminescence, and sonochemistry,” in review), a very detailed computational model of a single bubble collapse was presented, taking into account phase change, mass diffusion, heat diffusion, nonuniform pressure, and chemical reactions. All of these phenomena are important in determining the conditions at collapse. The present work involves development of a much simpler model that includes all the relevant physics. Comparisons to the more detailed computations are made and the current model is found to provide reasonable results. Further, it is shown that many of the observed trends in sonochemistry are reflected in the results with a single “representative” bubble. [This work was supported by NSF and LLNL.]

2:15

3pPA6. Video investigation of the effects of gravity on single-bubble sonoluminescence. Ben Dzikowicz, David B. Thiessen, and Philip L. Marston (Dept. of Phys., Washington State Univ., Pullman, WA 99163-2814, dzikowic@wsunix.wsu.edu)

Using NASA's KC-135 aircraft flying parabolic trajectories, we can observe SBSL as the effective gravity changes from near 0 g to about 1.8 g. A spherical resonator operated at 32.5 kHz is suspended in a constant ambient pressure chamber. Then, a photomultiplier measures light emission while a strobed CCD camera obtains images of the bubble collapse. In normal gravity, buoyancy displaces the bubble from the velocity node of the sound field so that water is flowing past the bubble at the time of collapse. Other gravitational effects include the variation in ambient pressure. During intervals of negligible drift of the SBSL intensity, there can be a rapid intensity rise of between 4% and 20% as the effective gravity decreases. The CCD images are then used to calculate the bubble radius. These measurements were fit to the Rayleigh–Plesset equation to infer the size and pressure conditions by assuming spherical bubble oscillations. These conditions were found to be similar to those reported by other groups for measurements in normal gravity. There are also indications that changes in the effective gravity are accompanied by small changes in the maximum bubble radius. [Work supported by NASA.]

2:30

3pPA7. Theory for sonoluminescence in cryogenic monatomic systems. Brian D. Storey, Howard Tseng, and Andrew J. Szeri (Dept. of Mech. Eng., Univ. of California, Berkeley, CA 94720-1740, bstorey@newton.berkeley.edu)

Single-bubble sonoluminescence involves the focusing of acoustic energy by a single bubble to produce light. Most successful sonoluminescence experiments are conducted in liquid water around standard atmospheric conditions. Recent numerical work has shown that water vapor trapped in the interior limits the peak temperatures produced in the bubble interior. The water vapor has a low specific heat ratio reducing the amount of compression heating, and endothermic chemical reactions absorb much of the focused acoustic energy. To overcome these problems, a system of gaseous helium dissolved in liquid argon is investigated. In this system the bubble interior is always a monatomic gas, therefore compression heating is maximized. Regions of stability (with respect to shape perturbations and rectified diffusion) are determined for single bubbles. It is of interest to determine whether the peak temperatures achieved can be significantly increased relative to experiments in water. [This work was supported by NSF and LLNL.]

Session 3pSP

Signal Processing in Acoustics and Underwater Acoustics: Algorithms and Performance II

James C. Preisig, Chair

Department of Applied Ocean Physics and Engineering, Woods Hole Oceanographic Institution,
Woods Hole, Massachusetts 02543-1053

Contributed Papers

1:00

3pSP1. Regularized matched-mode localization with environmental mismatch. Stan E. Dosso and Nicole E. Collison (School of Earth and Ocean Sci., Univ. of Victoria, Victoria, BC V8W 3P6, Canada, sdosso@uvic.ca)

This paper considers a new approach to matched-mode processing (MMP) for source localization. The MMP consists of decomposing far-field acoustic data to obtain the modal excitations, then matching these with modeled replica excitations. A potential advantage of MMP over matched-field processing (MFP) is that subsets of the complete mode set can be considered. For example, if geoacoustic properties are poorly known, the matching can be applied only to low-order modes that interact minimally with the seabed. However, modal decomposition can be ill posed and unstable if the sensor array does not adequately sample the acoustic field. For such cases, standard decomposition methods yield minimum-norm solutions that are biased towards zero. Although these methods provide mathematical solutions (stable solutions that fit the data), they may not represent physically meaningful solutions. The new approach of regularized MMP (RMMP) carries out an independent decomposition prior to comparison with the replica excitations for each grid point, using the replica itself as the prior estimate in a regularized inversion. This provides a more meaningful decomposition near the actual source location. In this paper, RMMP, MMP, and MFP are compared for realistic test cases, including various sensor array configurations, as well as environmental mismatch in seabed properties.

1:15

3pSP2. Matched-field source localization using data-derived modes. Paul Hursky, W. S. Hodgkiss, and W. A. Kuperman (Marine Physical Lab., Scripps Inst. of Oceanogr., UCSD, 9500 Gilman Dr., La Jolla, CA 92093-0238, hursky@mpl.ucsd.edu)

Acoustic modes can be derived from the cross-spectral density matrix of a vertical line array in an ocean wave guide, when spatially distributed ambient noise or a moving source causes the modes at the array to be uncorrelated. This typically produces modes that are noisy and incomplete (only some modes are recovered, at some frequencies, and only at depths spanned by the array). Furthermore, without wave numbers, only incoherent localization algorithms are possible, providing less gain than coherent MFP using wave numbers. It will be shown how to overcome these problems, assuming the sound speed profile has been measured in the water column but assuming no *a priori* knowledge of bottom properties. A data-dependent shooting method is presented which extrapolates the data-derived modes to the entire water column and estimates both the wave numbers and bottom properties. Having bottom parameters enables modes to be calculated at all frequencies, rather than being limited to frequencies of sources of opportunity. The MFP results will be presented from the SWellEX-96 experiment using data-derived modes and compared to MFP based on modes from a normal mode model.

1:30

3pSP3. Source ranging with minimal environmental information using the virtual receiver and waveguide invariant concepts. Aaron Thode (MIT, 77 Massachusetts Ave., Cambridge, MA 02139)

A method is presented for determining the range of an unknown broadband acoustic source in a waveguide, using a vertical array and a signal sample from another broadband source at a known location relative to the array. The method requires no knowledge or modeling of the ocean environment for range-independent environments. The strategy employed is to combine waveguide invariant theory [e.g., D'Spain and Kuperman, *J. Acoust. Soc. Am.* **106**, 2454–2468 (1999)] with “virtual receiver” concepts [Siderius *et al.*, *J. Acoust. Soc. Am.* **102**, 3439–3449 (1997)] to create a “virtual aperture.” In other words, a vertical array can effectively convert a source at known range R into a continuum of receivers lying between ranges $(1 \pm \alpha)R$, where $\alpha \approx 0.1$. When the intensity field across this virtual aperture is plotted in the range–frequency plane, the slopes of the intensity maxima yield the range to the unknown source from the guide source, and thus the range to the vertical array. The concept is illustrated in both simulation and data collected from a 75–150-Hz bandwidth pseudorandom noise source during the shallow-water experiment SWellEx-3. The virtual aperture can be reformulated for range-dependent environments, provided that adiabatic propagation can be assumed, and the bathymetry surrounding the array is known. [Work also conducted at Marine Physical Lab., Scripps Inst. of Oceanogr., La Jolla, CA 92093-0704.]

1:45

3pSP4. Localization method based on the higher-order statistics. Salah Bourennane (ENSPM Domaine, Universitaire de Saint-Jrme, 13397 Marseille, Cedex 20, France, salah.bourennane@lsi.u-3mrs.fr)

In the context of the narrow-band or wideband array processing problem, in this paper we develop a robust algorithm to improve the accuracy of the estimation of the direction of arrival of the sources. It is well known that when the noise covariance matrix is unknown, these estimates may be grossly inaccurate. Using both the higher-order statistics for suppression of the additive Gaussian noise, the transformation matrices for estimating the coherent signal subspace, and the high-resolution algorithm a robust and fast algorithm for the source parameters estimation in the presence of noise with an unknown cross-spectral matrix is developed. We shall show that the performance of bearing estimation algorithms improves substantially when our algorithm is used. Several simulation results are presented with different noise spectral matrices.

3pSP5. Multiple neural networks-integrated underwater target classification based on fuzzy theory. Tong Su (Marine Systems Eng. Res. Inst., China State Shipbuilding Corp., Beijing 100036, PROC)

In the area of underwater target recognition, in order to process various target feature data extracted from underwater radiating noise signals, a target classification method based on the integration of multiple fuzzy clustering neural networks and fuzzy comprehensive analysis is put forward in this paper. An adaptive fuzzy clustering neural network classifier is constructed and a relative studying algorithm is described for the first class classification, which is applied separately to different kinds of feature data from LOFAR spectrum, bispectrum, and wavelet transformation. Based on fuzzy comprehensive analysis of the outputs and probabilities of the above neural network classifiers and the category and distribution of test samples, the second class fuzzy classifying rules are designed to preprocess and assess the categories of ships. More than 4000 passive sonar signal samples are collected and many tests are done, and the results show the feasibility, adaptability, and effectiveness of the present method.

3pSP6. Application of matched filtering for environmental compensation. Charles F. Gaumond (Naval Res. Lab., Washington, DC 20375-5320)

The problem of improving the performance of broadband sonar through the use of modeling is posed. An approach to this problem is presented in terms of matched filtering. Specific filtering equations are derived from a coherent sonar equation as a signal model and a broadband sound propagation model. The approach assumes a statistical distribution of target positions with consequently stochastic response that is then shown to be describable with a signal space. The basis of the signal space is derivable from a broadband propagation model. The resultant signal space is used with a separable-kernel detector. The significant improvement of detection over conventional matched filtering is demonstrated for an impulsive signal in a simple, dispersive channel with additive, Gaussian, white noise. Furthermore, this signal-processing methodology has the potential for depth discrimination if the channel response is significantly different in different depth intervals. This difference is expressed as the overlap of signal spaces. The relationship of the overlap to depth discrimination is shown. [Work supported by the Office of Naval Research.]

3pSP7. Blind deconvolution for multipath mitigation in shallow water acoustics. George B. Smith (Naval Res. Lab., Code 7185, Stennis Space Center, MS 39529-5004, gsmith@nrlssc.navy.mil)

Extraction of a signal received on a hydrophone array at tactical ranges in shallow water is made difficult by distortions caused by dynamic multipath. Blind deconvolution techniques can adapt the processing to the environment if reasonable assumptions about the nature of the Green's functions used to represent the multipath propagation can be made. An interesting class of blind techniques is based on the fact that the desired signal is common to all channels, whereas the Green's functions usually vary across channels. These techniques attempt to extract the desired source signal as the common factor among all data channels. Two techniques from this class are presented here: Rietsch's algorithm, which is an exact algebraic solution to the blind deconvolution problem, and a method

which attempts to derive filters that produce signal estimates from each data channel that are as consistent with each other as possible in a least squares sense. Using normalized correlations of output with ground truth as a performance measure, the performance of these two techniques will be compared to each other and to unprocessed output. [Work supported by ONR. Technical management was provided by the Naval Research Laboratory, Stennis Space Center, Mississippi.]

3pSP8. A Bayesian approach to model-based source localization using arrival time information. Zoi-Heleni Michalopoulou and Xiaoqun Ma (Dept. of Mathematical Sci., New Jersey Inst. of Technol., Newark, NJ 07102, elmich@m.njit.edu)

In this work, the potential of a model-based time delay estimation approach for source localization in the ocean is explored. Source localization in an underwater waveguide is typically approached with full-field matching techniques such as matched-field processing. The proposed method is similar to matched-field processing in that it matches measured arrival times to replica arrival times; it is, however, more efficient because it does not rely on exhaustive replica computations. The method links identified arrival times to source location parameters through an approximate linear model. Using a Bayesian approach and through the maximization of the *a posteriori* probability density function of the source location parameters, the coordinates that satisfy best the linear model are identified yielding the source location estimates. The method is tested with synthetic data and is shown to be accurate and fast. In addition to source localization in a known medium, the method is extended to uncertain media cases; specifically, the case of unknown water column depth is investigated. The model-based time delay method is applied successfully in that case both for source localization and ocean depth estimation.

3pSP9. Localization analysis based on subband processing of broadband LFM signals. Georgios Haralabus (SACLANT Undersea Res. Ctr., Viale San Bartolomeo, 400 I-19138, La Spezia, Italy)

The increased utilization of broadband signals in underwater research and development is related to sonar performance enhancement in littoral waters, optimum operational frequency selection, and interplatform, multisensor experimental settings requiring frequency compatibility. Preliminary wide-band experiments indicate potential performance improvement in reverberation-limited environments. Here the frequency dependence of active broadband detection/localization is examined based on 1200-Hz LFM signals (2300–3500 Hz). A subband matched filter scheme is devised according to which a replica of the transmitted pulse is segmented into ten 120-Hz subbands. Although comparison of source localization indicates comparable performance of all subbands, ping-to-ping variability of the ten correlator outputs suggests that the detection performance maybe improved by employing incoherent processing schemes. Doppler effects and subband detection synchronization problems that lead to performance degradation in large time-bandwidth signal scenarios are addressed. A method to estimate range rate (relative velocity between source and receiver) based on single ping differential time delay between subband matched filter outputs is developed. This intraping technique is an alternative to the standard interping method which requires multiping detection history.

Session 4aAAa**Architectural Acoustics: Recent Authors of Books on Architectural Acoustics**

William J. Cavanaugh, Cochair

Cavanaugh Tocci Associates, Inc., 327F Boston Post Road, Sudbury, Massachusetts 01776

Ronald R. Freheit, Cochair

Wenger Corporation, 555 Park Drive, Owatonna, Minnesota 55060

Chair's Introduction—9:00

In connection with the Acoustical Society's first "Book Fair" ongoing at this Spring 2000 Atlanta Meeting, the Technical Committee on Architectural Acoustics invited authors and editors of recent publications of interest to share their thoughts, motives and objectives in producing their works and the extent to which they feel these objectives had been achieved. This session will provide an excellent opportunity to explore the "state-of-the-art" in architectural acoustics at the conclusion of this most productive 20th century and to think about where we might be headed in the century ahead. We hope for a lively discussion to prepare ourselves for Red Wetherill's lecture in the following session (4aAAb) on the remarkable history of the field. Authors and/or editors representing the books listed below have been invited to participate. Authors who are unable to attend the Atlanta meetings have been invited to submit short remarks which will be read at the panel discussion.

Apfel, R. E., *Deaf Architects and Blind Acousticians....A Guide to the Principles of Sound Design*, Apple Enterprises Press, New Haven, Connecticut (1998)

Barron, M., *Auditorium Acoustics and Architectural Design*, E&FN Spon, an imprint of Chapman and Hall, London and New York (1993)

Beranek, L. L., *Concert and Opera Halls....How They Sound*, Acoustical Society of America, Melville, New York (1996).

Cavanaugh, W. J. and Wilkes, J. A., Ed., *Architectural Acoustics....Principles and Practice*, Wiley, New York (1998)

Cowan, J. T., *Architectural Acoustics* (CD-ROM), McGraw-Hill, New York (1999)

Crocker, M. J., *Handbook of Acoustics*, Wiley, New York (1998)

Egan, M. D., *Architectural Acoustics*, McGraw-Hill, New York (1988).

Harris, C. M., *Noise Control in Buildings...A Practical Guide for Architects and Engineers*, Ins. Noise Control Engr., Poughkeepsie, New York, 1997 (orig. 1994, McGraw-Hill) and *Handbook of Acoustical Measurements and Noise Control*, Acoustical Society of America, Melville, New York (1998)

Irvine, L. K. and Richards, R. L., *Acoustics and Noise Control Handbook for Architects and Builders*, Krieger, Melbourne, Florida (1998)

Izenour, G. C., *Theater Design/Theater Technology/Roofed Theaters of Classical Antiquity*, Yale University Press, New Haven, Connecticut (1996)

Kopec, J. W., *The Sabines at Riverbank*, Acoustical Society of America (by Peninsular Press), Melville, New York (1997)

Knudsen, V. O. and Harris, C. M., *Acoustical Designing in Architecture*, Acoustical Society of America, Melville, New York (1980) (orig. 1950, Wiley)

Lubman, D. and Wetherill, E. A., Eds., *Acoustics of Worship Spaces*, Acoustical Society of America, Melville, New York (1985)

McCue, E. R. and Talaske, R. H., Eds., *Acoustical Design of Music Education Facilities*, Acoustical Society of America, Melville, New York (1990)

Mehta, M., Johnson, J. and Rocafort, J., *Architectural Acoustics...Principles and Design*, Prentice Hall, Upper Saddle River, New Jersey (1999)

Ollswang, J. and Ambrose, J., *Simplified Design for Building Sound Control*, Wiley, New York (1995)

Pelton, H., *Noise Control Management*, Van Nostrand Reinhold (1993)

Sabine, W. C., *Collected Papers on Acoustics*, Acoustical Society of America (by Peninsular Press), Melville, New York (1994) (orig. 1921, Harvard Univ. Press; 1964, Dover)

Salter, C. M., Ed., *Acoustics, Architecture, Engineering, the Environment*, William Stout Publishers, San Francisco, California (1998)

Sendra, J. J., Ed., *Computational Acoustics in Architecture*, WIT Press, Southampton, UK (1999)

Talaske, R. H. and Boner, R. E., Eds., *Theatres for Drama Performance: Recent Experience in Acoustical Design*, Acoustical Society of America, Melville, New York (1987)

Talaske, R. H., Wetherill, E. A., and Cavanaugh, W. J., *Halls for Music Performance: Two Decades of Experience—1962–1982*, Acoustical Society of America, Melville, New York (1982)

Uzzle, T., Bushnell, R. A., and Bouliane, T. G., *Technical Fundamentals of Audio*, Intertec Publishing, Overland, KS, 1999

FRIDAY MORNING, 2 JUNE 2000

TOWER 1206, 11:00 A.M. TO 12:05 P.M.

Session 4aAAb

Architectural Acoustics and Committee on Archives and History: History of Architectural Acoustics

William J. Cavanaugh, Chair

Cavanaugh Tocci Associates, Inc., 327F Boston Post Road, Sudbury, Massachusetts 01776

Chair's Introduction—11:00

Invited Paper

11:05

4aAAb1. The flowering of architectural acoustics in the twentieth century. Ewart A. Wetherill (Paoletti Assoc., 649 Mission St., San Francisco, CA 94105)

Arising from a single building deficiency, the art and science of architectural acoustics came into being as a practical discipline at the start of the twentieth century through the painstaking exploration and profound intuition of Wallace Clement Sabine and other pioneers in acoustics. It now faces the dawning of the twenty-first century as a universal scientific discipline and an integral component of advanced building design. For the past 70 years of its often-fragile coexistence with the building professions, of architecture, much of the growth of understanding and education in architectural acoustics has found a focus in the Acoustical Society of America in collaboration with its companion societies throughout the world. While the understanding of building design for good hearing is still uneven and far from complete, the story of this development, as seen in the history of the Acoustical Society, indicates rich promise for the years ahead.

A NOTE ABOUT THE ASA HISTORY LECTURE SERIES

In 1997, the ASA Committee on Archives and History conceived a plan for a series of invited lectures on each of the technical areas of the Society which would memorialize the significant achievements and milestones of each of its twelve technical committees and one interdisciplinary technical group during the first three quarters of the Society's first century.

With the cooperation of the technical committees, distinguished individuals are selected to review the history of their particular technical specialty and present a lecture which shows how that activity has developed and has contributed to the Society at large and to the broad field of acoustics as well. At the 138th meeting in Columbus, Ohio, the first two History Lectures were presented: Gabriel Weinrich on Musical Acoustics and Robert Beyer and David Blackstock on Physical Acoustics. At each subsequent meeting two additional lectures will be scheduled including those in Architectural Acoustics and Engineering Acoustics at this meeting in Atlanta.

The invited lecturers have been asked to prepare a written manuscript of their lectures which will be published in a commemorative book for the 75th Anniversary of the Society to be celebrated in 2004. The Archives and History Committee and the individual technical committees/group welcome comments and suggestions on both the History Lecture Series and on the proposed ASA Diamond Anniversary Book. Volunteers to assist the committees would be most welcome too. Contact Henry Bass, Chair, Committee on Archives and History, pabass@sunset.backbone.olemiss.edu

4a FRI. AM

Session 4aEA

Engineering Acoustics: Acoustic Barriers and Absorbers

Kirk E. Jenne, Cochair

Naval Undersea Warfare Center, Newport, Rhode Island 02841

Joseph Vignola, Cochair

Naval Research Laboratory, Washington, DC

Chair's Introduction—8:00

Contributed Papers

8:05

4aEA1. Acoustic barrier: BEM prediction and experimental verification. Samir Gerges and Calza Arlinton (Mech. Eng. Dept., Federal Univ. of Santa Catarina, Cx.P. 476, Florianopolis, SC, Brazil)

This paper presents the noise attenuation of a barrier by BEM prediction, analytical formulas, and also experimentally. The analytical formulas used consider the diffraction of semi-infinite barrier, based on source image and phase change caused by the reflection of the sound waves from rigid ground surface. Numerical simulation by BEM and analytical calculation of the barrier model are compared and later validated by experimental tests with a barrier built in a semianechoic chamber. The results obtained by BEM prediction, analytical formulas, and experimental measurements are compared and the divergences between them are discussed.

8:20

4aEA2. Acoustic characterization for quality or damage evaluation in concrete structures. Hasson M. Tavossi (Dept. of Eng. Sci. and Mech., The Penn State Univ., 227 Hammond Bldg., University Park, PA 16802-1401)

Concrete structures such as building columns, parking platforms, concrete highways, and bridges are under continuous action of two major types of cyclic stresses. First, thermal stress cycles due to seasonal temperature variations, resulting in cyclic thermal expansion and contraction. Second, mechanical stress cycles due to cyclic load variations; for example, resulting from the passage of a vehicle on a concrete platform. The cyclic stress, after a certain number of cycles, causes damage in concrete by fatigue. In this study, fatigue damage in concrete is evaluated by acoustic techniques, using wave speed and attenuation. Acoustic sensors of 50-kHz center frequency are coupled to the concrete, as transmitter and receiver, to evaluate concrete quality after concrete samples are subjected to thermal stress cycles. Results obtained in the laboratory will be presented, showing a significant weakening of the concrete strength after only a few stress cycles. The cyclic stress causes a significant drop in acoustic wave speed of the concrete, represented by a low value of its Young's modulus. A quantitative relationship between fatigue damage and acoustic characteristics of concrete can be established, to evaluate the quality of the concrete structures.

8:35

4aEA3. Effective wide-band, low-frequency multiple resonator sound absorber. Fathy B. Shenoda and Mohammed Abd-Elbasseer (Dept. of Acoust., Natl. Inst. for Standard "NIS," P.O. Box 136, Giza 12211, Egypt)

Resonator-type sound absorbers are incorporated to provide relatively large sound attenuation in narrow bands. They are applied in several environments and have great practical value for the attenuation of low frequencies.

A detailed experimental study leads to the construction of a multiple resonator sound absorber. It is made of metal sheet and has a total volume of 0.0016 m³. The acoustic performance of the designed device, namely, the sound absorption coefficient and the normalized input acoustic impedance, were measured in 20-Hz steps at normal sound incidence in room temperature and under heating conditions up to 4000 °C. The designed device provides good sound-absorption-frequency characteristics not only in the high values of the sound-absorption coefficient, but also in covering a wide frequency band (60 to 1000 Hz) without attenuation breaks. Also, its acoustic characteristics are changed very little under heating conditions. Applying earlier results obtained by Lord Rayleigh and simplified by Crandall for the propagation of a sound wave in a tube, and recently applied by Maa to characterize microperforated sound absorbers, the acoustic performance of the designed multiple resonator was predicted. The measured and predicted results are in good agreement.

8:50

4aEA4. Sound transmission through perforated panel with internal Helmholtz resonators. Pedro E. Solana Quiros, Miguel A. Picard Lopez, and Juan V. Arizo Serrulla (Universidad Politécnica de Valencia; E.T.S.I. Industriales, Edif. D4-D5; Camino de Vera 14; 46022 Valencia, Spain)

The study of the passive control technique of noise through a perforated panel with resonators is useful in acoustic barrier design for both industrial and traffic noise. The loss mechanisms that reduce the energy of sound have been considered in the method for predicting the TL of compound flat panels containing Helmholtz resonators. A transmission coefficient is obtained for the system that relates the resonance frequency of the resonators for nonlinear resistance at the nozzle. The outer and the interior panels are assumed to act as simple mass reactances. In order to examine the validity of the analytical results, comparisons were made with experimental data. To provide acoustic excitation, a loudspeaker array was used. The excitation and transmitted sound fields were measured with a microphone array located on each side of the panel. The model predicts that a significant increase in TL can be expected at or near the mean resonance frequency of the system multilayer-panel-resonator array. In addition, a significant decrease in the TL is predicted at higher frequency. Also, it has been observed that as the ratio of the nozzle area divided by the resonator surface area increases, the frequency at which the peak TL occurs also increases.

9:05

4aEA5. Wave propagation experiments using point load excitation in a multilayered structure. Chunnan Zhou (General Electric Co., 534 Griffin Rd., Bangor, ME 04401), S. John Popovics (Drexel Univ., Philadelphia, PA 19104), and D. Jan Achenbach (Northwestern Univ., Evanston, IL 60201)

Transient wave propagation measurements on a single plate, a two-layer plate, and a two-layer-on-a-half-space structure are reported. The experiments are performed using point load excitation and point detection at the surface of the structure. The experimental results are used to verify

the accuracy of a new model for wave propagation in layered structures. The basis of the model is first introduced. Experimental measurements using different wave sources, plate materials, and wave sensors are compared with each other and with theoretical results. The actual source functions for pencil-lead-break wave sources are obtained. Next, experiments performed on an aluminum plate bonded to stainless steel plate are reported. Both cases of stainless steel atop aluminum and aluminum atop stainless steel are considered. The experimental results show good agreement with the theoretical predictions. The effects of the second layer on the transient displacement of the top surface are discussed. Finally, a two-layer-on-a-half-space structure consisting of a stainless steel plate as the top layer, an aluminum plate as the second layer, and a thick acrylic resin block as the half space is studied. The presented experimental transient measurements agree with the predicted results, thus verifying the accuracy of the model.

9:20

4aEA6. Study of acoustic emission in a plate using laser ultrasonics and the finite element method. Zhiqiang Shi, Jacek Jarzynski (Dept. of Mech. Eng., Georgia Tech, Atlanta, GA 30332-0405), and Laurence Jacobs (Georgia Tech, Atlanta, GA 30332-0355)

Acoustic emission (AE) is the spontaneous release of transient energy due to changes in localized stresses or incremental crack growth. This research studies AE generation, propagation, and detection in platelike components. The modeling of a variety of AE sources (such as micro-crack extension) and their resulting wave propagation signature, in a two-dimensional component, are investigated. The experimental technique features a laser-ultrasonic system for the point source generation and broadband, point-wise detection of elastic waves. The numerical algorithm uses the finite element method (FEM), which has the flexibility to simulate a variety of AE source types, and thus quantify source and geometric effects—geometric effects are specific component effects such as boundaries—on the predicted AE waveforms. These results provide a quantitative understanding of AE mechanisms, and assist in the interpretation of field-measured AE signals.

9:35

4aEA7. Noise and vibration control for land mine detection. R. Daniel Costley (Miltec, Inc., Natl. Ctr. for Physical Acoust., University, MS 38677), James M. Sabatier, and Ning Xiang (Univ. of Mississippi, University, MS 38677)

Land mines can be detected by scanning the ground with a laser doppler vibrometer (LDV) to measure the motion (velocity) of the ground as it is being sonified. However, the system at times suffers from a low signal-to-noise ratio, especially when trying to detect deeply buried mines. If the signal-to-noise ratio could be improved, these mines could be more easily detected. One contribution to the noise is due to the acoustically induced vibration of the LDV. This noise is difficult to filter digitally or electronically since it occurs at the same frequency as the ground motion being measured. Passive noise and vibration control remedies have been incorporated with some success. These include enclosing the LDV within a box and mounting it on damped springs. However, it is impossible to make the LDV completely stationary at all frequencies. The translation of the LDV beam over anything other than a completely smooth surface will appear to the LDV as an out-of-plane motion of the surface. The noise and vibration control strategies will be presented and discussed, along with measurements showing both their effectiveness and limitations.

10:05

4aEA8. Mode analyses in a fluid-filled borehole surrounded by a concentrically layered formation. Xiuming Wang and Kevin Dodds (CSIRO Petroleum, P.O. Box 1130, Technology Park, Bentley, WA 6102, Australia)

Acoustic wave modes including normal modes and leaky modes in a fluid-filled borehole surrounded by a concentrically layered medium are analyzed. Dispersions and excitation spectra of various modes related to monopole, dipole, including hexapole, octepole, and decapole sources, are calculated for typical well logging environments. The numerical results demonstrate that the cutoff frequencies of nonsymmetrical modes vary greatly with the orders. A Stoneley mode has no cutoff frequency in a fast formation, while at a very slow formation, it possesses a cutoff frequency. Multiple pseudo-Rayleigh modes, flexural modes, and the other higher modes have their correspondent cutoff frequencies. As frequency decreases from the cutoff frequencies, these modes correspond to the complex pole on the other Riemann sheets and become leaky modes. Also, followed by our previous work (Wang *et al.*, 1994, 1995), it is shown that head compressional and shear arrivals cannot be totally described by contributions of vertical branch cuts in a wave field representation when the Poisson's ratio is larger than 0.36 for dipole and quadrupole sources. The numerical results also show that the radial penetration of the normal modes is very shallow. The Stoneley modes are limited to 10 cm for a typical logging environment for a lower frequency than 3 kHz.

10:20

4aEA9. Recent developments in the theory of ground vibration boom from high-speed trains. Victor V. Krylov (Dept. of Civil and Structural Eng., The Nottingham Trent Univ., Burton St., Nottingham NG1 4BU, UK, Victor.Krylov@ntu.ac.uk)

Generation of ground vibration boom by high-speed trains was theoretically predicted in 1994 by the present author. The first experimental observation of this phenomenon was reported in 1997 by C. Madshus, who worked with his team on the assessment of the newly opened high-speed railway line from Gothenburg to Malmo in Sweden. The ground on the site of observation was very soft, with Rayleigh wave velocity of only 45 m/s. Therefore, for train speeds as low as 160 km/h, this Rayleigh wave velocity could be exceeded and the ground vibration boom observed. It is now well understood that ground vibration boom represents a serious hazard for the built environment, especially in the cases where high-speed lines are built on very soft soil. The present paper reviews the current status of the theory of ground vibration boom from high-speed trains. Among the problems to be discussed are contributions of different generation mechanisms, effect of track wave resonances on generated ground vibrations, effects of layered geological structure of the ground, waveguide effects of the embankments, and focusing of generated waves due to the track curvature. The results of theoretical calculations are compared with the existing experiments. [Work supported by EPSRC.]

10:35

4aEA10. New geometric sound absorbers. Glenn E. Warnaka (Future Technologies, L.L.C., 1612 S. Allen St., State College, PA 16801)

In a previous paper of the same name, two new geometric sound absorbers were discussed. The geometric absorbers use closely coupled quarter-wave resonators to absorb sound. Because the structures absorb sound by their configuration, they do not require the use of conventional porous or fibrous sound-absorbing materials and can be made of nearly

any substance such as plastic, wood, metal, etc. The absorbers can be made very rugged and long-lasting for industrial, military, and outdoor use. The previous paper discussed two basic sound absorbers, configuration delta and flat absorber, but the paper concentrated on the former sound-absorber configuration. This paper will review configuration delta but will concentrate on describing the structure and performance of the flat absorber. This absorber can be made in a thin treatment, less than 25 mm thick, that provides high acoustic absorption. A number of different designs will be shown, and the results of standing wave tube and reverberation room tests will be given. Combining the two treatments described in this paper provides a broad frequency range absorber that is both compact and rugged. Helmholtz resonators can also be used with the flat absorber concept, and this is discussed.

10:50

4aEA11. Noise attenuation using a parabolic muffler. Jonathan Zalben (Yale Univ., New Haven, CT 06520-5893, jonathan.zalben@yale.edu)

This project is a patent-pending muffler design with paraboloid baffle chambers to reduce noise pollution from internal combustion engines (especially those of lawn mowers and leaf blowers). After theorizing the most reasonable design, experiments were conducted on wood and steel (18 gauge) models. The experiments show that for a wood model, white noise produced by a radio is reduced from 80 to 64 dB (97.5% intensity reduction). The wood model amplifies sound for pure frequencies around 1000 Hz and attenuates at higher frequencies. The steel model attenuates engine noise by 75% with varying rates of attenuation for pure frequencies. [Work supported by Residents for a More Beautiful Port Washington, the Nassau-Suffolk Landscapers Association, and Port Washington School System.]

11:05

4aEA12. Improvement of one- and two-layer liner design with bias flow. Jesse Ian Follet (Virginia Consortium of Eng. and Sci. Universities, 303 Butler Farm Rd., Ste. 101, Hampton, VA 23666, jfollet@vces.larc.nasa.gov)

The coupling of the resistance and reactance in acoustic liners has in the past limited aircraft engine liner design performance. This is due to the parametric dependence of the impedance on the perforated sheet geometry and cavity depths both having a significant effect on both parts of the

impedance. For normal incidence liner impedance, designing a liner with its reactance approaching zero and resistance approaching unity is essential to the goal of producing maximum broadband absorption. Introducing a mean flow through the liner has shown that the liner resistance can be changed with minimal effect on the reactance. Liner geometry can be designed to achieve a broadband reactance that is as close to zero as possible. The resistance can then be increased to near unity by allowing a mean bias flow through the liner to produce the maximum broadband absorption. With sufficient modeling of liner impedance, an optimization routine can be used to find the liner characteristics that produce maximum absorption over a broad range of frequencies and sound-pressure levels. It will be shown that using bias flow can improve normal incidence liner absorption by 17% for single layer liners and 7% for double layer liners.

11:20

4aEA13. Variational formulations for sound in enclosed spaces and ducts. Mario Zampolli, Allan D. Pierce, and Robin O. Cleveland (Dept. of Aersp. and Mech. Eng., Boston Univ., Boston, MA 02215)

The present formulation differs from those appearing in earlier literature in that the boundary conditions are explicitly incorporated, so that not all admissible trial functions need to satisfy them. If the class of considered trial functions is restricted so that each is a finite sum of simple functions with arbitrary coefficients, then the answer selected by the variational principle attempts to satisfy the boundary conditions plus the partial differential equations imposed in the interior simultaneously, but only achieves a trade-off. An example is a waveguide of variable cross section, with one end open and the other closed. The Rayleigh integral solution for sound radiated by a baffled aperture with an arbitrary distribution of velocity amplitude on the aperture yields the open-end boundary conditions. Individual trial functions are combinations of cross-sectional mode-shape functions with axially dependent coefficients. Numerical instabilities encountered in previous direct integration of the coupled-mode equations are avoided by taking the axial coefficient functions to be piecewise linearly along the axis. The amplitudes at the nodal points satisfy a sparse system of linear algebraic equations, this system resulting directly from the variational principle. [Work supported by DARPA/USNWRG.]

FRIDAY MORNING, 2 JUNE 2000

PEACH ROOM, 10:00 A.M. TO 12:00 NOON

Session 4aED

Education in Acoustics: Take Fives—Sharing Ideas for Teaching Acoustics

Uwe J. Hansen, Chair

Department of Physics, Indiana State University, Terre Haute, Indiana 47809

Do you have a novel demonstration, a new laboratory experiment, a favorite video, a recorded sound example, or a new idea for teaching acoustics which you are willing to share with your colleagues? At this session a sign-up board will be provided for scheduling presentations. No abstracts are printed. Presenters are encouraged to have handouts to distribute. Multiple presentations are acceptable (not consecutively). Presentations are limited to 5 minutes. Keep them short! Keep them fun!

Session 4aMU

Musical Acoustics and Signal Processing in Acoustics: Audio Signal Data Compression for Musical Applications

James W. Beauchamp, Chair

*School of Music, Department of ECE, University of Illinois Urbana-Champaign, 2136 Music Building, 1114 West Nevada, Urbana, Illinois 61801***Chair's Introduction—8:25***Invited Papers***8:30****4aMU1. Psychoacoustics as the basis for modern audio signal data compression.** Tilmann Zwicker (Valpichlerstr. 70, D-80686 Munich, Germany, tzwicker@epo.org)

To almost everyone involved in music, whether they create it or listen to it, storing and transmitting music signals is of great importance. Recent progress makes it possible to store audio signals in small files and transmit them with very small transmission bandwidths, so that standardized data compression schemes, such as ATRAC (Adaptive Transform Acoustic Coding) or MP3 (Motion Picture Expert Group Layer 3 Format), are now ubiquitous. Successful development of audio storage and transmission technology has always taken into account how the final receiver of audio signals and final judge of their quality, the human auditory system, can actually appreciate them. For example, standards for frequency response and wow and flutter for analog studio tape recorders were originally designed to satisfy known limits of hearing. The use of digital signal processing to exploit newly discovered intricacies of the auditory system has increased the accuracy with which stored and transmitted audio signals can be tailored to the requirements imposed by our hearing. Such accurate matching provides the basis for highly effective audio-signal data compression. This talk provides an overview of the related idiosyncrasies of human hearing, in particular, masking effects in the frequency and temporal domains.

9:00**4aMU2. Three approaches to the perceptual evaluation of audio compression methods.** Steven van de Par and Armin Kohlrausch (Philips Res. Labs. Eindhoven, Prof. Holstlaan 4, NL-5656 AA Eindhoven, The Netherlands and IPO-Ctr. for User-System Interaction, 5600 MB Eindhoven, The Netherlands)

The compression of raw digital audio data is commonly accomplished by coding the audio data into a representation where part of the information present within the original data is lost. Upon playback this loss of information reveals itself in terms of various distortions, whose audibility depends on the coding method used. Therefore, the success of a coding method is determined by the human auditory perception of coding distortions. Three approaches to the perceptual evaluation of coding methods will be discussed. The first deals with the audibility of coding errors and is useful for methods that aim to achieve audio that is indistinguishable from the original. The second approach measures the degree of perceived degradation of coded audio, allowing comparison of different coding schemes which introduce different types of clearly audible distortions. The third approach deals with perceived intrinsic quality. For the assessment of quality, the observer is assumed to have no knowledge of the original signal. This approach is useful in situations where the original is not available to the observer. Here a judgment about quality is reached on the basis of the observers' expectations with regard to the type of audio signal heard.

9:30**4aMU3. A malleable audio representation for data compression.** Scott N. Levine (Liquid Audio, 2221 Broadway, Redwood City, CA 94603)

The boundaries between audio data compression and audio synthesis/modifications are blurring as research in both fields progresses. New results using a sines + transients + noise parametric representation allows for coding gains as efficient as the most recent nonparametric transform coding schemes. Because the audio input is decomposed into separate sinusoidal, transient, and noiselike signals, aggressive quantization can be performed independently on each signal in a perceptually meaningful manner. While the sines + transients + noise representation allows for good compression rates, it also allows for high-quality time-scale modification in the compressed domain. The multiresolution sinusoidal modeling and the Bark-band noise modeling utilized in this representation have been used in both the music synthesis and audio/speech coding arenas. The transient modeling in this representation uses similar algorithms to those used in transform coding systems.

10:00

4aMU4. New approaches to sound compression that use algorithmic synthesis. Eric D. Scheirer (Machine Listening Group, MIT Media Lab., E15-401D, Cambridge, MA 02139-4307, eds@media.mit.edu)

Research into the use of perceptual models to compress sounds has achieved fruitful success in recent years. However, quantized filterbanks and linear-predictive coding are not the only techniques that can be used for audio compression. Recent research has tightened the connection between sound processing normally conceived as “compression” and that normally conceived as “synthesis.” By making this connection explicit, it becomes possible to apply methods taken from the broad literature on the synthesis of sound signals to applications in sound compression [B. L. Vercoe *et al.*, Proc. IEEE **85**, 622–640 (1998)]. New techniques that use audio synthesis to enable compression have recently emerged. One of these is termed *algorithmic structured audio*, and builds on research into software synthesis and sound-description languages. A particular implementation of this method forms the new MPEG-4 Structured Audio standard, in which the software-synthesis language SAOL is used to represent sound algorithmically. Recent research also suggests that the development of so-called *generalized audio coding* techniques, in which algorithmic-structured-audio formats carry customized perceptual decoders, might increase the overall efficiency of the marketplace of coding by allowing a more flexible approach to the selection of coding models.

10:30

4aMU5. Internet audio as seen by the producer/musician. Eric Somers (Dept. of Performing and Visual Arts, Dutchess Community College of the State Univ. of New York, Poughkeepsie, NY 12601)

In the battle over Internet audio transmission schemes, each format seeks to become the universal standard of the internet. Yet the author, a sound designer, composer, and producer, asserts that no one compression technology can best serve the various needs of the musician. The paper classifies these needs, suggests which schemes serve each best, and suggests musical and production issues that designers of audio compression technologies need to keep in mind.

11:00

4aMU6. Encoding considerations for MP3 and MPEG-2/MPEG-4 advanced audio coding. Karlheinz Brandenburg (Fraunhofer Institut Integrierte Schaltungen, 91058 Erlangen, Germany and Ilmenau Tech. Univ., 98684 Ilmenau, Germany)

Perceptual coding of high-quality audio has found widespread applications for broadcasting, internet delivery of music and storage of music as used in portable music players. The success of formats like MPEG-1 Layer-3 (AKA MP3) can be attributed to the standardization philosophy of MPEG: A universal format is specified and available for licensing for everybody. This leaves the encoder up to technical improvements. The presentation will (i) introduce the basics of high-quality audio coding, (ii) give an overview on MP3 and MPEG-2/4 AAC, (iii) explain encoding strategies, and (iv) give special consideration to parameters affecting coding quality.

FRIDAY MORNING, 2 JUNE 2000

TOWER 1205, 9:00 TO 10:45 A.M.

Session 4aNS

Noise: Topics in Noise—Noise and Vibration Effects and Noise Exposure Modeling

Richard L. McKinley, Chair

*Air Force Research Laboratory, AFRL/HECB Building 441, 2610 Seventh Street,
Wright-Patterson Air Force Base, Ohio 45433-7901*

Contributed Papers

9:00

4aNS1. The effects of airborne vibration on human body vibration response. Suzanne D. Smith (Air Force Res. Lab., AFRL/HECB, 2610 Seventh St., WPAFB, OH 45433-7901, suzanne.smith@wpafb.af.mil)

Aircraft ground operations and maintenance personnel can be exposed to whole-body vibration via the airborne transmission of acoustical energy. The purpose of this study was to characterize human body vibration response during exposures to airborne vibration generated by military fighter aircraft during high-power ground engine runs. Miniature triaxial accelerometers were mounted on the head (bitebar), chest, spine, and lower leg of the subject. Measurements were made for selected locations

along a line parallel to the longitudinal axis of the aircraft at specified engine-power settings. The highest accelerations occurred in the fore-and-aft (X) chest response. One-third-octave analysis showed a distinct peak in the chest between 50 and 100 Hz not observed in the noise levels, strongly suggesting the presence of a chest resonance. These peaks increased as the subject moved aft of the aircraft. While the noise levels also increased, it was difficult to determine a relationship between acceleration and noise without additional data due to differences in the results between power settings. The subject did report an increased sensation of vibration in the upper torso, which coincided with the increased noise levels. These data will be used in the development of human airborne vibration exposure guidelines.

4aNS2. A study of human response to helicopter interior noise.

Brenda M. Sullivan (NASA Langley Res. Ctr., M/S 463, Hampton, VA 23681, b.m.sullivan@larc.nasa.gov), Mark W. Davis, David L. Young, Douglas G. MacMartin (United Technologies Res. Ctr., East Hartford, CT 06108), and Thomas A. Millott (Sikorsky Aircraft Corp., Stratford, CT 06615)

Two studies investigating low-frequency tonal content of rotorcraft noise were performed in a joint project between NASA Langley Research Center, United Technologies Research Center, and Sikorsky Aircraft. The first study investigated preference reported by subjects exposed to noise representing that experienced inside helicopters passenger cabins or crew compartments. The second study examined the use of computer-based tools to measure possible fatigue caused by long-term exposure to helicopter interior noise. In the first study subjects heard sounds with modified low-frequency rotor-generated tones, high-frequency gear-mesh tones, and mid-frequency broadband noise. Sounds were presented in pairs and subjects reported degree of preference. The rank ordering of preference shows the perceived benefit of reducing low-frequency tones versus higher-frequency broadband noise. In the second study subjects were exposed for one 6-h period to a helicopter-type sound environment at 88 dBA and for one 6-h period of exposure to a control condition of minimal noise exposure. During each of these test-exposure days, the subjects were required to complete computer-based tests at specified intervals throughout the day, to monitor any change in fatigue or alertness levels. The most relevant conclusions from the tests will be presented in the paper.

4aNS3. The investigation of vibroacoustical activity of the combined fuel pump.

Valeri V. Lenchine (Yale Univ., P.O. Box 208206, New Haven, CT 06520-8206, valeri.lenchine@yale.edu), Alexander N. Kruchkov, Andrei B. Prokofiev, and Evgeniy V. Shakhmatov (Samara St Aersp. Univ., Samara, 443086, Russia)

A theoretical and experimental investigation of the fuel pump for a gas turbine engine is described. The combined pump consists of blade and gear units. Under the pump exploitation the graphite bearing of the blade unit was being destroyed due to the high level of vibroacoustical loading at the resonant working conditions. The complex dynamic model of the pump was created to estimate pulsation and vibration productivity of the pump. The mechanism and sources of intensive vibroacoustical stresses were identified. The complex test of the fuel system allowed us to define parameters of pulsing liquid, vibration of the pump (including natural frequencies), oscillation of torque and rotation of the driving shaft, and sound power of the pump. The test's results have good coincidence with the modeling data and confirm the hypothesis regarding the reason for the bearing destruction. The facilities to decrease vibroacoustical stresses were proposed after analyzing the dynamical characteristics of the fuel system. The effectiveness of the facilities was proved by modeling and natural experiments. As a result of the research, vibration, pulsation, and noise activity of the pump were reduced significantly, and the defect connected with the destruction of the bearing was eliminated.

4aNS4. Psychoacoustic comparison of two methods for the evaluation of prominent discrete tones.

Anne C. Balant (Dept. of Commun. Disord., State Univ. of New York at New Paltz, New Paltz, NY 12561 and IBM Hudson Valley Acoust. Lab., balanta@matrix.newpaltz.edu), Kristin Barringer (State Univ. of New York at New Paltz, New Paltz, NY 12561), and Matthew Nobile (IBM Hudson Valley Acoust. Lab., Poughkeepsie, NY 12601-5400)

Two objective methods have been developed to evaluate the potential annoyance of prominent discrete tones in the noise emissions of products: (1) the tone-to-noise ratio (TNR) method [ECMA-74-1997, ANSI S1.13-1995] and (2) the prominence ratio (PR) method [ANSI S1.13]. A task group of the Inter-Committee Working Group on noise from information technology and telecommunications equipment (ITTE) is studying both procedures, with the goal of optimizing a single method. In some cases the

procedures give conflicting results, especially when there are harmonics and/or multiple tones within a single critical band. ECMA-74-1997 employs a modification of the TNR method for cases in which there are two tones within the critical band. This talk will present the results of two pilot studies comparing naive listeners' subjective ratings of signals containing discrete tones to the measured TNR and PR values. In the first pilot study, stimuli were a series of recorded machine noises. In the second study, stimuli were constructed by mixing the noise produced by axial fans with a pure tone close to the fundamental frequency of the fans. The level and frequency of the tone was varied. The results of these pilot studies and recommendations for future work will be discussed.

4aNS5. Designing process plants to meet a noise limit based on verifying by a noise model rather than measurements.

Frank Brittain (Bechtel Corp., 50 Beale St., San Francisco, CA 94105, ffbritta@bechtel.com)

Normally, in the US process plants are designed to meet a noise limit based on verifying the limit has been met using noise measurements. An alternative approach is to design the plant based on confirming the noise limit has been met using a noise prediction model. The latter approach is often used in Europe and Australia. The strategy used in designing the plant depends in part on how meeting the noise limit is to be verified. Verification using a model has some distinct advantages and some disadvantages. The advantages of verification using a model include independence from variations in atmospheric conditions, and ambient levels (absolute and changes) at locations where the limit must be met. Verification using a model necessitates very close advance agreement among all parties on how the model is developed and approved. Further, the issues of experimentally verifying the model and updating the model with actual noise levels of equipment need to be addressed by all parties. These issues will be discussed, and alternative design strategies identified.

4aNS6. Use of structured, interactive interviews in retrospective noise exposure assessment in an occupational epidemiologic study.

Mary M. Prince, Martha A. Waters (CDC/NIOSH, 4676 Columbia Pkwy., R-16, Cincinnati, OH 45226, mmp3@cdc.gov), Robert R. Anderson, and Richard R. James (James, Anderson, & Assoc., Okemos, MI 48864)

As part of a NIOSH study examining factors affecting hearing conservation program (HCP) effectiveness, a job-noise exposure matrix was constructed using these data sources: (1) task-based sound level survey data; (2) noise exposure data (dosimetry and sound level surveys) provided from plant historical reports; (3) information on process changes and engineering controls; (4) interviews with plant personnel (in engineering and safety departments); and (5) detailed work history data from personnel records for each employee in the plant HCP. For plants in which changes in exposures have occurred due to engineering control and process changes, exposure estimation becomes a challenge when there are data gaps in exposure and process changes. This paper discusses how data from structured employee interviews can be used in conjunction with available quantitative, records-based data to reconstruct processes and machinery/layout history. The goal was to characterize how noise exposure determinants (manufacturing environment, equipment, processes, shift lengths) have changed and to estimate exposure by department, job, and era. The first step was to identify employees for interview whose jobs were located in departments where exposure data were sparse or nonexistent. Data collection efforts were then geared towards reconstructing noise exposure for specific departments and jobs over time.

4aNS7. Using Caltrans noise analysis protocol methodology to determine insertion loss of classrooms at a high school. Michael Greene (URS Greiner Woodward Clyde, 2020 E. First St., Ste. 400, Santa Ana, CA 92705)

The construction of a new freeway adjacent to an existing high school in eastern San Diego County, California, prompted the need for a rigorous analysis of the noise effects on the school. The insertion loss of structures (with windows and doors open and closed) at a high school was measured using the recently published California Department of Transportation (Caltrans) Noise Analysis Protocol. Both the school district and Caltrans

agreed upon the details of the measurement methodology prior to the tests. The test setup consisted of two commercial-grade loudspeakers mounted atop a manually operated lift, associated amplifiers, pink-noise generator, a real-time noise analyzer, and sound-level meters. Noise levels were measured at equivalent distances in the absence of and then inside the room of interest, to derive the structure's insertion loss. This was done at incident angles of 30, 45, 60, and 75 deg to the building façade. The resultant data from these measurements required the use of specially designed spreadsheets to effectively analyze and present the results. The results of the measurements indicated that improvements to the older classrooms near the freeway would be necessary in order to meet the indoor noise standard for classroom spaces.

FRIDAY MORNING, 2 JUNE 2000

SPANISH ROOM, 9:00 TO 11:35 A.M.

Session 4aPA

Physical Acoustics: Acoustics of Sand, Paper and Foam

Julian D. Maynard, Jr., Chair

Department of Physics, The Pennsylvania State University, 104 Davey Laboratory, State College, Pennsylvania 16802

Chair's Introduction—9:00

Invited Papers

9:05

4aPA1. Impulse propagation in granular media. Surajit Sen (Dept. of Phys., State Univ. of New York at Buffalo, Buffalo, NY 14260-1500)

Grains can be described as elastic objects. When two grains are pushed against one another, they repel via Hertz' nonlinear force law, $F \propto \delta^n$, $n > 2$, $\delta \geq 0$, being the overlap between the grains. We show that the propagation of an impulse of any magnitude in a 1D chain of grains at zero loading can be described as a solitary wave. The width of the solitary wave, $L(n) \rightarrow 1$ as $n \rightarrow \infty$, $L \sim 5$ grain diameters for typical granular contacts and $L(n) \rightarrow \infty$ when $n \rightarrow 2$. The condition $\delta \geq 0$ leads to the formation of secondary solitary waves when two identical solitary waves propagating in opposite directions collide. Randomness in grain densities and sizes and restitutional losses lead to approximately exponential decays in distance of the energy of a propagating solitary wave. It turns out that impulses can be exploited to fingerprint a buried impurity mass in a 1D chain with and without gravity. In closing, we shall discuss the backscattering of an impulse from a buried object in 3D beds at shallow depths. Implications of this work with respect to humanitarian demining and related applications will be presented. [Work supported in part by U.S. Army Corps of Engineers and Sandia National Laboratories. Work done in collaboration with Marian Manciu (SUNY-Buffalo) and Alan J. Hurd (Sandia National Labs).]

9:25

4aPA2. The noise from a crumpled candy wrapper as a probe of a disordered system. Eric M. Kramer (Simons Rock College, Great Barrington, MA 01230)

We discuss the origin and properties of the crackling sound emitted by a crumpled sheet of Mylar as it is strained. These sheets possess many qualitative features of a traditional disordered system, including frustration and discrete memory. The sound can be resolved into discrete clicks, which are emitted during rapid changes in the conformation of the sheet. Observed click energies range over six orders of magnitude. The measured energy autocorrelation function for the sound is consistent with a stretched exponential, $C(\tau) = A \exp[-B\tau^{0.35}]$. The probability distribution of click energies has a power law regime, $p(E) \sim 1/E$, independent of sheet size and material. We also discuss future directions for this research.

4aPA3. Avalanches, Barkhausen noise, and disorder-induced critical behavior. Karin Dahmen (Dept. of Phys., Univ. of Illinois at Urbana—Champaign, Urbana, IL 61801-3080)

Hysteresis loops are often seen in experiments at first-order phase transformations when the system goes out of equilibrium. They may have a macroscopic jump, roughly as seen in the supercooling of liquids, or they may be smoothly varying, as seen in most magnets. We have studied the nonequilibrium zero-temperature random-field Ising model as a model for hysteretic behavior at first-order phase transformations. As disorder is added, one finds a transition where the jump in the magnetization (corresponding to an infinite avalanche) decreases to zero. At this transition the model exhibits power law distributions of noise (avalanches), universal behavior, and a diverging length scale, which should be detectable in acoustic emission and other noise measurements. We study universal properties of this critical point using renormalization group methods and numerical simulations. Connections to experimental systems such as athermal martensitic phase transitions (with and without “bursts”) and the Barkhausen effect in magnetic systems will be discussed. Similar ideas can also be applied to the interpretation of the Gutenberg–Richter scaling law in the statistics of earthquakes.

10:05

4aPA4. Attenuating stress waves during the fracture of a brittle carbon foam using ferrofluid damping. L. C. Krysc (Dept. of Phys., Univ. of the Pacific, Stockton, CA 95211)

Brittle fracture is a complex problem where details of defects and disorder on the microscopic scale affect the macroscopic strength of a material. We have found that during the fracture of disordered brittle carbon foam samples, stress waves produced by the breaking of individual struts may initiate avalanches of further strut-breaking events. If conditions are right for a macroscopic, sample-sized avalanche to be produced in this manner, the sample ruptures. If the stress waves are damped, by acoustic damping techniques, it may be possible to delay the onset of fracture in this material. The effect of damping the stress waves using viscous and magnetic ferrofluids on the measured fracture strength will be discussed.

10:25–10:35 Break

Contributed Papers

10:35

4aPA5. A laser-based ultrasonic technique to measure the dependence of the bending stiffness of copy paper on moisture and temperature. David A. Griggs and Yves H. Berthelot (Woodruff School of Mech. Eng., Georgia Inst. of Technol., Atlanta, GA 30332-0405, yves.berthelot@me.gatech.edu)

A noncontact system is used to generate and detect, optically, a broad-band ultrasonic pulse propagating in paper under controlled temperature and moisture conditions. The generation laser (Nd–Yag) produces a short optical pulse which is directed through an optical window on the paper sample inside an environmental chamber where temperature and relative humidity are controlled. Ultrasonic Lamb waves propagating in the paper are detected by a fiberoptic Mach–Zehnder interferometer (argon–ion laser) where the Doppler-shifted light exits the chamber through a small-core multimode fiber. The moisture content is measured by monitoring the weight of the paper sample relative to a dry sample. The dispersive nature of the A0 Lamb mode of the ultrasonic signal contains information about the bending stiffness normalized by the density and thickness of the paper averaged over the propagation path. The frequency dependence of the group velocity is found from the analytic wavelet transform of the signal (complex Morlet wavelet). [Work supported by the Department of Energy and the Institute of Paper Science and Technology.]

10:50

4aPA6. A measure of the dynamic tortuosity in stratified fibrous porous materials. Pedro E. Solana Quiros, Miguel A. Picard Lopez, and Juan V. Arizo Serrulla (Universidad Politécnica de Valencia, E.T.S.I. Industriales Edfs. D4-D5, Camino de Vera 14, 46022 Valencia, Spain, psolana@fis.upv.es)

Fibrous porous materials are an important element for noise control. The dynamic tortuosity is a magnitude that contains phenomenological and physical information of these materials and represents a tool of great usefulness for acoustical models and studies of fluid mechanics within them. From a phenomenological approach, this dynamic parameter includes the density of the fluid in the pores and the dynamic flow resistance and, from a microscopic point of view, contains information of the structure shape and the dynamic structure factor. It can be obtained theoretically through microscopic approximations, but it can also be calculated through acoustic procedures. From a phenomenological approximation in this work, an acoustic method of calculation is proposed to obtain it based on the measurement of the complex flow impedance of the thin samples. This will be more accurate when the sample is more rigid. Finally, the experimental procedure by a two-microphone method and the existence of relationships to microscopic parameters can facilitate the studies of acoustic behavior and optimization of these materials for the manufacturer as well as for the technician that employs them in projects of noise control.

4aPA7. The use of compressional and shear wave velocities to measure the shear strength of soil. F. Douglas Shields and Jim Sabatier (Natl. Ctr. for Physical Acoust., Coliseum Dr., University, MS 38677, dshields@olemiss.edu)

Compressional and shear wave velocities have been measured in soil contained in a rigid-walled cylinder in both the axial and radial direction. This oedometer cell was fitted with a piston so that the soil could be axially compressed. The velocities and axial strain in the soil were measured as the axial stress was increased from zero to 35 lbs./sq. in. Measurements were made on both dry and wet soil. The axial and radial velocities were used to determine the radial stress. The ratio of the radial to axial stress gave the coefficient of earth pressure at rest and the angle of shearing resistance. For both the wet and dry soil, this angle was approximately 33 degrees. The velocities in wet and dry soils had approximately the same dependence upon axial stress, but the wet soil was much more compressible than the dry soil. Both wet and dry compressed soils maintained their strain and strained velocities upon decompression. However, when water was introduced into the decompressed soil, it expanded significantly and the velocities dropped to their precompression values. [Work supported by the USDA.]

4aPA8. The effect of moisture on compressional and shear wave speeds in unconsolidated granular materials. F. Douglas Shields, Jim Sabatier (Natl. Ctr. for Physical Acoust., Coliseum Dr., University, MS 38677, dshields@olemiss.edu), and Mark Wang (Sensys Instruments Corp., Sunnyvale, CA 94089)

The effect of water vapor on the shear and compressional wave speeds in two different kinds of glass beads and in Ottawa sand has been measured. The nominal diameter of the glass beads was 125 μm and of the sand, 500 μm . The measurements were made as the water vapor was introduced slowly into the evacuated material. The vapor pressure isotherm for the beads made of titanium-barium glass was fit reasonably well by the simple BET theory. For the Ottawa sand the BET theory fit the vapor pressure isotherm if the surface area of the grains was assumed to be three times the area calculated, assuming all of the grains were spheres with a diameter of 500 μm . In these two materials the vapor had little effect on the wave speeds. For beads made of lime glass, however, the wave speeds approximately double with the introduction of water vapor, and the vapor pressure isotherm had the BET shape only if the saturated vapor pressure was assumed to be lowered by 20%. These results have been explained by assuming that a chemical reaction occurred between the lime glass and the water to form a gel. [Work supported by the USDA.]

FRIDAY MORNING, 2 JUNE 2000

STATE ROOM, 8:30 TO 11:45 A.M.

Session 4aPP

Psychological and Physiological Acoustics: Complex Sound Perception

Jennifer J. Lentz, Chair

Army Audiology and Speech Center, Walter Reed Army Medical Center, Washington, D.C. 20307-5001

Contributed Papers

8:30

4aPP1. Joint detection-recognition of amplitude modulation. Stanley Sheft and William A. Yost (Parham Hearing Inst., Loyola Univ., 6525 N. Sheridan Rd., Chicago, IL 60626, ssheft@luc.edu)

Joint detection-recognition ability was measured for amplitude modulation. On each trial there were two subject responses, one for detection and the other for recognition. Stimuli were samples of wideband noise sinusoidally modulated at either 4, 16, 64, or 256 Hz. Conditions included all pairwise combinations of the four AM rates. Rate uncertainty led to only a slight decrement in detection ability with little effect of the extent of the separation between the two rates. Recognition performance was always poorer than detection ability with recognition near chance on trials in which the detection response was incorrect. When one of the AM rates was 4 Hz, the detection-recognition theorem [Starr *et al.*, *Radiology* **116**, 533–538 (1975)] provided a reasonable prediction of recognition ability. For the three other pairwise combinations of AM rate, recognition performance was poorer with the theorem failing to account for the extent of the decrement. The theorem requires orthogonality of signals. Results then suggest independence of 4 Hz versus higher rate processing, and also correlation in processing the higher rates. This pattern of orthogonality is consistent with two rate-dependent cues in modulation processing. [Work supported by NIH.]

8:45

4aPP2. An effect of auditory-filter envelope modulation variability on masking. William Treurniet (Communications Research Ctr., P.O. Box 11490, Sta. H, Ottawa, ON K2H 8S2, Canada, bill.treurniet@crc.ca)

In an earlier report [W. C. Treurniet and D. R. Boucher, *J. Acoust. Soc. Am.* **106**, 2146 (1999)], the threshold for a noise probe masked by a complex masker was shown to be lower when the masker partials were harmonically related than when the frequencies of the partials were perturbed by small amounts. The results suggest that the improved detection may arise from differences in the pattern of modulations across auditory-filter outputs in response to harmonic and inharmonic maskers. That is, envelope modulations are identical across filters for a harmonic masker, but the modulations vary for an inharmonic masker or a noise probe. The variable modulation rates across filters due to the noise probe are easier to detect against the invariant background rates generated by the harmonic masker than against the variable background rates resulting from the inharmonic masker. In support of this hypothesis, the difference in masked threshold was strongly related to the variance of the dominant period of envelope modulations measured across auditory-filter outputs. Further, the masking level differences were minimal when the average variance of modulation amplitude was small, as in the case of a single partial per auditory filter.

4aPP3. The effect of speech reading on masked detection thresholds for spoken sentences. Ken W. Grant (Walter Reed Army Medical Ctr., Army Audiol. and Speech Ctr., Washington, DC 20307-5001)

Detection thresholds for spoken sentences in steady-state noise are reduced by 1–3 dB when synchronized video images of movements of the lips and other surface features of the face are provided. In a previous report [K. W. Grant and P. F. Seitz, *J. Acoust. Soc. Am.* **103**, 3018 (1998)], we showed that the amount of masked threshold reduction, or *bimodal coherence masking protection* (BCMP), depended on the degree of correlation between the rms amplitude envelope of the target sentence and the area of lip opening. In the present study, we extend these results by directly manipulating this cross-modality correlation through either bandpass filtering or amplitude adjustments of selected words contained in the target sentences. A control condition was also included in which visual orthography was provided to explicitly identify the target sentence prior to each test trial. Results showed that orthographic information reduced detection thresholds by about 0.5 dB for all target sentences. Preliminary results for filtered and amplitude-adjusted sentences suggest that the magnitude of the BCMP depends primarily on the cross-modality correlation between lip-area function and rms amplitude envelope computed over windows of approximately 300 ms. [Work supported by NIH and the Department of Clinical Investigation, Walter Reed Army Medical Center.]

4aPP4. Masking period patterns in listeners with cochlear hearing loss. Magdalena Wojtczak, Anna C. Schroder, and David A. Nelson (Clinical Psychoacoustics Lab., Univ. of Minnesota, 516 Delaware St. SE, Minneapolis, MN 55455)

Masking period patterns (MPPs) were measured in listeners with hearing loss of cochlear origin, using a short 6-kHz probe presented at different times within the envelope of a 100% sinusoidally amplitude-modulated (AM) masker (4-Hz modulation rate). The carrier frequency of the AM masker was 3 kHz (off-frequency masking) or 6 kHz (on-frequency masking). MPPs obtained from hearing-impaired listeners were compared to those obtained from normal-hearing listeners using the same stimuli. A model used previously to simulate MPPs from normal-hearing listeners [Wojtczak *et al.*, *J. Acoust. Soc. Am.* **106**, 2147 (1999)] was used to predict data from the hearing-impaired listeners. In contrast to MPPs from normal-hearing ears, MPPs from ears with moderate to profound hearing loss did not exhibit large differences in MPP shapes for on- and off-frequency maskers, i.e., they did not exhibit longer MPP valleys for the off-frequency masking situation. Results of the modeling suggest that the similar shapes for on- and off-frequency MPPs reflect a lack of compression (or reduced compression) in the hearing-impaired listeners. The implications of this effect for speech understanding in the presence of competing talkers will be discussed. [Work supported by NIH NIDCD DC00149 and the Lions 5M Hearing Foundation.]

4aPP5. Monaural phase effects: Timing versus level cues. Martin P. Law, Mark A. Stellmack, and Neal F. Viemeister (Dept. of Psych., Univ. of Minnesota, Minneapolis, MN 55455, mpl@nextear.psych.umn.edu)

As shown in previous studies, detection thresholds for a 500-Hz signal vary as a function of its phase relative to a 250-Hz masker. Due to differences in detectability, phase discrimination between suprathreshold 250- and 500-Hz components may be possible using the difference in effective stimulus level as a cue. Alternatively, phase discrimination may be based on changes in temporal fine-structure. If level were the cue, then the psychometric function [P(C) versus level of the 500-Hz component in the signal interval] should be V-shaped with a minimum at chance when the effective levels of the 500-Hz components in the signal and nonsignal intervals are equal. While the data did exhibit the predicted V-shaped function, performance dropped to chance only when the masker intensity

level was below 70 dB, suggesting that fine-structure information can be used, but only above 70 dB. In another experiment, phase discrimination was assessed when the overall level of the two-tone complex was roved on each presentation. Small rove ranges significantly degraded performance, consistent with the use of a level cue. Overall, these data suggest that even at low frequencies monaural phase effects are not primarily based on fine-structure timing. [Work supported by NIDCD DC00683.]

4aPP6. Monaural phase effects in the discrimination of the spectral shape of transients. Gregory H. Wakefield (Dept. of Elec. Eng. and Comput. Sci., The Univ. of Michigan, 1301 Beal Ave., Ann Arbor, MI 48109-2110, ghw@eecs.umich.edu), Laurie M. Heller (Naval Submarine Medical Res. Lab., Groton, CT 06349-5900), Laurel H. Carney (Boston Univ., Boston, MA 02215), and Maureen Mellody (The Univ. of Michigan, Ann Arbor, MI 48109)

Limits on temporal resolution in auditory perception range from 25 ms, for the judgment of temporal order, to a lower limit of 3–4 ms for the discrimination of monaural phase. Results are presented on the discrimination of spectral shape for signals concentrated within this lower bound. A wideband, 4-ms noise was compared with spectrally smoothed versions. The phase spectrum was controlled by assigning the same random phase spectrum to both the original and smoothed signals. Depending on the choice of phase spectrum, discrimination thresholds for spectral shape were found to vary from 2–3 dB to as much as 15–17 dB, at which point the smoothed spectrum is essentially flat. This dependence on phase can be eliminated by presenting a train of transients, rather than a single transient. Such findings do not appear to be accounted for by synthesis artifact, nor by any simple features of either the complex spectrum or waveform envelope of these wideband signals. Phase effects are observed, however, in a computational model of the auditory nerve. [Work supported by the following grants from ONR: 61153.04114.00, MURI Z883402, and US DOD N66604-96-C-H366.]

4aPP7. Growth of spectral contrast enhancement in Schroeder-phase harmonic complexes. Laura E. Dreisbach, Marjorie R. Leek, and Jennifer J. Lentz (Army Audiol. and Speech Ctr., Walter Reed Army Medical Ctr., Washington, DC 20307, ldreis@attglobal.net)

Discrimination of spectral differences in harmonic complexes is better for waveforms that are highly modulated and presented at high levels relative to lower-level stimuli with flatter temporal envelopes [M. R. Leek and V. Summers, *J. Acoust. Soc. Am.* **94**, 2074–2082 (1993)]. This may reflect an enhancement of contrast between spectral peaks and valleys related to nonlinear cochlear processing. To further explore this relationship, discrimination of harmonic complexes with slightly different peak frequencies was measured at several stimulus levels. Peak-to-background differences required for discrimination were taken as a measure of the spectral contrast preserved in the auditory system. Phases were selected to generate either flat or peaked internal waveforms (negative or positive Schroeder phase), and peak frequency regions were near 2, 3, or 4 kHz. In normal-hearing listeners, the positive Schroeder waveforms produced lower spectral contrast thresholds than the negative Schroeder stimuli at all tested intensities and peak regions, with the greatest differences at moderate stimulus levels. Generally, hearing-impaired listeners demonstrated smaller threshold differences due to phase selection. These results are consistent with normal cochlear processing that is nonlinear at moderate levels, but more linear at low and high levels, and with more linear processing in impaired ears. [Work supported by NIH.]

10:30

4aPP8. Hearing discrete elements of a temporal sequence: Learning and generalization. Robert S. Schlauch, Jeffrey J. DiGiovanni, and Dennis T. Ries (Dept. of Commun. Disord., Univ. of Minnesota, 115 Shevlin Hall, Minneapolis, MN 55455, Schla001@tc.umn.edu)

The ability to hear discrete elements of a temporal sequence was assessed. Subjects judged whether the first element of a three- or four-burst sequence of 50% duty cycle noise that alternated in level from low to high (or high to low) began with a low-level (60 dB SPL) or a high-level (80 dB SPL) burst. The rate was varied from trial-to-trial [selected values between 5 to 40 pulses per second (pps)] to prevent absolute duration and loudness cues. Initially, listeners were at chance performance for rates above roughly 20 pps. However, after 6 h of practice with feedback performance was near 90% for the highest tested rate (45 pps). After training with the stimulus that alternated in level, subjects were presented with a new condition; the three- or four-pulse sequences began with either a high-level or a low-level burst of noise, as before, but the level of each subsequent burst in the sequence was selected randomly between 60 and 80 dB SPL. Performance remained high for this condition, suggesting that listeners did not learn patterns of sequences but, rather, were able to hear discretely the first element in the sequence. [Work supported by NIH.]

10:45

4aPP9. Specificity of learning in an auditory temporal-order task. Beverly A. Wright and Julia A. Mossbridge (Audiol. and Hearing Sci. Prog., 2299 North Campus Dr., Northwestern Univ., Evanston, IL 60208-3550)

The ability to determine the onset order of two tones at different frequencies is a measure of temporal acuity and an analog of voice-onset time. Learning in this temporal-order task and its generalization to untrained frequencies and temporal-acuity tasks was explored. Temporal-order thresholds of six listeners trained 1 h per day for 6 days with tones at 0.25 and 4 kHz improved from 58 ms before to 12 ms after training, yielding a proportional improvement greater than on all untrained conditions. Comparisons between proportional improvements of trained and control listeners revealed that trained listeners learned considerably more than controls on the trained condition, and only slightly or no more than controls on: (1) the trained task at two sets of untrained frequencies (0.5 and 1.5 kHz, 0.75 and 1.25 kHz), and (2) three untrained temporal-acuity tasks (offset temporal order, onset asynchrony, offset asynchrony) at the trained frequencies. The lack of robust generalization to untrained frequencies and tasks suggests learning in onset temporal-order tasks is mediated by a frequency-dependent mechanism different from the mechanism(s) underlying performance in other temporal-acuity tasks. By increasing our understanding of plasticity in temporal-acuity mechanisms, these data may guide the treatment of related disorders. [Work supported by NIDCD.]

11:00

4aPP10. Attention and the build-up of auditory stream segregation. Robert P. Carlyon, Rhodri Cusack, Jessica M. Foxton, and Ian H. Robertson (MRC Cognition and Brain Sci. Unit, 15 Chaucer Rd., Cambridge CB2 2EF, UK)

In the baseline condition of experiment 1, subjects were presented with a repeating ABA–ABA sequence of 125-ms tones in their left ear, and indicated, throughout the entire 20-s sequence, whether they heard one or

two streams. The proportion of “two stream” responses increased during the course of each sequence. This build-up of stream segregation was greatly reduced when, for the first 10 s of each sequence, subjects performed a competing task on a train of noise bursts presented to the right ear. The competing task involved identifying the amplitude envelope of each noise burst as either rising or falling. The noise bursts had no effect in a third condition, in which subjects were instructed to ignore them. Experiment 2 showed that a substantial build-up of streaming can still be observed when subjects switch tasks halfway through the tone sequence, provided that both tasks require them to continually attend to the tones. We conclude that attention is crucial for the build-up of stream segregation, and that our results are inconsistent with models of streaming based purely on peripheral, “automatic” neural mechanisms.

11:15

4aPP11. The influence of noise, harmonic structure, and temperament on tonal consonance judgments. James Kondash and Rodney Hallgren (Dept. of Psych., Wright State Univ., Dayton, OH 45435, jkondash@sdl.psych.wright.edu)

Although a large body of research exists on tonal consonance in which dyadic musical intervals have been systematically rank ordered via psychophysical experiments and models, none of these studies has addressed how tonal consonance judgments may change in the presence of a masking stimulus. Presenting musical intervals and maskers simultaneously could reduce elements potentially contributing to dissonance, such as beats between upper partials or distortion products. This study investigated consonance ratings of musical dyads as influenced by (1) the level of broadband and band-limited noise, (2) the number and level of upper partials in the tones, and (3) the use of equal or just intonation. In a randomized factorial experiment, participants rated musical dyads in the various conditions on a consonance/pleasantness scale. Preliminary results for equal-tempered dyads presented in the quiet show good agreement with consonance ratings reported in similar studies [e.g., Kameoka and Kuriyagawa, *J. Acoust. Soc. Am.* **45**, 1460–1469 (1969)]. The presence of noise decreased consonance ratings across most intervals. [Work supported by a NASA fellowship.]

11:30

4aPP12. Extended perceptual spaces for pitched and percussive timbres. Stephen Lakatos (Washington State Univ., 14204 NE Salmon Creek Ave., Vancouver, WA 98686, lakatos@vancouver.wsu.edu) and James Beauchamp (Univ. of Illinois at Urbana—Champaign, Urbana, IL 61801)

A series of listening tests attempted to better isolate higher-order perceptual dimensions of timbre by building on past findings showing that spectral centroid and rise time represent principal acoustic correlates of the primary timbral dimensions. Listeners with varying levels of musical training rated the timbral similarity of three sets of pitched and percussive instrument sounds that were equated for centroid and rise time using signal processing techniques. Multidimensional scaling analyses yielded two- and three-dimensional perceptual spaces for the three stimulus sets. Several aspects of the spectral fine structure of the timbres correlated with the dimensions of these spaces, but the nature of the acoustic correlate varied somewhat depending on the stimulus set in which the timbre was presented. The results suggest that additional perceptual dimensions of timbre exist, but that their precise acoustic correlates are context dependent and therefore less perceptually “primary” than centroid and rise time. [This research was funded by Air Force Office of Scientific Research Grant No. F49620-99-1-0293.]

Session 4aSA

Structural Acoustics and Vibration: Acoustics of Uncertain Structures

Joseph W. Dickey, Cochair

Center for NDE, Johns Hopkins University, 3400 North Charles Street, Baltimore, Maryland 21218

Alison B. Flatau, Cochair

Aerospace Engineering and Engineering Mechanics, Iowa State University, 1200 Howe Hall, Ames, Iowa 50011

Contributed Papers

8:00

4aSA1. A hybrid finite-element formulation for midfrequency analysis of systems with excitation applied on short members. Xi Zhao and Nickolas Vlahopoulos (Dept. of Naval Architecture and Marine Eng., Univ. of Michigan, 2600 Draper Rd., Ann Arbor, MI 48109-2145)

The hybrid finite-element analysis combines conventional finite-element analysis with energy finite-element analysis for midfrequency computations. A hybrid FEA formulation has been published for systems with excitation applied on long members. [N. Vlahopoulos and X. Zhao, "A Basic Development of a Hybrid Finite Element Method for Mid-Frequency Computations of Structural Vibrations," *AIAA J.* **37**(11), 1495–1505 (1999).] In this presentation, theoretical developments associated with applying the excitation on short members will be discussed. Validation cases will be presented for several systems by comparing analytical solutions to hybrid FEA results. It will be demonstrated that the resonant behavior of the short members and the interaction between long and short members are captured correctly by the hybrid FEA. [Work supported by the Automotive Research Center established at the University of Michigan by the US Army Tank Automotive Command.]

8:15

4aSA2. Investigation of power flow in the midfrequency range for systems of colinear beams based on a hybrid finite-element formulation. Xi Zhao and Nickolas Vlahopoulos (Dept. of Naval Architecture and Marine Eng., Univ. of Michigan, 2600 Draper Rd., Ann Arbor, MI 48109-2145)

A hybrid finite-element analysis (hybrid FEA) is employed for investigating power-flow characteristics for systems of colinear beams in the midfrequency range. The importance of capturing power reinjection and power reradiation effects in the solution is demonstrated. The dependency of the power-flow characteristics of a system in the midfrequency range on the rigidity, mass, and damping properties of its components is determined. Both the hybrid FEA and analytical solutions are employed for analyses in order to establish the viability of the hybrid FEA as a simulation technology in the midfrequency range. [Work supported by the Automotive Research Center established at the University of Michigan by the US Army Tank Automotive Command.]

8:30

4aSA3. Structural-acoustic analysis of 3-D curved shell with a discontinuity using analytical/numerical matching (ANM). Christopher Park, Linda Franzoni, and Donald Bliss (Duke Univ., Box 90300, Durham, NC 27708-0300)

Analytical/numerical matching (ANM) is used to determine the structural vibration of a curved shell driven at a single support. The ANM solution decomposes the problem into global, local, and matching sub-problems. The global problem addresses large-scale effects with structural discontinuities replaced by smooth distributed forces. The local problem models the rapidly changing region around a structural discontinuity. These constituent problems are solved independently by the most efficient method available. Here, the local problem is solved numerically using 3-D solid finite element analysis (FEA) and the matching problem is solved analytically, using the Love–Timoshenko shell equations. The global problem is decomposed modally in azimuth, leaving a corresponding FEA problem in the axial direction for each mode. Due to the smoothness of the global problem, relatively few modes are required for the modal decomposition and the affiliated FEA problem can be solved using a low resolution. An advantage of the ANM method is its ability to capture the detailed response of structures with geometric complexities without having to incorporate the complexity (via higher-order elements or high mesh densities) in the modeling of the large global problem. [Work sponsored by ONR.]

8:45

4aSA4. Vibration power transmission coefficients for the coupling of circular cylindrical shells to flat plates. Benjamin F. Willis and Courtney B. Burroughs (Grad. Prog. in Acoust., The Pennsylvania State Univ., P.O. Box 30, State College, PA 16804)

Bending and in-plane waves are coupled by curvature of structures and at junctions between structures. To estimate the effects of wave coupling on the transmission of vibrational power through junctions between curved and flat elastic structures, an analytical model of circular cylindrical shells coupled to flat plates along a circumference of the shell was developed. Donnell–Mushtari shell equations are used to model the infinite and semi-infinite cylindrical shells. Classical theory is used to model both bending waves and in-plane longitudinal and shear waves in the plates. Models include plates inside and outside the shell as well as continuous through the shell. All plates are perpendicular to the shell surface. Plates coupled to both semi-infinite shells, where the shell terminates at the junction with the plate, and to infinite shells are modeled. Power transmission coefficients are presented as a function of circumferential mode number and frequency for different types of incident waves either in the shell or in the plate, and for different shell/plate configurations.

4aSA5. A new algorithm for modal identification using mode isolation. Michael V. Drexel and Jerry H. Ginsberg (Woodruff School of Mech. Eng., Georgia Inst. of Technol., Atlanta, GA 30322-0405)

Multiple degree of freedom (MDOF) algorithms are the dominant methods for identifying modal properties from measured response data. Although there are numerous MDOF methods, all are based on the notion that because the response of a linear dynamic system is the sum of many modal contributions, the extraction technique must deal with all of the modal parameters in a simultaneous fashion. Thus, these methods do not exploit the fact that each mode has unique characteristics that can be used to isolate that mode from the contribution of the other remaining modes. The mode isolation method proposes an algorithm that extracts the modal parameters of each mode in a recursive search, and then refines the estimation of each mode by isolating its effect from the other modal contributions. The mode isolation method has an advantage over MDOF methods in that MDOF algorithms require an *a priori* guess of the number of significant modes, or degrees of freedom, of the system. An error in that guess can have serious consequences with respect to the accuracy of the extraction process. Several MDOF methods attempt to remedy this situation by utilizing peak counting to find the number of modes contained in the system. In contrast, the mode isolation method offers a simple and consistent modal extraction methodology that is inherently automatic in nature. A numerical example of the mode isolation method is presented for a system that was previously studied with the eigensystem realization algorithm (ERA) and enhanced ERA. Eigenvalues and mode shapes are compared for each algorithm. The results suggest that the mode isolation method is more robust in the treatment of noisy data.

9:15

4aSA6. Causal recovery of the nonminimum phase from the measured magnitude of one-dimensional acoustic reflections. Cory L. Clarke and J. Gregory McDaniel (Dept. of Aerosp. and Mech. Eng., Boston Univ., 110 Cummings St., Boston, MA 02215)

A new approach is presented for obtaining a reflection coefficient phase from magnitude using frequency-domain causality relations. This is useful because in many situations the reflection coefficient phase is more difficult to accurately measure than the magnitude; however, a knowledge of the phase is often vital to understanding the dynamics of the reflecting object. If a transfer function is minimum phase, then causality requires that its magnitude and phase are related by the Hilbert transform. Such relations do not exist for nonminimum phase-transfer functions. Previous work found that reflection coefficients are not necessarily minimum phase [McDaniel, J. Acoust. Soc. Am. **105**, 2710–2716 (1999)]; however, the drive-point impedance of a passive reflecting object is. For one-dimensional systems, the drive-point impedance of the reflecting object is algebraically related to the reflection coefficient. The approach combines this relationship with causality relations for the impedance to find the complex reflection coefficient given only its magnitude. It is applied to analytic data for a one-dimensional structural acoustic system so that the actual phase may be compared to that predicted by the method. Extensions of the approach to multidimensional scattering will be discussed. [Work supported by ONR.]

9:30

4aSA7. Determining impact source location in structural networks. Joseph Dickey (Ctr. for Nondestruct. Eval., Johns Hopkins Univ., 3400 N. Charles St., Baltimore, MD 21218) and Gideon Maidanik (Taylor Model Basin, Carderock, MD 20817)

An impact, or similar event, generates a transient in a structural network consisting of connected one-dimensional systems. A sensor, attached at an arbitrary point in the network, will “hear” a complicated time series of pulses. A few of these pulses are recorded and used to generate a time-reversed series which is injected at the sensor location. If the network

is stationary in time, the impulse response function (IRF) can be used to show that the introduced series will constructively interfere at the original impact point and this location can thereby be determined. The procedure is demonstrated computationally using a time-domain IRF. The relationship between the above and “reciprocity” and “inverse filtering” will be discussed.

9:45

4aSA8. The influence of substructure modeling on the predicted vibroacoustic response of a fluid-loaded plate system. W. Steve Shepard, Jr. (Mech. Eng. Dept., The Univ. of Alabama, Box 870276, Tuscaloosa, AL 35487, sshepard@coe.eng.ua.edu) and Kenneth A. Cunefare (Georgia Inst. of Technol., Atlanta, GA 30332)

When predicting the vibroacoustic behavior of complex systems, one must decide how much effort to spend on modeling the features with dimensions smaller than those of the primary structure. To investigate these issues, this work examines changes in the vibroacoustic response of a semi-infinite fluid-loaded plate due to variations in modeling details associated with an attached substructure. The substructure consists of a smaller plate supported by springs along each edge. To examine the impact of spring scales, the manner in which the elasticity is distributed over the main plate is varied. Substructure modeling issues are examined by varying the number of degrees-of-freedom allowed in the substructure model. Both the combined system response and acoustic radiation from the main plate are computed using the Acoustic Surface Variational Principle and Hamilton’s Principle. Furthermore, sensitivity relationships that express changes in the system response to changes in the scale of the spring elements are presented. For the cases considered, it is shown that details associated with the scale of the spring element are only important for frequencies near or below the resonances of the isolated subsystem. Furthermore, only the dynamics of the substructure that includes rigid-body-type motions is important. [Work supported by Georgia Tech.]

10:00–10:15 Break

10:15

4aSA9. Improved power input methods in SEA modeling. Robert C. Haberman (BBN Systems and Technologies, Union Station, New London, CT 06320, rhaberman@bbn.com)

The traditional method for determining SEA input power to structures subjected to point-force excitation is to consider the directly driven structural element as infinite. This allows for a relatively simple calculation of input power. For example, the power input to three-dimensional frame structures is to treat the driven beam as infinite and use the corresponding real part of the infinite beam mobility along with the mean-square force. Similarly, the power input to a structure consisting of interconnected plates is to likewise consider the driven plate as infinite and use the corresponding real part of the infinite plate mobility and force. Errors are introduced by this method when the drive point approaches the boundaries (in terms of flexural wavelengths) of interconnected structural elements. At these locations, the power is influenced by near-field effects and the transmission of power into the connected structure. To reduce these errors, a method based on modeling the driven member and directly connected structural elements is investigated. All the members are modeled as semi-infinite. Examples are given of interconnected beam and plate structures, and guidelines are introduced for improved SEA power input calculations.

4aSA10. Sound transmission through finite laminated composite panels using statistical energy analysis. Avinash R. Patil and Malcolm J. Crocker (Mech. Eng. Dept., 201 Ross Hall, Auburn Univ., Auburn, AL 36849)

Laminated composite panels are widely used in aerospace structures and other applications where weight reduction of the structure is desired without compromising the strength. They can be made up of a number of layers of unidirectional laminas. Sound radiation and sound transmission properties of these panels are different than those for isotropic panels. This is because of the variation of the material properties along the direction of bending wave propagation. During this work, analysis of those finite laminated composite panels which exhibit uncoupling of bending/twisting phenomenon is done. A bending wave number diagram is drawn to study the variation of bending wave number with respect to the angle of bending wave propagation for a given frequency. Instead of a single critical frequency as in the case of an isotropic panel, a band of critical frequencies is observed. In this band of critical frequencies, coincidence occurs only at a certain angle of bending wave propagation for a particular frequency. Frequency average radiation efficiency of such a panel is predicted by considering the variation of wave number along different directions of bending wave propagation. Modal density is predicted from the bending wave number diagram. Transmission loss of the panel is predicted using the statistical energy method.

4aSA11. Partitioning, substructuring, and imposition of impedance: Are they the same? John J. McCoy (The Catholic Univ. of America, Washington, DC 20064)

Formulation partitioning, substructuring, and the imposition of an impedance operator, or DtN map, all refer to solution methods with the same purpose. This is to distinguish those coordinates necessary for describing the response of a structure, which are not directly forced by an external agent, and to eliminate these as primary variables in a reduced formulation. It is increasingly common to encounter use of the three terms interchangeably. This raises the question. Are the solution methods identical? Fundamental differences are identified and the significance of these differences is discussed.

4aSA12. Probabilistic and possibilistic models of uncertainty in structural dynamics. Dexter F. Johnson and Robin S. Langley (Dept. of Eng., Univ. of Cambridge, Cambridge CB2 1PZ, UK)

The prediction of structural reliability under static loading has received much attention in the field of civil and structural engineering, and well-established methods such as FORM and SORM are available to estimate the failure probability. If the quality and quantity of the available data do not justify a probabilistic analysis, then the reliability can alternatively be assessed by using less precise "possibilistic" methods such as interval analysis or convex modeling. None of these probabilistic or possibilistic methods has as yet been applied extensively to noise and vibration problems, and the present work considers this issue. Initially it is shown that the probabilistic and possibilistic algorithms can all be expressed in the form of a constrained optimization problem, and the methods are then applied to a vibration problem involving a beam with an uncertain mass distribution. The beam is modeled both experimentally and theoretically to highlight various types of uncertainty that can arise: in addition to "controlled" uncertainties involving adjustable lumped mass positions, there are "uncontrolled" uncertainties arising from items such as boundary conditions and the support stiffnesses. The relative merits of the various reliability assessment techniques are discussed in light of this example.

4aSA13. The vibration of uncertain structures: Random matrix theory and non-Poisson statistical models of the system's natural frequencies. Robin S. Langley (Dept. of Eng., Univ. of Cambridge, Cambridge CB2 1PZ, UK)

At medium to high frequencies, the noise and vibration levels in an engineering structure can be sensitive to small changes in the properties of the system. This means that the response of a set of nominally identical units (for example, automobiles) from a production line can be radically different, and knowledge of the response statistics would significantly aid quality assurance and control. Unfortunately, the prediction of the noise and vibration response statistics for a complex built-up random structure is an extremely difficult task, as is the associated problem of determining the statistics of the natural frequencies and mode shapes. In recent years it has been suggested that progress might be made by employing the techniques of random matrix theory previously developed for applications in nuclear physics. In particular, the Gaussian Orthogonal Ensemble (GOE) has been shown to yield good agreement with the statistics of the natural frequencies of various metal blocks. In this paper the applicability or otherwise of random matrix theory to realistic engineering structures is considered, and a comparison is made with other approaches, such as a non-Poisson point process model of the natural frequencies. Several examples are considered to illustrate the theoretical developments.

4aSA14. Le Chatelier's Principle in noise control. G. Maidanik (Carderock Div., Naval Surface Warfare Ctr., 9500 MacArthur Blvd., West Bethesda, MD 20817-5700)

Le Chatelier's Principle is perhaps better known to those who studied chemistry at the university. Therefore, I find that few in the noise control community are familiar with this principle. This paper features some examples of noise control problems in which "the Principle" plays a role. For noise control purposes and for the record, Le Chatelier's Principle states that "when a dynamic system is modified with the intention of achieving a high degree of change in the noise issued by this dynamic system, the achievement is mitigated by the implementation of the modification."

4aSA15. When is a "fuzzy" not a fuzzy (continued)? M. Strasberg (Code 702, David Taylor Model Basin, NSWCDD, West Bethesda, MD 20817-5700)

The term "fuzzy substructures" as originally coined by C. Soize includes four categories of substructures whose influence on the vibratory behavior of the master structure to which they are attached may be significantly different. Either (1) the substructures are sufficiently numerous and their modal antiresonant frequencies sufficiently close to each other to result in modal overlap over the frequency range of interest; or (2) they do not have modal overlap; also, either (a) successive modal frequencies are separated by a constant frequency increment, or (b) the frequency increments are irregular. The effect of category (1) substructures on the master structure is independent of the detailed spacing between modal frequencies, provided modal overlap exists, for both steady-state and transient excitation. Contrariwise, the effect of category (2) substructures on the transient behavior of the master structure does depend on whether or not the modal frequencies are equally spaced. Although the statistical average behavior of an ensemble of such (2b) substructures may be similar to that of category (1), individual members of the ensemble may behave quite differently from their average. The differences will be exhibited by examples. To avoid confusion, it is suggested that the term "fuzzy structures" be limited to substructures with modal overlap.

Session 4aSC

Speech Communication: Hearing Impairment and Audio-Visual Cues in Speech (Poster Session)

Lynne E. Bernstein, Chair

House Ear Institute, 2100 West Third Street, Los Angeles, California 90057

Contributed Papers

All posters will be on display from 9:00 a.m. to 11:30 a.m. To allow contributors an opportunity to see other posters, contributors of odd-numbered papers will be at their posters from 9:00 a.m. to 10:15 a.m. and contributors of even-numbered papers will be at their posters from 10:15 a.m. to 11:30 a.m.

4aSC1. Combining acoustic and electric hearing: Simulations and real-patient results. Christopher W. Turner, Bruce J. Gantz, Sue A. Karsten, and Bom Jun Kwon (Univ. of Iowa, Iowa City, IA 52242)

For patients with severe high-frequency sensorineural hearing loss, providing audible high-frequency, acoustic speech information often results in no increase, or in some cases a decrease, in speech recognition performance. One approach to this problem is to transmit the high-frequency speech information via a multichannel cochlear implant located in the base of the cochlea, while allowing the patient to continue to use the remaining low-frequency acoustic hearing. Data from simulation experiments are presented to show the amount of speech information that can be transmitted via this combination. The simulation of electric hearing was accomplished by using multiple channels of narrow-band noise modulated by the high-frequency speech signal. In general, speech recognition scores have the potential to be quite high if low-frequency (acoustic) information (below 500 or 1000 Hz) is combined with electrical stimulation for frequencies of 3000 Hz and above. An adverse effect of a mismatch between the speech frequencies and carrier band frequencies was observed, similar to the effect of assigning "wrong" frequencies to the cochlear implant. Some promising results from several human patients who have been implanted with this device are also shown. [Work supported by NIDCD.]

4aSC2. Bone conduction transmission and head-shadow effects for unilateral hearing losses fit with transcranial CIC hearing aids. Marc A. Fagelson (Dept. of Commun. Disord., East Tennessee State Univ., Johnson City, TN 37614), Colleen Noe, Jennifer Blevins, and Owen Murnane (James H. Quillen VAMC, Mountain Home, TN 37694)

Bone conduction transmission and head-shadow effects were determined with transcranial completely-in-the-canal (TCCIC) CROS hearing aids. Five subjects with documented profound unilateral hearing loss and experience with traditional CROS/BICROS fittings (TCROS) were tested with a CIC hearing aid placed in their poorer ear. Peak SPL was measured at the tympanic membrane and ranged from 105–115 dB SPL at 2000 Hz. Pure-tone crossover thresholds and functional gain tested at frequencies from 250–8000 Hz varied considerably more than the SPL measures. The pure-tone results indicated that sensitivity in the better ear was moderately associated with functional gain across frequency. Speech recognition was then tested in the sound field in two conditions: direct (noise in the poorer ear, speech in the better ear) and indirect (noise in the better ear, speech in the poorer ear) at S/Ns of -6, 0, +6, +12, and quiet. The TCCIC fittings were more effective than TCROS aids across S/Ns, particularly in the direct condition. In the indirect condition, the two fittings performed similarly. When data were pooled across conditions, the TCCIC aids provided better word recognition than the TCROS aids, particularly for those subjects with greater sensitivity in the better ear.

4aSC3. Phoneme recognition before and after hearing aid fitting. Amy T. Neel and Ayaskanta Rout (Dept. of Audiol. and Speech Sci., Purdue Univ., 1353 Heavilon Hall, West Lafayette, IN 47907-1353, atneel@purdue.edu)

Many listeners with sensorineural hearing loss experience only modest increases in speech recognition after receiving their hearing aids and very small increases in hearing aid benefit up to one year after hearing aid fitting. In this study, phoneme recognition patterns before and after hearing aid fitting were examined to determine which phonemes improved most with amplification and which phonemes continued to improve with hearing aid experience. Preliminary results using a nonsense syllable recognition test revealed that voiceless stops and fricatives improved the most immediately after hearing aid fitting although accuracy remained below 50% for these phonemes. Aided recognition for these phonemes, however, improved very little from one month to one year postfitting. Variability in phoneme recognition before and after hearing aid fitting across subjects was large. Phoneme recognition results from an ongoing study of first-time hearing aid users with comparable levels of signal processing will be presented. Differences in phoneme recognition before and after hearing aid fitting can help explain differences in hearing aid benefit across hearing aid users. In addition, phoneme recognition profiles can be used to design individualized, efficient auditory training programs for new hearing aid users.

4aSC4. The vowel spaces of normal-hearing listeners and patients with cochlear implants. James D. Harnsberger, David B. Pisoni (Speech Res. Lab., Dept. of Psych., Indiana Univ., Bloomington, IN 47405), Mario A. Svirsky, Adam R. Kaiser (Indiana Univ. School of Medicine, Indianapolis, IN), and Richard Wright (Univ. of Washington, Seattle, WA)

Cochlear implant (CI) users show substantial individual differences in their ability to understand speech in general, and vowels in particular. These differences may result from widely different abilities in identifying formant frequencies or in adapting to the more basal than normal spectral information presented by the implant. In this study, we administered a vowel perception test, using a method-of-adjustment (MOA) paradigm, to 8 CI users and 43 normal-hearing listeners. The MOA vowel test consisted of 330 steady-state synthetic vowel stimuli, varying in $F1$ and $F2$, arranged in a visual two-dimensional grid. Subjects were asked to label and rate on a 7-point scale those stimuli that matched the vowels contained in ten visually-presented words, "heed," "hid," "aid," "head," "had," "hut," "odd," "whod," "hood," "owed," and "odd." Plots of subjects' responses for all ten words constituted the vowel spaces of the subjects. With one exception, no systematic shift was observed across all vowel categories of CI users, suggesting that these subjects were able to

adapt completely to the spectral shift introduced by the implant. However, the CI users' spaces differed substantially from normal vowel spaces in terms of the relative size of the vowel categories and their location in perceptual space.

4aSC5. Perception of prosodic phrasing by hard-of-hearing listeners.

Dragana Barac-Cikoja, Sally Revoile, and Michelle Waters (Gallaudet Univ., 800 Florida Ave. NE, Washington, DC 20002, Dragana.Barac-Cikoja@gallaudet.edu)

Severely to profoundly hard-of-hearing (HOH) listeners ($n=12$) were tested for their perception of prosodic phrase continuity in clear and conversational speech. In two experiments, fluently spoken sentences were compared with utterances artificially assembled from isolated words (experiment 1) and prosodic phrases (experiment 2) that replicated the sentences' lexical composition. The difference in the overall duration of the assembled and fluent utterance was eliminated via signal processing, but the amplitude envelope and the intonational contour differences were left intact. Randomized repetitions of each utterance pair (72 pairs, 12 sentences \times 2 speaking styles \times 3 speakers, in experiment 1, and 40 pairs, 10 sentences \times 2 speaking styles \times 2 speakers, in experiment 2) were presented to each listener over several experimental sessions. In a 2IFC procedure, the listener identified the utterance that sounded more like fluent speech. Four normal-hearing (NH) controls were tested in white noise (S/N ratio = -20 dB). NH listeners identified the fluent sentences with accuracy greater than 90%. Across HOH listeners, identification accuracy varied from completely random to less than 75% correct, and was dependent on a listener's sensation level and preferred communication mode (oral versus manual). [Work supported by NIDCD.]

4aSC6. Auditory-visual context effects in a speech and nonspeech condition.

Linda W. Norrix, Iris Oved (Dept. of Psych., Univ. of Arizona, Tucson, AZ 85721), and Lawrence D. Rosenblum (Univ. of California, Riverside, CA 92521)

Perception of an auditory continuum can be changed by a preceding visual context [Norrix and Green, in Proc. AVSP'99, Santa Cruz, CA, August, 1999]. This study asked whether this perceptual shift requires that listeners experience the stimuli as speech. A sine-wave continuum (/ara-ala/) was presented to three groups of trained participants for identification (auditory-only condition). Tokens from the continuum were also paired with a point-light display of a talker saying /aba/ (auditory-visual condition) and presented to the same listeners for identification. Observers in group one (speech) identified the sounds as containing an /r/ or /l/ and reported after the experiment that they heard the sounds as speech. Group two, also instructed to identify the speech sounds as containing /r/ or /l/, reported they did not hear the sounds as speech. Group three (nonspeech) was instructed to identify the environmental sounds as most similar to the "first" or "last" of the continuum. Results indicated a reliable shift for the auditory-visual compared to the auditory-only identification function only in the speech group, suggesting that auditory-visual context effects might depend on observers interpreting the stimuli as speech. [Work supported by NSF Grant #SBR9809013 awarded to Kerry P. Green (deceased).]

4aSC7. Effects of auditory and visual stimuli on motor facilitation of speech muscles.

Aravind, N. K. (Dept. of Speech Pathol., Univ. of Toronto, Tanz Neurosci. Bldg., 6 Queens Park Crescent W., Toronto, ON M5S 3H2, Canada, a.namasivayam@utoronto.ca), Megha Sundara (McGill Univ., Montreal, QC H3G 1A8, Canada), and Robert Chen (Toronto Western Hospital, Toronto, ON M5T 2S8, Canada)

Transcranial magnetic stimulation (TMS) was applied to the left motor cortex during presentation of video and audio speech stimuli. Motor evoked potentials (MEPs) were recorded from the right orbicularis oris muscle which is activated during pronunciation of the consonant /ba/ but

not the consonant /ta/. Subjects were instructed to pay attention to six types of stimuli: video presentation of the consonant /ba/, audio presentation of /ba/, audio and video presentation of /ba/, video presentation of /ta/, audio presentation of /ba/ with time-locked video presentation of video /ta/, and the face of the speaker when he was silent. The MEPs were enhanced only when subjects were visually presented with the consonant /ba/. Auditory presentations of the same consonant failed to elicit similar enhancement. The results from visual observation of the consonant /ba/ support other studies of motor facilitation in limb control research in monkeys and humans. In these studies observation of a motor action present in the repertoire of the animal results in enhanced MEPs representing automatic, covert retrieval of the matching action in the premotor cortex. This retrieval is, however, modality dependent. Auditorily presented signals do not trigger action retrieval from the premotor cortex in an identical way.

4aSC8. Development of a facility for simultaneous recordings of acoustic, optical (3-D motion and video), and physiological speech data.

Lynne E. Bernstein, Edward T. Auer, Jr., Brian Chaney (Dept. of Commun. Neurosci., House Ear Inst., 2100 W. Third St., Los Angeles, CA 90057, lbernstein@hei.org), Abeer Alwan, and Patricia A. Keating (Univ. of California—Los Angeles, Los Angeles, CA 90095)

A multidisciplinary, multilaboratory project is underway whose focus is optical and acoustic phonetic signals and their relationships to each other in speech production and perception. The goals are to quantitatively characterize optical speech signals, examine how optical phonetic characteristics relate to acoustic and to physiologic speech production characteristics, study what affects the intelligibility of optical speech signals, and apply obtained knowledge to optical speech synthesis and automatic speech recognition. We describe the data acquisition facility, which includes recording of video, acoustic, electromagnetic midsagittal articulography, and 3-D motion capture data simultaneously and in synchrony. We also outline the design of the database of recordings being obtained and show examples from the database. [Work supported by NSF 9996088.]

4aSC9. Confidence ratings in auditory-visual speech perception.

Bart R. Clement, Sarah K. Erickson, Su-Hyun Jin, and Arlene E. Carney (Dept. of Commun. Disord., Univ. of Minnesota, 164 Pillsbury Dr. SE, Minneapolis, MN 55455)

In an earlier report, Carney *et al.* [J. Acoust. Soc. Am. **106**, 2270 (1999)] showed that the perception of the McGurk effect, the result of a visual syllable biasing the perception of a simultaneously presented auditory syllable, was affected by the use of multiple talkers. In addition, individual listeners could be strongly auditory in their perception, even when a visually biasing stimulus was present. During the original experiment, listeners were asked to rate each of their perceptions in the auditory-only (AO), visual-only (VO), and auditory-visual (AV) modes for syllabic stimuli. Listeners were very confident of their perceptions for bilabial stimuli in the VO mode but not for alveolar or velar stimuli. Confidence ratings were also high for all AO stimuli and for AV stimuli for matching visual-auditory tokens. However, confidence ratings varied for tokens with mismatched A and V stimuli. Some individual listeners were always confident in their perception regardless of speaker or condition. Other listeners varied in confidence with speaker and condition.

Across listeners, certain talkers evoked different ratings of confidence as well. Results support the notion of a graded perception of bimodal stimuli. [Work supported by NIDCD.]

4aSC10. Recognizing words from optical phonetic signals. Sven Mattys, Lynne E. Bernstein, and Edward T. Auer, Jr. (Dept. of Commun. Neurosci., House Ear Inst., 2100 W. Third St., Los Angeles, CA 90057, smattys@hei.org)

Lipreading is thought to be extremely difficult, and the difficulty is attributed to lack of perceptual distinctiveness among most visible spoken words. Therefore, it is thought that accurate lipreading requires reliance on postlexical processes, such as guessing. However, we [E. T. Auer and L. E. Bernstein, *J. Acoust. Soc. Am.* **102**, 3704–3710 (1997)] predicted, based on computational modeling of the lexicon, that many words maintain their visual distinctiveness. In this study, we took only visual phoneme confusability and word frequency into account to predict the relative intelligibility of lipread words. We predicted that words without visual lexical competitors can be perceived as readily as heard words, and a systematic change in intelligibility would be observed as a function of word frequency and number of competitors. Predictions of near-100% accuracy for many words were confirmed in a perceptual word identification experiment with skilled deaf American lipreaders. Accuracy was not influenced by word length. When phoneme confusability in words was statistically controlled, the number of lexical competitors correlated significantly with word score, as did word frequency. These results will be compared with those of hearing participants, and the implications discussed with attention to bottom-up processes versus lexical competition in spoken word recognition. [Work supported by NIH/NIDCD 02107.]

4aSC11. On the visual distinctiveness of English words. Margaret MacEachern (Dept. of Linguist., Univ. of Pittsburgh, Pittsburgh, PA 15260)

Lipreading is facilitated when words are visually distinct. It has been established [Auer and Bernstein, *J. Acoust. Soc. Am.* **102**, 3704–3710 (1997)] that words in the English lexicon manifest a low degree of visual similarity: most words remain visually unique, even when distinctions are collapsed across phoneme classes, as is characteristic of visually perceived speech. But, is the high degree of visual uniqueness a distinctive (nonaccidental) property of the English lexicon? Perhaps not: the lexical space of English is large, both in crosslinguistic terms and in absolute terms. In other words, many lexical slots exist; a fairly low percentage of them are filled by actual words [Hockett, 288–290 (1958)]. Therefore, the high degree of visual distinctiveness might be an “accidental” property of the lexicon, deriving from the segment inventory and phonotactics of the language. This study will explore this question by comparing the visual similarity properties of words in simulated (hypothetical) lexicons to those of words in the actual English lexicon. Each novel lexicon will be generated on the basis of the phonotactic patterns of the actual lexicon. The modeled lexicon is PhLex, a 30 000-word phonologically transformable lexicon of American English [Seitz, Bernstein, and Auer (1995)].

4aSC12. Phonetic structure is similar across auditory and vibrotactile speech perception. Christopher T. Kello and Lynne E. Bernstein (Dept. of Commun. Neurosci., House Ear Inst., 2100 W. Third St., Los Angeles, CA 90057, ckello@hei.org)

Vibrotactile vocoders have long been viewed as a possible method for conveying speech information to deaf perceivers. However, modest levels of speech transmission have led to numerous alternative notions concerning limitations of vibrotactile devices. This study investigated whether perceived acoustic phonetic structure is preserved in the transformation of the *F*₂ speech region via a linear vibrotactile vocoder array. In experiment

1, naive hearing subjects discriminated pairs of vibrotactile word stimuli immediately before and after 2 h of practice. At both times, relative vibrotactile discrimination levels were well predicted by identification of phonemes presented by an analogous *F*₂ acoustic vocoder, suggesting that structure is preserved. In experiment 2, subjects separately rated auditory and vibrotactile consonant similarity of stimuli presented by auditory and vibrotactile vocoders, respectively. Analysis of phonological feature information showed similarities and differences in perceptual structure across modalities. Vibrotactile stimuli were rated to be more similar to each other than were auditory stimuli, suggesting that phonetic information was degraded relative to auditory information. Based on these results, current investigations are directed at methods for inducing perceptual learning of vibrotactile phonetic information. [Work supported by NIH/NIDCD 00695.]

4aSC13. Sources of variability for talker and listener effects in auditory–visual speech perception. Su-Hyun Jin and Arlene E. Carney (Dept. of Commun. Disord., Univ. of Minnesota, 164 Pillsbury Dr. SE, Minneapolis, MN 55455)

Carney *et al.* [*J. Acoust. Soc. Am.* **106**, 2270 (1999)] reported on talker and listener effects for auditory–visual speech perception for syllabic stimuli that were disparate [a visual /gi/ paired with an auditory /bi/ or a visual /bi/ paired with an auditory /gi/]. Listeners varied in the ratios of fused and combination responses (the McGurk effect) to auditory responses, depending upon the talker and their own most frequently observed mode of perception. The purpose of this experiment was to determine if listener and talker variability could be manipulated by recombining talkers’ voices with different talkers’ faces. Green *et al.* [Percept. Psychophys. **50**, 524–536] had demonstrated that the McGurk effect was observed even when voices and faces did not match in gender. The same 11 talkers (5 male and 6 female) were used from the Carney *et al.* experiment. Twenty-four listeners heard the combination of all male talkers with each other and all female talkers with each other. Similar patterns of listener and talker variability were observed even when voices and faces did not match, supporting the earlier conclusion that the McGurk effect is a graded phenomenon, depending upon listener and talker. [Work supported by NIDCD.]

4aSC14. The effects of audiovisual speech on sentential phoneme detection. Ethan A. Cox (Dept. of Psych., Univ. of Arizona, Tucson, AZ 85721, ecox@u.arizona.edu)

Many previous studies in audiovisual speech perception concentrate on situations where auditory information is degraded, or stimuli with conflicting auditory and visual information. However, relatively little work has addressed the issue of how the availability of visual information contributes to postphonetic levels of processing. Presented here are results from a series of studies targeting audiovisual processing in lexical and sentential structures. Subjects in a phoneme-monitoring task for target words in sentences showed significantly faster RTs for audiovisual conditions as compared to auditory-only conditions. Additionally, a lexical frequency effect was observed in the auditory-only, but not the audiovisual condition. These initial findings focused on the bilabial viseme group and were compatible with an interpretation based on the earlier temporal availability of the visual information as compared to the auditory. Subsequent experimentation with the same general procedures incorporated catch trials with nontarget bilabials, and indicates a more complex account. Comparison of the original findings with other viseme groups also supports the view that the availability of visual information contributes more than a simple temporally earlier trigger for target identification. Monitoring performance cannot be well accounted for by treating the auditory and visual information as separate channels for controlling the response.

Session 4aUW

Underwater Acoustics, Signal Processing in Acoustics and Acoustical Oceanography: Model-Based Processing of Sources in Motion I

Peter G. Cable, Cochair

BBN Technologies, Union Station, New London, Connecticut 06320

William M. Carey, Cochair

*Aerospace and Mechanical Engineering Department, Boston University, 110 Cummington Street,
Boston, Massachusetts 02215*

Chair's Introduction—8:25

Invited Papers

8:30

4aUW1. Inference of source–receiver motion from phase measurements in the Modal Mapping Experiment. George V. Frisk (Dept. of Appl. Ocean Phys. and Eng., Woods Hole Oceanogr. Inst., Woods Hole, MA 02543)

In March 1997, the Modal Mapping Experiment (MOMAX) was conducted aboard the R/V Endeavor in about 70 m of water off the New Jersey coast. Three drifting MOMAX buoys, each containing a hydrophone, GPS navigation, and radio telemetry received signals up to ranges of 10 km from a NUWC J15-3 source suspended from the moving ship and transmitting pure tones at 50, 75, 125, and 175 Hz. A striking feature of the data is the remarkable stability and regularity of the phase, even though the magnitude displays a complex multimodal interference pattern. This phase behavior occurs even at the higher frequencies and higher source/receiver speeds (up to 3 kts) measured in the experiment. A phase model is developed which indicates that the leading-order behavior of the time rate-of-change of the phase is simply equal to the product of a typical wave number in the water column and the source-receiver speed. This model accurately predicts the source-receiver speed from the phase measurements obtained in MOMAX. The implications of this model for other scenarios, such as long-range, deep-water situations are also discussed. [Work supported by ONR.]

8:50

4aUW2. Synthetic aperture processing to extract ocean bottom acoustic properties. Subramaniam D. Rajan (Science Solutions, Inc., 845 106th Ave. NE, Ste. 200, Bellevue, WA 98004)

A review of an approach to determine ocean bottom acoustic properties in a shallow-water environment from data acquired on a synthetic aperture horizontal array will be presented. The performance of this method will be demonstrated using data from a number of field experiments. The feasibility of using this approach for rapid environment assessment will also be discussed.

9:10

4aUW3. Model-based and data-based processing of moving sources in a shallow water channel. Lisa M. Zurk and James Ward (MIT Lincoln Lab., 77 Massachusetts Ave., Cambridge, MA 02139)

Adaptive MFP algorithms are applied to passively detect and localize a moving source while mitigating interference from surface shipping. A model-based motion compensation algorithm is applied to prevent smearing loss from target motion and extend the observation interval for adaptation. Moving interference is removed by applying a time-varying spatial filter at each data snapshot. The time-variation of the filter decreases the size of the instantaneous interference subspace and allows for a larger signal subspace. The interference subspace is formed by one of three methods. In the first method, knowledge of the ship position is combined with a propagation model to predict the time-varying acoustic interference. A sector null is formed to protect against mismatch degradation. In the second method, a data decomposition is used to determine the strong interferers. The third method combines model-based and data-based approaches. For each method, interference rejection is quantified by presenting SINR values for the candidate algorithm and comparing to conventional approaches. The results use data obtained from the Santa Barbara Channel Experiment. [This work was sponsored by DARPA under Air Force Contract F19628-95-C00022E. Opinions, interpretations, conclusions, and recommendations are those of the author and are not necessarily endorsed by the United States Air Force.]

9:30

4aUW4. The effects of source motion on the performance of matched field processors. Ahmad T. Abawi, Newell O. Booth, Phil Schey (SPAWAR Systems Ctr., 53560 Hull St., San Diego, CA 92152-5001, abawi@nosc.mil), and W. S. Hodgkiss (Scripps Inst. of Oceanogr.)

To obtain a reliable and accurate estimate of the cross-spectral matrix, it is necessary to average individual matrices over a number of snapshots. However, an upper limit in the number of snapshots that can be used to average the cross-spectral matrix is imposed by source and interferer motion. Averaging while the target moves through resolution cells introduces signal gain degradation, spreading

the beam-former peak response through the resolutions cells occupied by the source. Averaging while interferences move smears sidelobes, increasing the number of eigenvalues in the cross-spectral matrix and reducing the effectiveness of adaptive noise cancellation processes. During the May 1996 Shallow Water evaluation cell Experiment (SwellEx-96), which was performed in shallow water over a relatively flat bottom off the coast of San Diego, data was obtained with towed sources and interfering surface ships moving at many different rates. Data were recorded on several arrays including two 118-m 64-phone vertical array, one tilted by 45 degrees. This paper presents an analysis of source and interferer motion effects using SwellEx-96 data. Signal gain degradation, target and side-lobe smearing, and adaptive noise cancellation effectiveness are examined as a function of integration time, and number of elements used in the processing.

9:50

4aUW5. A hybrid plane-wave beamforming/matched-field processing, track-before-detect technique for detection of broadband moving sources with horizontal line arrays. Paul A. Baxley and Randall Brannan (SPAWAR Systems Ctr. San Diego, 53560 Hull St., San Diego, CA 92152-5001, baxley@spawar.navy.mil)

A hybrid plane-wave beamforming/matched-field processing (MFP), track-before-detect (TBD) technique for the automated detection and tracking of broadband moving sources in shallow water is presented. The approach, applicable to bottom-mounted horizontal line arrays (HLAs), provides estimates of the full track parameters (range, depth, course, and speed) with a processing load substantially lower than that required for a fully three-dimensional MFP search. The method begins with the production of a bearing-time record (BTR) for the HLA. By applying a TBD algorithm to the BTR, a detection is made and track bearing as a function of time is estimated with some bias and uncertainty. The MFP range-depth ambiguity surfaces are then produced for an envelope of bearings bracketing the bearing estimate uncertainty at each time. The TBD processing of these surfaces, involving a search over tracks defined by all combinations of bearings in the envelopes at the beginning and end of candidate tracks, allows for an estimation of the full track parameters as well as a reduction of bearing uncertainty and bias. Results of the application of this technique to broadband source-tow data in the Black Sea are presented. [Work supported by the ONR 321US.]

10:10–10:30 Break

10:30

4aUW6. Broadband shallow-water sea-test results with a towed vertical aperture array. John P. Ianniello and John M. Tattersall (Naval Underwater Systems Ctr., 1176 Howell St., Newport, RI 02841)

We review matched-field processing (MFP) sea test results obtained using a towed vertical aperture array. The array consisted of five horizontal lines towed in a vertical plane, nominally one over the other, with an approximately 5-m separation between the lines. We review results from two tests, one conducted in the Gulf of Mexico in about 120 m of water, and the other conducted southwest of Key West, FL in about 600 m of water. The vertical aperture array was towed on straight-line courses, at various ranges, parallel to a towed sound source. The radiated signal was a Gaussian pseudorandom waveform in the 100–600-Hz band. Results using replica correlation techniques are first used to study the channel transfer function; next, MFP results obtained by treating the signal as a random process are shown. These results demonstrate accurate range and depth localizations from 5 to 20 km as well as the ability to resolve a loud surface interferer from a quieter source at depth which is on the same bearing. These test results demonstrate the inherent robustness of broadband MFP. [Work supported by ONR 321.]

10:50

4aUW7. Model-based towed array processor: Experimental results. Edmund J. Sullivan (Naval Undersea Warfare Ctr., 1176 Howell St., Newport, RI 02841), Dieter Brecht (Federal Armed Forces Underwater Acoust. and Marine Geophys. Res. Inst., Germany), and Leif Persson (Natl. Defence Res. Establishment, Stockholm, Sweden)

Results of a towed array bearing estimation experiment using a model-based processor are presented. The experiment took place in the southern Baltic using an Atlas towed array. The towship was the Swedish corvette GAVLE. Results for the case of a 121-Hz source are presented. The towship ran in a straight line, past the source, where the distance of closest approach was 500 m. A short overview of the model-based approach, which is based on a state space formulation using a Kalman estimator, will be presented. By formulating the problem in state space form, it is possible to include sophisticated models in the processor. This, in effect, is a self-consistent means of including *a priori* information into the processing. It will be shown that the inclusion of the array's forward motion provides significant improvement in performance over the conventional processor in terms of the variance of the estimate. The improvement in performance arises from the bearing information contained in the forward motion of the array. This information obtains, from the time scale, change in the signal or, in the narrow-band case, the Doppler. The experimental results clearly show the improvement over the conventional processor.

11:10

4aUW8. Imaging a moving target using forward-look synthetic aperture sonar. Geoffrey S. Edelson (Sanders, A Lockheed Martin Co., Adv. Systems & Technol., P.O. Box 868, Nashua, NH 03061-0868, geoffrey.s.edelson@lmco.com), Charles J. Gedney, Philip A. Abbot, Ira Dyer (OASIS, Inc., Lexington, MA 02173), and Kenneth D. Rolt (Sonetech Corp., Bedford, NH 03110)

Mid-frequency active, surface ship sonar data from a recent shallow-water sea trial have been processed using synthetic aperture sonar (SAS) algorithms. The target of opportunity closed head-on directly toward the surface ship at a relative speed of 13 knots through an acoustic environment that was characterized by a downward refracting sound speed profile in 100-m-depth water over a sand/mud bottom during calm seas. This situation represents an endfire (90-degree squint angle) scenario for which the resolution of any line array (synthetic or real) is considerably degraded relative to the broadside case. The test was not designed for SAS application so the aperture is significantly undersampled. The processing approach utilizes signal adaptive, ping-to-ping correlation autofocusing

to correct for the effects of platform and target motion, and for ocean instabilities. The processing is tuned to the specific environment by setting the cross-correlation window size based on the two-way group arrival structure of the channel. The aliasing that results from the undersampled aperture is reduced to acceptable levels through the appropriate use of cross-range matched field processing. Images formed by processing the data from a forward-look synthetic aperture measuring 2200 wavelengths with a mean target distance of 9.5 km are presented.

Contributed Paper

11:30

4aUW9. Array measurements and characterizations with sources in motion. William M. Carey (Dept. of Aerosp. and Mech. Eng., College of Eng., Boston Univ., 110 Cummington St., Boston, MA 02215, wcarey@bu.edu)

Traditional processing has treated space and time as independent random variables with assumed ergodicity; however, dynamic-noise fields are produced by moving sources. In these cases, can velocity and range information be used to increase gain, cancel noise, and improve resolution? High-resolution arrays have measured the dynamic ambient noise fields at low frequencies, over long time intervals and at basin/margin spatial

scales, and have resolved/tracked shipping. This paper summarizes experiments with arrays which used source-relative velocity as a processing variable, that is, space-time processing. Results from a closed-basin shipping dominated noise field showed the relative velocity acted as a filter that improved the signal-to-noise ratio for sources with known relative velocity and reduced it for sources having different relative velocities. Results from a shallow water experiment with a vertical array, moving source, and the Bessel transform were shown to determine the wave number spectrum and were comparable to the modeled spectrum in both amplitude and phase. Thus the contention that improved array ship-noise measurement and cancellation is possible if the velocity is known may have merit, especially if modal variations can be adaptively modeled.

FRIDAY AFTERNOON, 2 JUNE 2000

AMERICAN ROOM, 2:30 TO 4:50 P.M.

Session 4pAA

Architectural Acoustics: Computer Modeling and Measurement Techniques for Large Room Acoustics

Paul T. Calamia, Chair

Kirkegaard & Associates, 801 West Adams Street, 8th Floor, Chicago, Illinois 60607

Chair's Introduction—2:30

Contributed Papers

2:35

4pAA1. An echoic room model for acoustic simulations by non-reflection boundaries using the Bergeron method. Hidemaro Shimoda (Inst. of Technol., Shimizu Corp., 3-4-17 Etchujima Koto-ku, Tokyo 135, Japan, shimoda@sit.shimzu.co.jp)

An analytical model of anechoic rooms by distributed equivalent circuits is constructed in the computer using the Bergeron method [Shimoda *et al.*, "Analysis of sound fields in rooms using Bergeron's method," Trans. I.E.C.E. Japan, 72 A, 1, 1–11 (1989); English transl., 1989 Electronics and Communications in Japan, Part 3, Vol. 72(12), Scripta Technica, Inc. (Wiley, New York, 1990)]. In this analytical method, the non-reflection boundaries are modeled by gradually increasing sound transmission losses in specific regions near boundaries just like a sand beach exposing sea waves. Also, the sound transmission characteristics or the sound pressure distributions in the room are investigated. The model can be applied to the sound scattering analysis by an arbitrary object.

2:50

4pAA2. From auditorium ray acoustics to wave acoustics: The next giant step. Herman Medwin (OCEANAC Assoc., 4021 Sunridge Rd., Pebble Beach, CA 93953)

For at least five decades, architectural acousticians have used straight-line, infinite-frequency rays to determine which regions of an auditorium areinsonified by a sound source and which are shadowed. In the next decade, leading acoustical designers of auditoriums will turn to predictions of the timing and amplitude of the different frequencies which partially reflect and bend around real auditorium components such as proscenium arches and boxes. The frequency dependence of wave reflection from rigid, finite surfaces has been developed by Clay *et al.* [J. Acoust. Soc.

Am. **94**, 2279 (1993)]. The calculation of the frequency dependence of the diffraction component, which some have called the BTM technique, is a digitally formulated, finite-wedge, frequency-spectral interpretation of the classic Biot–Tolstoy (1957) theory for impulse scatter from an infinite wedge. The BTM technique, which we described first at an ASA meeting in 1978, has been applied to shadowing by highway noise barriers [Medwin, J. Acoust. Soc. Am. **69**, 1060–1064 (1981)] and to multifrequency reverberation from randomly rough, ocean surfaces [Keiffer and Novarini, in *Computational Acoustics*, edited by D. Lee *et al.* (1990), Vol. 1, pp. 67–81]. The procedure is explained in detail in Medwin and Clay [*Fundamentals of Acoustical Oceanography* (Academic, New York, 1998)]. The justification and analytical extension of this method have recently been formulated by Svensson [J. Acoust. Soc. Am. **106**, 2331–2344 (1999)], who has made the program available on his web site and has suggested applications to auditoriums. The physical interpretation of the BTM technique will be described in terms of the spectral and temporal behavior of finite plates and wedges as modules of more complex components of auditorium surfaces.

3:05

4pAA3. Subjective relevance of objective measures for spatial impression. Lily M. Wang and Anders C. Gade (Dept. of Acoust. Technol., Tech. Univ. of Denmark, Bldg. 352, DK-2800, Lyngby, Denmark)

Several objective measures have been proposed to describe the feeling of spatial impression in concert halls, including Lateral Energy Fraction (LF) and Interaural Cross-Correlation Coefficient (IACC). However, previous studies have shown that LF and IACC values did not highly correlate with each other at individual seat positions in real halls [J. S. Bradley, J. Acoust. Soc. Am. **96**, 3525–3535 (1994)]. To investigate the listener

to correct for the effects of platform and target motion, and for ocean instabilities. The processing is tuned to the specific environment by setting the cross-correlation window size based on the two-way group arrival structure of the channel. The aliasing that results from the undersampled aperture is reduced to acceptable levels through the appropriate use of cross-range matched field processing. Images formed by processing the data from a forward-look synthetic aperture measuring 2200 wavelengths with a mean target distance of 9.5 km are presented.

Contributed Paper

11:30

4aUW9. Array measurements and characterizations with sources in motion. William M. Carey (Dept. of Aerosp. and Mech. Eng., College of Eng., Boston Univ., 110 Cummington St., Boston, MA 02215, wcarey@bu.edu)

Traditional processing has treated space and time as independent random variables with assumed ergodicity; however, dynamic-noise fields are produced by moving sources. In these cases, can velocity and range information be used to increase gain, cancel noise, and improve resolution? High-resolution arrays have measured the dynamic ambient noise fields at low frequencies, over long time intervals and at basin/margin spatial

scales, and have resolved/tracked shipping. This paper summarizes experiments with arrays which used source-relative velocity as a processing variable, that is, space-time processing. Results from a closed-basin shipping dominated noise field showed the relative velocity acted as a filter that improved the signal-to-noise ratio for sources with known relative velocity and reduced it for sources having different relative velocities. Results from a shallow water experiment with a vertical array, moving source, and the Bessel transform were shown to determine the wave number spectrum and were comparable to the modeled spectrum in both amplitude and phase. Thus the contention that improved array ship-noise measurement and cancellation is possible if the velocity is known may have merit, especially if modal variations can be adaptively modeled.

FRIDAY AFTERNOON, 2 JUNE 2000

AMERICAN ROOM, 2:30 TO 4:50 P.M.

Session 4pAA

Architectural Acoustics: Computer Modeling and Measurement Techniques for Large Room Acoustics

Paul T. Calamia, Chair

Kirkegaard & Associates, 801 West Adams Street, 8th Floor, Chicago, Illinois 60607

Chair's Introduction—2:30

Contributed Papers

2:35

4pAA1. An echoic room model for acoustic simulations by non-reflection boundaries using the Bergeron method. Hidemaro Shimoda (Inst. of Technol., Shimizu Corp., 3-4-17 Etchujima Koto-ku, Tokyo 135, Japan, shimoda@sit.shimzu.co.jp)

An analytical model of anechoic rooms by distributed equivalent circuits is constructed in the computer using the Bergeron method [Shimoda *et al.*, "Analysis of sound fields in rooms using Bergeron's method," Trans. I.E.C.E. Japan, 72 A, 1, 1–11 (1989); English transl., 1989 Electronics and Communications in Japan, Part 3, Vol. 72(12), Scripta Technica, Inc. (Wiley, New York, 1990)]. In this analytical method, the non-reflection boundaries are modeled by gradually increasing sound transmission losses in specific regions near boundaries just like a sand beach exposing sea waves. Also, the sound transmission characteristics or the sound pressure distributions in the room are investigated. The model can be applied to the sound scattering analysis by an arbitrary object.

2:50

4pAA2. From auditorium ray acoustics to wave acoustics: The next giant step. Herman Medwin (OCEANAC Assoc., 4021 Sunridge Rd., Pebble Beach, CA 93953)

For at least five decades, architectural acousticians have used straight-line, infinite-frequency rays to determine which regions of an auditorium areinsonified by a sound source and which are shadowed. In the next decade, leading acoustical designers of auditoriums will turn to predictions of the timing and amplitude of the different frequencies which partially reflect and bend around real auditorium components such as proscenium arches and boxes. The frequency dependence of wave reflection from rigid, finite surfaces has been developed by Clay *et al.* [J. Acoust. Soc.

Am. **94**, 2279 (1993)]. The calculation of the frequency dependence of the diffraction component, which some have called the BTM technique, is a digitally formulated, finite-wedge, frequency-spectral interpretation of the classic Biot–Tolstoy (1957) theory for impulse scatter from an infinite wedge. The BTM technique, which we described first at an ASA meeting in 1978, has been applied to shadowing by highway noise barriers [Medwin, J. Acoust. Soc. Am. **69**, 1060–1064 (1981)] and to multifrequency reverberation from randomly rough, ocean surfaces [Keiffer and Novarini, in *Computational Acoustics*, edited by D. Lee *et al.* (1990), Vol. 1, pp. 67–81]. The procedure is explained in detail in Medwin and Clay [*Fundamentals of Acoustical Oceanography* (Academic, New York, 1998)]. The justification and analytical extension of this method have recently been formulated by Svensson [J. Acoust. Soc. Am. **106**, 2331–2344 (1999)], who has made the program available on his web site and has suggested applications to auditoriums. The physical interpretation of the BTM technique will be described in terms of the spectral and temporal behavior of finite plates and wedges as modules of more complex components of auditorium surfaces.

3:05

4pAA3. Subjective relevance of objective measures for spatial impression. Lily M. Wang and Anders C. Gade (Dept. of Acoust. Technol., Tech. Univ. of Denmark, Bldg. 352, DK-2800, Lyngby, Denmark)

Several objective measures have been proposed to describe the feeling of spatial impression in concert halls, including Lateral Energy Fraction (LF) and Interaural Cross-Correlation Coefficient (IACC). However, previous studies have shown that LF and IACC values did not highly correlate with each other at individual seat positions in real halls [J. S. Bradley, J. Acoust. Soc. Am. **96**, 3525–3535 (1994)]. To investigate the listener

envelopment aspect of spatial impression further, subjective paired-comparison tests have been run using signals which have various values for LF, early IACC (from 5–80 ms), late IACC (from 80 ms–1 s), and late lateral relative sound level (GLL). Another proposed measure, called Inter-aural Level Fluctuations (IALF), has also been included, which is based on the rate of change over time of the level difference between the ears. The binaural test signals were generated by the room acoustics simulation program ODEON, using four different musical motifs, and were presented to test subjects via headphones. Results from the experiments are presented. [Work supported by ASA Hunt Postdoctoral Research Fellowship.]

3:20

4pAA4. Measurements of the subjective acoustical parameters on an orchestra platform. Jin Yong Jeon (School of Architectural Eng., Hanyang Univ., Seoul 133-791, Korea)

This paper deals with the relationship between the subjective room-acoustical parameters and the objective properties of the sound field on an orchestra platform. A set of acoustical parameters is investigated and subjective preference tests are conducted both on a real platform and a simulated platform of a concert hall. Parameters EDT, RT, C80, D, Ts, 1-IACC (or LEF) are used for assessments of the quality of the stage acoustics. Advantages and limitations of computer simulations versus field experiments are discussed and both experiments are reviewed, in which the EDT/RT and spatial impressions are investigated. From the subjective preference tests, it is found that musicians form their stage acoustic judgments based on their perception of reverberance, clarity, envelopment, loudness, warmth, brilliance, and overall impression. In addition, the relative reverberance EDT/RT is revealed to have high correlation with the subject-averaged judgments and the spatial impression indices.

3:35

4pAA5. Music-masked maximal-length sequences for auditorium acoustic measurements. Ning Xiang and Charles Sabatier (Natl. Ctr. for Physical Acoustics, Coliseum Dr., Univ. of Mississippi, University, MS 38677, nxiang@olemiss.edu)

Maximal-length sequence (M-sequence) measurement technology has gained acceptance within the architectural acoustics community for quite some years. Owing to the high noise immunity and computational efficiency of the fast M-sequence transform, acoustic qualities in auditoria can be determined quickly and with satisfactory precision. The present work investigates the use of music signals to mask M-sequence excitations. The inherent cross-correlation mechanism of the fast M-sequence transform makes it possible to determine room impulse responses with a positive music/M-sequence ratio (expressed in dB). The feasibility, limitations, and potential applications for using music-masked M-sequences in taking auditorium impulse responses measurements will be discussed in the presentation.

3:50

4pAA6. Acoustic design for University of Southern California, Annenberg School of Communications, lecture hall remodel. Fiona Gillan (Arup Acoustic, 901 Market St., Ste. 260, San Francisco, CA 94103) and Nathan Sevens (Arup Acoustic, Los Angeles, CA 90064)

In 1998, the redesign of the 250-seat lecture hall at the USC Annenberg School of Communications was undertaken to update the aesthetics of the 1974 design and meet the modern functional requirements of the University. Primary uses of the hall were to be lectures, audiovisual presentations, and panel discussions. The existing fan-shaped space had a concrete ceiling, plastered concrete side walls, and a concave rear wall with a slatted wood acoustic finish. The existing reverberation time (RT60) was 0.95 s at mid-frequency, but there was a sense of over-reverberance and a lack of speech clarity. This was attributed to focusing by the curved rear wall and to excessive bass reverberation. The acoustic design targets included a modest reduction of the mid-frequency RT60, a significant reduction of the bass reverberation, and mitigation of the ef-

fects of focusing, as well as a maximum noise level of NC30 for the new HVAC system. Simple materials, including gypsum board wall linings, fabric wrapped panels, and acoustic foam in inconspicuous areas, were used to achieve cost effectiveness. Acoustics and architecture were fully integrated for efficiency of design. The project, completed in 1999, was an acoustic and architectural success. [Architectural Design by Siegal Diamond Architecture.]

4:05

4pAA7. Some acoustic considerations in the use of Kerto-wood in the construction of Sibelius Hall. Christopher A. Storch (Artec Consultants, Inc., 114 West 26th St., 9th Fl., New York, NY 10001-6812)

Sibelius Hall is a 1300-seat concert hall and congress center which was built by the City of Lahti, Finland, with the sponsorship of the Finnish wood industry. The wood industry's support in this construction project was conditional based on using wood in new and experimental ways, not only as a finish material, but as a structural material. A proprietary product called Kerto-wood was used in a wide variety of structural and finish applications. Sand-filled Kerto-wood panels 300 mm thick were used in the slanted exterior walls. Details of the sand-filled panels' construction, the joining details, and the transmission loss testing of the assembly will be discussed. Prefabricated Kerto-wood elements were also used to construct the ceiling, balconies, and reverberation chamber doors. Discussion of the reflective properties will also be presented.

4:20

4pAA8. Adjustment of reverberation in the rehabilitation of theaters. Its application to the architectural patrimony of theaters in Andalusia, Spain. Angel Luis Leon, Jaime Navarro, and Juan Jose Sendra (School of Architecture of Seville, Av. Reina Mercedes, 2 Sevilla 41012, Spain, leonr@arquitectura.us.es)

One hundred years after its formulation by W. C. Sabine, the reverberation time continues to be, we feel, the one most useful indicator for architects when undertaking the rehabilitation of theaters in which the main objective is to obtain good acoustic conditions. The reverberation time is easily quantifiable during the planning phase, thus proving to be of great interest in the process of theater design. The adjustment of reverberation times in a theater to the different frequencies at intervals regarded optimum is considered to be an indispensable determinant, though not in itself sufficient, to obtain a good acoustics. This adjustment is intimately related to the decisions proper to the work of the architect in interior design, such as the election of surface materials or the volume of the locale. The authors of this study, who have participated in the rehabilitation of 35 theaters in Andalusia as acoustic consultants, propose to relate the results of their work with the different proposals of the principal researchers in room acoustics, about optimum values of reverberation, with the aim of obtaining a series of conclusions to offer to professionals of rehabilitation.

4:35

4pAA9. Analysis of spatial acoustic qualities from listening in rooms. Martin A. Gold (321 ARCH, Dept. of Architecture, Univ. of Florida, Gainesville, FL 32611, mgold@ufl.edu)

Statistical analysis was conducted on data from in-room listening studies to evaluate the effects of spatial impression and auditory source width as independent variables with regard to overall acoustic impression. The listening studies were conducted with groups of listeners (30+) in the rooms under evaluation. Subjects in each circumstance were asked to evaluate the acoustic environment using a bi-polar semantic rating scale for the acoustical qualities of overall impression, loudness, clarity, reverberance, treble response, bass response, spatial impression, auditory source width, echoes, and background noise. Ratings were taken in a

variety of listening environments ranging from outdoors to a 2000-seat auditorium to elicit a range of qualitative assessments. Both live music and recorded music and speech sources were evaluated in the study. The study also includes a comparison of in-room ratings. Results indicate that, in

most cases, it is difficult to statistically isolate independent spatial variables with regard to overall room ratings in complex environments. In specific room locations, these variables appear to be statistically independent.

FRIDAY AFTERNOON, 2 JUNE 2000

CANADIAN ROOM, 2:30 TO 5:10 P.M.

Session 4pAO

Acoustical Oceanography: Fine Scale Ocean Processes, Bubbles and Internal Waves

Grant B. Deane, Chair

*Marine Physical Laboratory, Scripps Institution of Oceanography, University of California—San Diego,
La Jolla, California 92093-0238*

Chair's Introduction—2:30

Contributed Papers

2:35

4pAO1. Nonlinear sound scattering by subsurface bubble clouds in sea. Lev A. Ostrovsky (Univ. of Colorado, CIRES/NOAA ETL, 325 Broadway, Boulder, CO), Zinovy Kluzek (Polish Acad. of Sci., Sopot, Poland), Alexander M. Sutin, Irina A. Soustova, Alexander I. Matveev, and Andrey I. Potapov (Russian Acad. of Sci., Nizhny Novgorod, Russia)

This presentation is based on a series of field experiments performed in a shallow sea (Baltic Sea). Investigated were bubble clouds of two types: one created in a motorboat's jet, and the other existing naturally in the subsurface layer. A two-frequency signal (30 and 35 kHz) was radiated from below (at a 15-m depth) and received upon backscattering at sum and difference frequencies (65 and 5 kHz). The intensity of the scattered signal averaged over five subsequent pulses was measured. Theoretical consideration is given for reverberation at sum frequency and coherent propagation of the difference-frequency wave with subsequent reflection from the surface. The method is based on integration of fields nonlinearly scattered by single bubbles. A good qualitative and, in many cases, quantitative agreement was obtained between theory and experiment. A possibility to use nonlinear scattering for bubble distribution evaluation is also discussed.

2:50

4pAO2. High-frequency propagation through the wave boundary layer. David Farmer, Svein Vagle (Inst. of Ocean Sci., 9860 W. Saanich Rd., Sidney, BC V8L 4B2, Canada, farmerd@dfp-mpo.gc.ca), and Grant Deane (Scripps Inst. of Oceanogr., La Jolla, CA 92093-0230)

The surf zone presents one of the most extreme and challenging environments for the study of shallow water propagation where the complexities of a sloping beach, wave amplification, wave breaking, and the presence of dense bubble clouds combine to block or modify acoustic signals in a variety of ways. Propagation measurements at 12 and 100 kHz reveal sporadic drop-outs during which the signal can be abruptly blocked for 1–5 min or may be weakly detected but with strong modulation by the surface waves. Bubbles introduced to the water column through breaking are subject to persistent vertical mixing through the turbulent wave boundary layer which becomes progressively more energetic closer inshore. The bubble suspension is further modified by buoyancy and dissolution effects. Different acoustic frequencies are attenuated by bubbles of different radii, which in turn respond differently to the competing influences of turbulent mixing, buoyancy, and dissolution. We explore the implication of these effects on high-frequency propagation with a model that includes bubble injection, vertical mixing by the wave boundary layer, and bubble cloud

evolution. The predictions are compared with observations of high-frequency propagation, wave amplification, wave breaking, and bubble distributions. [Work supported by the US ONR.]

3:05

4pAO3. A maximum likelihood estimation algorithm for tracking two multipaths. Alan Barton and Daniela Di Iorio (Dept. of Marine Sci., Univ. of Georgia, Athens, GA 30602)

A high-frequency (67 kHz) acoustical scintillation experiment was carried out in the northern entrance to Hood Canal, Puget Sound, WA. This experiment made use of a 4-transmitter and 4-receiver array configured as a T-shape. The two-dimensional feature of this array was designed to measure both along-channel small-scale properties as a result of advection and vertical properties as a result of refraction from temperature/salinity stratification. Sampling was carried out at 5 Hz for all transmitters so as to understand the turbulent effects of the medium on acoustic scattering. This experiment tested if acoustical scintillation techniques can be extended to long distances (2380 m) and in deep (100 m) coastal channels having strong tidal currents and temperature/salinity stratification. With the longer path lengths and stable stratification, acoustic propagation resulted in multipath arrivals which were separable for most of the measurement period. In order to study the oceanographic effects on the direct path signal at 20-m depth and the upward refracted signal into the surface mixed layer, a maximum likelihood estimation algorithm is developed that tracks the two paths and calculates amplitude, phase, and travel time for each. [Experiment was carried out by SAIC of Bellevue, WA, and student support was from the University of Georgia.]

3:20

4pAO4. The time-domain behavior of coupled resonator scatterers. C. Feuillade (Naval Res. Lab., Stennis Space Center, MS 39529-5004, cf@nrlssc.navy.mil)

In previous work [J. Acoust. Soc. Am. **98**, 1178–1190 (1995)] a coupled oscillator formalism was introduced for describing collective resonances, scattering, and superresonances of multiple air bubbles in a fluid. Subsequent time-domain investigations, of the impulse response of coupled systems, have disclosed the conditions which exactly determine whether the ensemble behavior is properly described using: (a) a multiple scattering paradigm or (b) a collective modal methodology. The important criteria are the Q of the individual scatterers, and their typical separations in the medium. For highly damped or sparse systems, e.g., scattering from loose schools of swimbladder fish, or from a gassy seabed containing entrained bubbles, the scatter counting approach may be applicable. For

more strongly coupled systems, e.g., a dense cloud of resonating bubbles in the water column, the individual scattering events are typically not resolvable and a modal approach must be used. The result has implications for both volume and bottom-scattering applications. Theoretical and numerical analysis will be presented. [Work supported by ONR.]

3:35–3:55 Break

3:55

4pAO5. Wideband determination of source functions using the Faran elastic sphere scattering model. C. Feuillade, R. W. Meredith (Naval Res. Lab., Stennis Space Center, MS 39529-5004, cf@nrlssc.navy.mil), N. P. Chotiros (Univ. of Texas, Austin, TX 78713-8029), and C. S. Clay (Univ. of Wisconsin—Madison, Madison, WI 53706)

During August 1998, a bottom scattering tank experiment was performed at ARL-UT to measure wideband reverberation from multiple objects (e.g., cobbles and pebbles) placed on a sediment simulation of the sea floor. As a first step in this experiment, a procedure was adopted to calibrate the source transfer function by measuring reflected signal pings from the water surface (which has a spike impulse response), with the transducer directed vertically upwards. Subsequent analysis showed that the strength of the reflections may have induced a nonlinear receiver response in the transducer. Fortunately, time-domain scattering measurements had also been made with a stainless steel ball suspended in the water column. These data were used to study the source drive by deconvolving the theoretical impulse response of a steel sphere, obtained via the Faran scattering model, from the scattered time signals. Both the characteristics of the experimental source signal and another, low-power, source measured later were examined. Results show the critical importance of accurate input–output system calibrations for wide signal bandwidth sound scattering research, and have implications for boundary and volume scattering applications. Experimental data and theoretical numerical analysis will be presented. [Work supported by ONR.]

4:10

4pAO6. Simultaneous measurements of the bubble size distributions and radiated noise from open-ocean whitecaps. Grant Deane and Dale Stokes (Marine Physical Lab., Scripps Inst. of Oceanogr., UCSD, La Jolla, CA 92093-0238, grant@mpl.ucsd.edu)

It is now widely acknowledged that wind-induced oceanic ambient noise comes from ringing bubbles formed by breaking waves at the ocean's surface. The numbers, sizes, creation rates, and spatial distribution of these bubbles are all important in determining the acoustic radiation signature of a breaking wave. Results from an experiment to measure the bubble size distribution within open-ocean whitecaps and simultaneously record the burst of sound radiated by forming whitecaps will be presented. The bubble size distribution measurements were made with an optical instrument mounted within 0.5 m of the ocean's surface beneath a surface-following frame. The instrument is capable of resolving the time-evolving bubble size distribution on millisecond time scales and centimeter length scales within a whitecap. The simultaneous optical and acoustical measurements are important for modeling the sound of breaking waves in the open ocean and the surf zone. [Work supported by ONR and NSF.]

4:25

4pAO7. Results of an acoustic normal-mode analysis of SESAME II data, and conclusions on internal wave activity. Jacob George and Robert L. Field (NRL Code 7185, Stennis Space Center, MS 39529)

An acoustic normal-mode analysis has been used to examine the effect of internal waves (IWs) both at the receiving array and along the path of acoustic transmission in SESAME II data (year 1996; off the coast of

Scotland). Power transfer from lower to higher modes has been observed to occur only intermittently. The sound-speed profiles (SSPs) at the receiving array during these occurrences are found to be nearly identical. Fourier analysis has revealed that a prominent peak in the frequency band 39–47 cycles/day in the acoustic reception at 800 Hz corresponds to a similar peak in the same frequency band in the thermal data at the receiver. That coincidence was investigated through normal-mode analysis, IW time delay between two nearby vertical arrays of thermal sensors, and a cylindrical wavefront model. The results show that the two peaks could not have been caused by the same IW. Therefore, the IW responsible for the acoustic peak should have originated farther north. [Work supported by ONR and NRL.]

4:40

4pAO8. Internal waves and temperature inversion in the East China Sea. Tao Wang (Eng. Sci. 104-44, California Inst. of Technol., Pasadena, CA 91125), Tian-Fu Gao, De-Jun Jiang, and Yun-Peng Zhang (Chinese Acad. of Sci., Beijing 100080, PROC)

Several shallow-water acoustics experiments were performed in the Zhoushan offshore area of the East China Sea from 1996 to 1998. The water depth is about 55 m at the mooring locations. Temperature chains sampling once per 6.4 s or 1 min provide the oceanographic data. One clearly sees the long-wavelength internal tides. The dominant component of the internal wave field is the semidiurnal tide of 10-m height. High-frequency internal waves with 10–15 min duration and 2-m amplitude ride on top of the internal tide. These rank-ordered waves are likely internal solitons, which is validated by recent SAR images. The asymmetry of the initial and final states indicates that the solibore description is somewhat applicable to this case, but more evidence is needed. The inversion phenomenon of temperature and salinity frequently occurs in the long and narrow East China Sea area off the coast of Jiangsu and Zhejiang. Temperature inversion is noted via the measurements and shows yearly variations. Its structure and generation mechanism are discussed. [Work supported by NSFC.]

4:55

4pAO9. Higher-order spectral analysis of shallow-water internal waves. Tao Wang (Eng. Sci. 104-44, California Inst. of Technol., Pasadena, CA 91125) and Tian-Fu Gao (Chinese Acad. of Sci., Beijing 100080, PROC)

Higher-order spectral analysis is a useful means for investigating nonlinear properties of shoaling random surface waves and a wide range of other phenomena, but has been considered to be fruitless for internal waves based on some studies in which the energies are represented by the deep-ocean GM model. However, it has been increasingly recognized that nonlinearity manifests itself in various shallow-water internal wave phenomena. The classical phase-blind power spectra contain information in the second-order statistics of a signal, and only suffice for describing Gaussian processes. Consequently, there are three main reasons for using higher-order spectra: to extract information due to deviations from Gaussianity, to recover correct phase properties, and thus to detect and quantify nonlinearity in time series. Existing techniques of bispectral analysis are particularly useful for studying quadratic interaction, the lowest order of nonlinear interactions. In this paper, it is found that internal waves on the continental shelf could produce a characteristic bispectral signal of sufficient strength to be verified by observations, and that real and imaginary parts of bispectra represent contributions to vertical and horizontal asymmetry of the internal wave elevations from each frequency pair, respectively. [Work supported by NSFC.]

Session 4pMU

Musical Acoustics: Musical Instruments and Scales

James P. Cottingham, Chair

Department of Physics, Coe College, Cedar Rapids, Iowa 52402

Contributed Papers

2:30

4pMU1. 2DOF model applied to the violin octet A0 and A1 cavity modes. George Bissinger (Dept. of Phys., East Carolina Univ., Greenville, NC 27858)

An optimized 2DOF model for the A0 and A1 cavity modes [E. A. Shaw, *J. Acoust. Soc. Am.* **87**, 398–410 (1990)] has been used to estimate these frequencies for the Hutchins–Schelleng violin octet. The 2DOF model is notable because it predicts coupling between the A0 and A1 cavity modes [G. Bissinger, *J. Acoust. Soc. Am.* **104**, 3608–3615 (1998)]. This coupling reduces the volume dependence of the A0 mode. 2DOF predictions of A0 frequencies are compared with measurements made on the four largest members of the octet during the course of rib-height modifications [C. M. Hutchins, *J. Acoust. Soc. Am.* **92**, 639–650 (1992)]. Comparisons with new A0 and A1 measurements for the octet will also be made. [Work supported by NSF DMR9802656.]

2:45

4pMU2. Acoustics of mandolins. David Cohen (J. Sargeant Reynolds Community College, Richmond, VA 23261) and Thomas D. Rossing (Northern Illinois Univ., DeKalb, IL 60115)

We compare the normal modes of vibration and sound radiation from two Gibson F-type mandolins, one with an elliptical sound hole and one with *f*-holes. The instruments were constructed by the first author, using carved plates and other parts supplied by the Gibson company. Normal modes of vibration were determined with the help of electronic TV holography. Air cavity modes were determined by imbedding the instruments in sand to immobilize the plates. Normal modes in a mandolin are somewhat akin to those observed in guitars, although the coupling between plates and the enclosed air is weaker in mandolins.

3:00

4pMU3. An objective method for determining soundboard material quality. Mark French (Robert Bosch Corp., Farmington Hills, MI 48335-1672) and John G. Cherng (Univ. of Michigan—Dearborn, 4901 Evergreen Rd., Dearborn, MI 48128)

For a stringed instrument of the guitar or violin family, the dynamic characteristics of the soundboard are extremely important in determining the sound quality of the finished instrument. Since an exceptional instrument can be worth several tens of thousands of dollars, the importance of getting such a basic step correct are obvious. This paper describes our efforts to describe the desirable characteristics of a guitar soundboard blank using what is currently some of the highest quality test equipment available. Our goal is to provide objective standards that can be implemented in an industrial guitar-making process to help ensure the highest quality possible sound quality in the finished instruments. Formal modal tests on a range of soundboard blanks ranging from inexpensive ones intended for use by those just learning their craft to master grade blanks suitable for the highest quality instruments have been performed. The subjective ratings used to grade these blanks with objective test results are correlated.

3:15

4pMU4. Modal analysis of a new steelpan: The ping. Thomas D. Rossing (Phys. Dept., Northern Illinois Univ., DeKalb, IL 60115), Uwe J. Hansen (Indiana State Univ., Terre Haute, IN 47809), Felix Rohner, and Sabina Schärer (Panart, Bern, Switzerland)

The ping, which has the same playing range as a tenor steelpan, is a new instrument in the steelpan family. Its playing surface is sandwich hardened steel having a uniform thickness of 0.9 mm, and its skirt is of stainless steel 1.25-mm thick. Each note area has a small elliptical dome at the center to provide stability and a durable strike point. We compare the modes of vibration of the ping to those of a tenor steelpan. In both pans, the (0,1) mode of each note is tuned an octave above the fundamental, and the (1,0) mode is tuned a twelfth (three times the frequency) above the fundamental in the lower notes. In the ping, it is possible to tune a twelfth throughout more of the range, and the frequency remains unchanged for a longer time during hard playing. Radiation of sound by the skirt is compared.

3:30

4pMU5. Acoustics of rototoms. Thomas D. Rossing and Juyong Kwon (Phys. Dept., Northern Illinois Univ., DeKalb, IL 60115)

Rototoms are single-membrane drums that can be quickly tuned by rotating the drum. The sound spectra and modal shapes recorded holographically show membrane modes quite similar to those observed in other drums, such as tom-toms, snare drums, and timpani. The sound is dominated by the fundamental radiated bathe (0,1) mode of vibration. The frequency ratio of the various partials changes with tension over the playing range of each drum, resulting in a changing timbre.

3:45

4pMU6. Lip and throat configuration in flute playing. Joon-Hee Beth Hwang^{a)} (Faculty of Health Sci., The Univ. of Sydney, Australia) and William Thorpe (The Univ. of Sydney, East St. Lidcombe, NSW 2141, Australia)

This study reports on simultaneous measurements of lip aperture and glottal aperture during varied performance tasks by each of a group of six expert flute players. Lip aperture was recorded using a video camera, and glottal aperture with a nasendoscope. Video images were selected from stable portions of the recording and digitized, and measurements of vocal fold and lip opening were made from the display. The results showed that there was a general trend of increasing lip aperture with increase in loudness and decrease in lip aperture with increasing pitch. Similarly, glottal aperture increased with increase in loudness and decreased with increase in pitch. The ratio of glottal to lip aperture area generally decreased with increase in pitch, the glottal aperture always being greater than the lip aperture. Although the results of this study generally agree with those of previous studies [J. W. Coltman, *J. Acoust. Soc. Am.* **40**, 99–107 (1966), **44**, 983–992 (1968), N. H. Fletcher, *ibid.* **57**, 233–237 (1975), S. Mukai, *J. Otolaryngol. Jpn.* **92**, 260–270 (1989)], players showed considerable individual variation. This suggests that various compromises in technique are possible among expert players. ^{a)}Present address: 7205 Mentor Ave., #C102, Mentor, OH 44060.

4:00

4pMU7. The acoustics of a symmetric free reed coupled to a pipe resonator. James P. Cottingham (Phys. Dept., Coe College, Cedar Rapids, IA 52402)

The Asian free-reed mouth organs employ symmetric free reeds mounted in resonating pipes, with the reed vibration strongly coupled to the pipe resonance. The sheng, the sho, and the khaen use a single pipe for each reed, constructed so that the pipe resonance frequency is fairly close to the natural frequency of the reed. The playing frequency is typically slightly above the resonant frequencies of both the reed and the pipe. The free-reed pipe with finger holes (known in China as the bawu) employs a pipe resonator of variable effective length in which both the pipe resonance and sounding frequency are normally well above the natural reed frequency, resulting in a striking change in tone quality. The operation of these instruments has been studied both experimentally and theoretically, with particular attention to the coupling of the reed vibration with the pipe resonator. Experimental measurements include both studies of reed vibration and impedance measurements of the pipes. In general, the experimental results can be shown to agree with the predictions of simple theoretical models.

4:15

4pMU8. The influence of upstream geometry on the activation pressure of free and restricted reed configurations. Marius O. Vermeulen, Jabus A. Wessels (Dept. of Medical Physiol., P.O. Box 19063, Tygerberg, 7505, Cape Town, South Africa, jaw@gerga.sun.ac.za), and Theodore W. von Backström (Univ. of Stellenbosch, Matieland, South Africa)

This paper describes a computationally efficient algorithm for the automatic calculation of reed activation pressure. The method was evaluated using both free and restricted reed configurations which were not connected to any external air column. It is shown how the reed activation is

influenced by intrinsic factors such as rounded and minute burrs on the edge of the reed as well as extrinsic factors such as upstream geometry of the experimental apparatus. The introduction of constrictions in the flow path upstream from the reed resulted in marked changes in the activation pressure which was highly dependent on the length of the constriction. Expansions of the same magnitude, however, had less influence on the reed activation pressure.

4:30

4pMU9. A complexity measure for musical scales. Alpar Sevgen (Dept. of Phys., Bogaziçi Univ., Bebek 80815, Istanbul, Turkey, sevgena@boun.edu.tr)

Equally tempered scales with N semitones and M notes and interval structures $\mathbf{n} = \{n_1, n_2, \dots, n_M\}$, where n_k is the number of semitones between the notes t_k and t_{k+1} , possess the following properties: Each distinct interval structure \mathbf{n} corresponds to a multiplet of N scales. Members of a multiplet can be labeled by a set of integers $\{c\}$, modulo N , called scale labels. Each scale label is the *difference* between the number of sharps and flats occurring in that scale and is unique within the multiplet if N and M are relative primes. This labeling does not differentiate between different scale structures. To do this, *complexity* is introduced as the *sum* of the number of sharps and flats occurring in a scale. For $N = 12$ and $M = 7$, out of 462 possible scale structures, the major scale and its cyclical permutations, called modes, have the minimum complexity which allows the practical use of the key signatures in music. Complementary scales where notes and no notes are interchanged have the same complexity. The minimum and maximum complexity scales occupy the opposite ends of the energy spectrum under the force laws $\pm n^\alpha$ ($\alpha \neq 0$), between the notes of a scale.

FRIDAY AFTERNOON, 2 JUNE 2000

SPANISH ROOM, 1:30 TO 5:15 P.M.

Session 4pPA

Physical Acoustics: Outdoor Sound Propagation and Acoustic Seismic Coupling

Gregg D. Larson, Chair

School of Mechanical Engineering, Georgia Institute of Technology, Atlanta, Georgia 30332

Contributed Papers

1:30

4pPA1. Systematic investigation on acoustic-to-seismic responses of landmines buried in soil. James M. Sabatier and Ning Xiang (Natl. Ctr. for Physical Acoust., Coliseum Dr., Univ. of Mississippi, University, MS 38677, sabatier@olemiss.edu)

Recently, acoustic-to-seismic coupling has been successfully applied to landmine detection [Sabatier and Xiang, *J. Acoust. Soc. Am.* **105**, 1383 (1999); **106**, 2143 (1999)]. When airborne sound penetrates the surface of ground it is refracted towards the normal. If a landmine is buried below the surface of an insonified patch, the transmitted waves will be scattered or reflected, resulting in increased ground surface vibrational amplitudes. These distinct acoustic-to-seismic coupled vibrational changes are sensed using a scanning laser Doppler vibrometer (LDV) device. To better understand this mine detection phenomenon, the present work is a systematic investigation of the acoustic-to-seismic response to different types of mines in different soil types and at different burial depth has been conducted. [This work is supported by U.S. Army Communications-Electronics Command.]

1:45

4pPA2. Air acoustic sensing of seismic waves. Gregg D. Larson, James S. Martin (School of Mech. Eng., Georgia Inst. of Technol., Atlanta, GA 30332-0405), Waymond R. Scott, Jr., and Cheng Jia (Georgia Inst. of Technol., Atlanta, GA 30332)

Propagation of elastic waves in damp, compacted sand involves pressure, shear, and Rayleigh waves. The associated seismic surface displacements can be detected by sensing the acoustic pressure immediately above the surface. Propagation speeds are very low in sand. The high wave numbers of seismic displacements are, therefore, evanescent in air. Thus, the acoustic pressure can only be measured well within a seismic wavelength of the surface. Planar near-field acoustic holography techniques can then be used to back-propagate these signals and calculate surface displacements. Measurements have been made using a laboratory experimental model to investigate the potential of using this technique to detect buried land mines. The experimental model utilizes a surface-coupled transducer to generate elastic waves in a sand-filled tank, which simulates the earth. The microphone and a radar system were used to independently

measure the surface displacements. Data taken with both sensors compare well and exhibit the signature of a buried inert antipersonnel mine. For a 100–800-Hz incident pulse, the mine signature can be seen in the raw microphone data when the height of the microphone is less than 3 cm. Holographic signal-processing techniques will be investigated to increase the allowable height for the microphone. [Work supported by ARO.]

2:00

4pPA3. Electric arc source for high-frequency seismic measurement.

James S. Martin, Gregg D. Larson, Peter H. Rogers (School of Mech. Eng., Georgia Inst. of Technol., Atlanta, GA 30332-0405, james.martin@me.gatech.edu), and Waymond R. Scott, Jr. (Georgia Inst. of Technol., Atlanta, GA 30332-0250)

An electric arc source was designed to study high-frequency seismic surface wave propagation. The noncontact nature of this source made it feasible for use in synthetic aperture transmit arrays. The transmit signal, which was not linearly controllable, was found to be predominantly in the 1- to 4-kHz band with a Gaussian spectrum. This is an octave below the simultaneously generated air acoustic pulse. The source was used to create a synthetic line array in conjunction with a stationary receiver. The experiment was conducted in a sand-filled tank. Surface wave speeds in the range of 80 m/s were measured on the resulting seismograms. Significant dispersion occurred in the propagating waveform at distances both near and far from the source. Surface wave arrivals were discernable over 1 m from the source and compressional head waves could be observed within 30 cm. The data were in good agreement with lower frequency measurements made by other techniques. The surface wave generation was studied and found to be a combination of the surface interaction of the arc and the air acoustic interaction. Strong hysteresis was observed in the first arcing event. Later, the signal was smaller but sufficiently stable for averaging. [Work supported by ARO.]

2:15

4pPA4. Seismic/electromagnetic system for landmine detection.

Waymond R. Scott, Jr. (School of Elec. and Computer Eng., Georgia Inst. of Technol., Atlanta, GA 30332-0250), Gregg D. Larson, James S. Martin, and Peter H. Rogers (Georgia Inst. of Technol., Atlanta, GA 30332-0405)

A system has been designed for the detection of buried landmines. The system uses a stationary seismic source in conjunction with a movable displacement sensor that is based on an 8 GHz CW radar. The sensor measures the surface displacement by analog demodulation of the radar signal, which is reflected from the soil surface and modulated by the surface motion. The sensor is not in direct contact with the soil surface and is, therefore, capable of interrogating surface motion immediately above a buried mine. This configuration provides the dual advantage of removing half the seismic propagation path that would be encountered with a classical pulsed echo technique and detecting localized fields that would not propagate to a remote receiver. The system has been used in the laboratory to image inert antipersonnel mines and simulated antitank mines buried in damp compacted sand. Signal processing in the wave number domain provides significant improvement in the contrast between mine-related and background motion. The simplest detection cue for antipersonnel mines was found to be low-frequency resonances of their trigger mechanisms. These responded to seismic excitations with substantial local displacement. The resonances made these mines easily discernable from buried clutter such as rocks and sticks. [Work supported by ARO.]

2:30

4pPA5. A three-dimensional model for elastic waves in the ground.

Christoph T. Schroeder, Kangwook Kim, and Waymond R. Scott, Jr. (School of Elec. and Computer Eng., Georgia Inst. of Technol., 777 Atlantic Dr., Atlanta, GA 30332, christoph.schroeder@ee.gatech.edu)

A three-dimensional finite-difference time-domain model for elastic waves in the ground has been developed and implemented on a massively parallel computer. The model is based on the three-dimensional equation

of motion and the stress–strain relation, from which a first-order stress-velocity formulation is obtained. The boundary between the soil and the air is modeled as a free surface. A perfectly matched layer is implemented at the remaining grid edges to absorb the outward traveling waves. The numerical model has been developed as part of a project in which elastic and electromagnetic waves are used synergistically to detect buried landmines. The numerical model is being used to study the interaction of the elastic waves with the buried mines. To verify that the model accurately predicts the mine–wave interaction, the eigenfrequencies of various solid bars and plates are determined numerically and compared to analytical solutions. Currently, the model is being refined to incorporate loss within the bulk medium. Results will be shown for various landmines buried in both loss-free and lossy ground. [Work supported by ARO and ONR.]

2:45

4pPA6. Detection of land mines in fluid-saturated unconsolidated soil:

Numerical modeling. Yanqing Zeng and Qinghuo Liu (Dept. of Elec. and Comput. Eng., Duke Univ., Durham, NC 27708)

Because of the strong interactions of waves with the solid grains and the fluid in the pore space, it is more appropriate to model the soil as a fluid-saturated unconsolidated material than a single-phase elastic medium. A three-dimensional finite-difference method for modeling acoustic waves propagating in fluid-saturated unconsolidated soil has been developed. Instead of the conventional elastic wave equations, Biot's equations are used for the poroelastic model. Based on the strain–stress relationship, Biot's equations are reformulated into a first-order hyperbolic system which is equivalent to strain-velocity formulation. A perfectly matched layer (PML) is used to absorb outgoing waves at the truncated boundary of an unbounded medium. The numerical method is validated by comparing numerical results to an analytical solution. Models of a land mine buried in fluid-saturated unconsolidated soil are developed. The numerical method is used to study the interaction of acoustic waves with the buried mines. Comparison of these results is made with those for a buried land mine in a conventional single-phase elastic soil model.

3:00–3:15 Break

3:15

4pPA7. Acoustically induced slow dynamics in nonlinear mesoscopic elastic materials.

Alexander M. Sutin (Stevens Inst. of Technol., 711 Hudson St., Hoboken, NJ 07030), Paul A. Johnson, and James A. TenCate (Los Alamos Natl. Lab., Los Alamos, NM 87545)

We have known about slow dynamics in rock due to continuous wave excitation drive for several years (<http://www.ees4.lanl.gov/nonlinear>). TenCate, Smith, and Guyer (see abstract, this meeting) have recently discovered that both the elastic modulus and the wave dissipation display log time recovery in granular solids, and that it may be thermally or mechanically induced. Much to our surprise, we have discovered that a CW or broad-frequency band acoustic source can also induce slow dynamical response. This response was observed as a variation of the ultrasonic probe wave amplitude, the resonance frequency, and Q factor after the action of a pump wave. The slow time recovery took place in materials such as powdered metals, damaged materials, concrete, and rocks as well. The variations of material properties due to the action of pump waves lead to transient amplification and an obscuration of CW probe waves. The observed behavior may be more universal than was first thought. The results have potential implications to many topics, including laboratory wave studies, earthquake strong ground motion, elastic waves emanating from a point source, damage detection, and manufacturing processes. [Work supported by Stevens and by the Department of Energy: Office of Basic Energy Sciences.]

4p FRI. PM

4pPA8. Quantifying a new form of acoustic creep in rocks and concrete. Eric Smith, James A. TenCate (Earth and Environ. Sci., Los Alamos Seismic Res. Ctr., Los Alamos Natl. Lab., Los Alamos, NM 87545, desmith@arlut.utexas.edu), and Robert A. Guyer (Los Alamos Natl. Lab., Los Alamos, NM 87545)

A new form of nonlinear elastic response, which we call “slow dynamics,” has recently been found to scale universally in all sandstones, limestone, and concretes we have tested. Slow dynamics is unique among known creep processes; the elastic modulus and inverse loss tangent drop in response to harmonic acoustic stress, so the response violates the symmetry of the source. We show here that the slow dynamic recovery rate is a function of temperature, consistent with a thermal activation model for creep requiring a spectrum of energy barriers. We also show that short-term modulus drop, with slow dynamic recovery, accompanies temperature change of either sign, again violating the source symmetry. These two modes of activation are compared quantitatively, and suggest a common origin for slow dynamics in disequilibrated internal strain fields. [Work supported by the Department of Energy: Office of Basic Energy Sciences and IGPP, the Institute for Geophysics and Planetary Physics at Los Alamos.]

4pPA9. Examination of surface roughness of porous soil via acoustic backscatter. Michael L. Oelze and James Sabatier (NCPA, Univ. of Mississippi, University, MS 38677, mloelze@olemiss.edu)

Work has been done to determine acoustically the pore properties of soils, such as porosity, permeability, and tortuosity. Early methods used probe microphones to determine the soil pore characteristics by looking at the attenuation and phase change of sound propagating through the soil pores. The attenuation and phase change of the sound were then related to a surface impedance that incorporated the soil pore properties. Seeking noninvasive means of finding these properties, forward scattering or acoustic level difference experiments were performed. It was seen that sound propagated over porous surfaces is reduced by the surface impedance and by the surface roughness. Theories have been developed to take into account the reduction in acoustic signal due to the finite surface impedance. Difficulties have arisen in separating the signal loss due to impedance and the signal loss from roughness scattering. Acoustic backscatter has been used to study the roughness statistics of soil surfaces in an effort to determine the effects of roughness apart from the surface impedance effects. The roughness statistics obtained via acoustic backscatter are then compared with alternate nonacoustic methods for examining surface roughness.

4pPA10. High-angle propagation modeling in the atmosphere. Xiao Di (Appl. Res. Lab., Pennsylvania State Univ., State College, PA 16804) and Kenneth E. Gilbert (Univ. of Mississippi, University, MS 38677)

Many important situations in outdoor sound propagation involve high-angle propagation, for example, elevated sources and ground receivers, or vice versa. For such cases, we have compared two parabolic equation models: the GFPE model of Gilbert and Di and an atmospheric version of the RAM algorithm of Collins. Although the GFPE model, which is based on a split-step Fourier method, gives exact results for a homogeneous atmosphere (constant sound speed and constant density), it was initially thought to be less accurate for an inhomogeneous atmosphere than the RAM algorithm, which is based on high-order Padé approximations. Surprisingly, it was found that for cases with sufficiently weak atmospheric gradients, the GFPE is considerably more accurate at high angles than the atmospheric RAM model, given equal computational effort with both models. With increased computational effort (i.e., higher-order Padé expansions and shorter range steps), the RAM model gives high-angle accuracy comparable to the GFPE. As the atmospheric gradients steepen, the accuracy of the RAM model exceeds that of the GFPE. Nevertheless, for many typical atmospheric conditions, the GFPE offers sufficient accuracy

and considerable computational savings. In this paper, we compare the GFPE and RAM models with benchmark calculations from a fast-field program (FFP) and discuss practical guidelines for using the GFPE and RAM models for high-angle propagation.

4pPA11. Detection and tracking of broadband airborne sources: An “acoustic fence” approach. Jay E. Williams, Chad M. Williams (Miltec, Inc., NCPA, Coliseum Dr., University, MS 38677, jwilliams@mil-tec.com), and Kenneth E. Gilbert (Univ. of Mississippi, NCPA, University, MS 38677)

An “acoustic fence” concept is discussed as an advance detection system for approaching subsonic airborne vehicles. Each “post” of the acoustic fence is an array of three closely spaced microphones at right angles to each other on the ground. The posts themselves are spaced approximately 1 kilometer apart. It is shown that the azimuthal and elevation angles for an airborne vehicle can be accurately determined from simple cross correlation of the broadband time series measured at a single array (i.e., at one post). With two arrays, the vehicle track can be determined by finding the stationary points of an iterative nonlinear equation. The equation is of the form $\tau_p = f(\tau_e)$, where τ_e and τ_p are, respectively, the estimated and predicted acoustic travel-time differences for two adjacent arrays. The stationary points, i.e., where $\tau_p = \tau_e$, allow the vehicle track to be computed directly from triangulation. Examples are discussed using both synthetic data and actual field measurements.

4pPA12. Sound propagation through thermal turbulence near a convex surface. Ph. Blanc-Benon, J. Wasier, and D. Juvé (Ecole Centrale de Lyon, LMFA UMR CNRS 5509, BP 163 69131 Ecully Cedex, France, acous@ec-lyon.fr)

Making use of an acoustic analogy, we study the sound propagation above a convex surface, to simulate the propagation of sound in an upward-refracting medium and to investigate the scattering of sound by turbulence into an acoustic shadow zone. An experimental study was achieved under well-controlled laboratory conditions. A heated grid is placed horizontally in a large anechoic room and the mixing of the free convection plumes above the grid generates a homogeneous isotropic random thermal field. The convex surface is simulated by a vertical cylinder. Experimental data are obtained by varying both the frequency of the acoustic source and the distance of propagation. Measurements will be presented for a rigid convex surface. In this paper, we concentrate on the mean sound-pressure levels and the intensity fluctuations. Experimental values will be compared with results of numerical simulations based on a wide-angle parabolic code.

4pPA13. A fast method for deducing ground-impedance parameters from measurements of excess attenuation spectra. Shahram Taherzadeh (Faculty of Technol., The Open Univ., Milton Keynes MK10 9BS, UK, s.taherzadeh@open.ac.uk)

An efficient numerical method for indirect deduction of ground-surface impedance [J. Acoust. Soc. Am. **105**, 2039–2042 (1999)] is used to deduce surface impedance of absorbing grounds and to obtain estimates for the effective flow resistivity of the ground. One-parameter and two-parameter impedance models are used for the parameter-fitting purposes. It is shown that the two-parameter model can be modified to give a better agreement with the measured data at grazing angles.

4pPA14. The estimation of excess attenuation of acoustic beam and its application. Naixian Pan (Dept. of Geophys., Peking Univ., Beijing 100871, PROC)

A theory of the excess attenuation for a conical beam is presented. The equation of attenuation coefficient shows the attenuation contributed by the item of C_v square is larger than that of C_t square about five times at least in the atmosphere. The theory shows that a significant difference of excess attenuation between emitting wave and back scattering on the same path. Excess attenuation and absorption both are taken into account in

sodar data processing and the C_t square measured by sodar has a reasonable agreement in statistics with that measured by a filament thermometer. Based on the quantitative measurement of C_t square, sodar was used in an experiment whose aim was estimating the seeing quality of a candidate observatory in China in 1997. The seeing measured by sodar has a better correlation with the seeing measured by a differential image motion monitor in our experiment than that in other experiment in Russia (1988), in which the attenuation of sound had not been considered. The value of the excess attenuation is estimated to be the same order as the absorption in a typical clear day [E. H. Brown and S. F. Clifford, J. Acoust. Soc. Am. **60**, 788–794 (1976)].

FRIDAY AFTERNOON, 2 JUNE 2000

STATE ROOM, 2:30 TO 4:30 P.M.

Session 4pPP

Psychological and Physiological Acoustics: Auditory Modeling and Physiology

Neal F. Viemeister, Chair

Department of Psychology, University of Minnesota, 75 East River Road, Minneapolis, Minnesota 55455

Contributed Papers

2:30

4pPP1. Directional microphone theory. Patrick C. Fournier (Prescription Design Hearing Aid Ctr., 2311 Tenth Ave. S., Great Falls, MT 59405)

An alternate to the cochlear amplifier and resonant tectorial membrane models, this theory models the cochlea as a type of directional microphone. The main source of cochlear nonlinearity is a mechanical canceling of more intense signals from the interaction of the cochlea's various diaphragms. These interactions allow both the inner and outer hair cells to focus a wide range of environmental amplitude into the relatively invariant encoded signal that is sent to the auditory brain. This cancellation of intense signals allows concurrent weak signals in the environment to be perceptively enhanced as compared to the more intense concurrent signals. This model explains Yates dual-wave mode in the cochlea, how and why the outer hair cells are stimulated at low intensities, and how the cochlea performs a fast Fourier transform. Loudness growth charts recalculated in pascals clearly show the enhanced perception of soft signals in comparison to intense ones. Detection of pure-tone signals delivered in a sound field of narrow-band white noise showed equivalent intensity differences for both impaired and nonimpaired ears for perceived loudness levels when calculated in pascals. Expected focus or enhancement of normal soft to intense sound perception was recreated artificially for the impaired ears.

2:45

4pPP2. Modeling across-channel processing of amplitude modulation. Stephan Ewert, Jesko Verhey, Birger Kollmeier (Carl von Ossietzky Universitaet Oldenburg, AG Medizinische Physik, D-26111 Oldenburg, Germany, se@medi.physik.uni-oldenburg.de), and Torsten Dau (Boston Univ., Boston, MA 02215)

In typical modulation detection experiments, a sinusoidal signal modulation is applied to broadband noise as the carrier. In such conditions, the carrier envelope at the output of different peripheral auditory channels might be considered as nearly uncorrelated while the signal modulation is coherent across channels. Models of auditory processing often combine signal information optimally in terms of signal detection theory, assuming independent "observations in different auditory channels." On the other

hand, various experiments demonstrated that modulations applied to spectrally separated carriers can interfere, indicating that (prior to the decision stage) amplitude modulation is not processed independently across carrier frequency in these conditions. In the present study, modulation detection and modulation masking conditions were examined where the signal and the masker modulation were applied either to the same carrier or to spectrally separated carriers. Simulations are presented based on an envelope power spectrum model [Ewert and Dau, J. Acoust. Soc. Am. (submitted)], which calculates the envelope power in the transfer range of a modulation filter tuned to the signal-modulation frequency. Within the model, different strategies of combining across-channel envelope information, as well as the influence of peripheral filtering and compression, are considered. [Work supported by the Deutsche Forschungsgemeinschaft.]

3:00

4pPP3. Modeling the temporal response of auditory nerve fibers with a leaky integrator. Kyle T. Nakamoto, Rita L. Smith, and Bruce G. Berg (Dept. of Cognit. Sci., Univ. of California, Irvine, CA 92612)

A leaky integrator model (e.g., bandpass filter, followed by half-wave rectification and a low-pass filter) describes the temporal response of auditory nerve fibers presented with multicomponent, periodic sounds. It is shown that the spectra of the leaky integrator output closely resemble the spectra of period histograms obtained from published studies. Five parameters define the model: a time constant defining the cutoff frequency of the low-pass, one-pole filter, a parameter representing internal noise, and three parameters describing the initial asymmetric roex(p) filter (center frequency, low- and high-frequency skirts). For high intensities, an algorithm was introduced to represent saturation of the fiber. In most cases, the model accounted for more than 90% of the variance. Parameter estimates are consistent with known characteristics of auditory nerve fibers: (1) the low skirt of the initial filter was shallower than the high-frequency skirt, (2) increases in intensity were accompanied by shifts in the center frequency of the initial filter, and (3) bandwidth of the initial filter increased with intensity. With respect to the nonlinearity, square wave and full wave rectification were found to be implausible alternatives.

4pPP4. The phases of auditory-nerve fiber responses to tones: Dependence on stimulus frequency and intensity. Andrei N. Temchin and Mario A. Ruggero (Hugh Knowles Ctr. and Inst. for Neurosci., Northwestern Univ., 2299 North Campus Dr., Evanston, IL 60208-3550)

As stimulus intensity increases, the phases of responses to tones of squirrel monkey auditory-nerve fibers [ANFs; Anderson *et al.*, J. Acoust. Soc. Am. **49**, 1131–1139 (1971)] and guinea pig inner hair cells [IHCs; Dallos, Hearing Res. **22**, 185–198 (1986)] undergo lags and leads, respectively, at frequencies lower and higher than the characteristic frequency (CF) and remain relatively constant at frequencies near CF. A similar pattern holds for basal sites of the basilar membrane (BM) in several species. In contrast, tectorial-membrane (TM) vibrations at apical sites of the chinchilla cochlea exhibit phase shifts of opposite polarity [Rhode and Cooper, Aud. Neurosci. **3**, 101–121 (1996)]. To explore this inconsistency between neural and mechanical behavior at apical cochlear regions, we investigated the phases of responses to tones of ANFs recorded in deeply anesthetized chinchillas. For stimulus intensities lower than 70 dB SPL, responses of most ANFs (regardless of CF) exhibited phase shifts similar to those observed in squirrel monkey ANFs, guinea pig IHCs, and basal BM sites in several species. Thus, the inconsistency between neural and TM mechanical data cannot be ascribed to species differences or a dependence of phase behavior on CF. [Work supported by NIH Grant DC-00419.]

3:30

4pPP5. Insight into the diagnosis of conductive hearing loss using wideband acoustic reflectance measurements of surgically modified human temporal bones. Antonio Miller (Mimosa Acoust., Inc. and the Penn State Grad. Prog. in Acoust., 382 Forest Hill Way, Mountianside, NJ 07092, ajm249@psu.edu)

There is mounting evidence to suggest that wideband acoustic impedance/reflectance measurements will aid in the clinical diagnosis of various middle ear pathologies. In preparation for clinical trials, wideband pressure reflectance measurements (100 Hz to 8 kHz) were made looking into the ear canals of three human temporal bones which were surgically modified to simulate five possible pathological conditions [unpublished data of J. Allen, G. Ball, S. Nishihara, and R. Goode]. The experimental procedure, data, and analysis are presented for the different modifications. The unmodified, air-filled tympanic cavity data characterize the baseline losses for the entire system. A fluid-filled tympanic cavity characterizes the effect of fluid loading on the tympanic membrane, simulating the condition of acute and chronic otitis media (middle ear infection). The fixed stapes condition simulates the effect of otosclerosis (calcification of the stapes footplate). Disarticulated stapes and the severed stapedius tendon conditions are both symptoms associated with chronic otitis media or cholesteatoma (middle ear cysts). Finally, the fixed malleus condition can be a warning sign of otitis media, but can easily be misdiagnosed as otosclerosis. It is hoped that this work will provide a physical basis for the interpretation of wideband pressure reflectance data collected in future clinical trials.

3:45

4pPP6. Is hearing all cochlear (revisited)? Frequency tuning and intensity thresholds in acoustically stimulated, myogenic vestibular-evoked potentials (MVEP). Neil P. McAngus Todd (Dept. of Psych., Univ. of Manchester, Manchester M13 9PL, UK, todd@fs4.psy.man.ac.uk), Fred W. Cody, and Jon Banks (Univ. of Manchester, Manchester M13 9PL, UK)

There has amassed considerable evidence that the sacculus, an organ of hearing in lower vertebrates, has retained some acoustic sensitivity throughout phylogeny. In humans, myogenic vestibular-evoked potentials

(MVEPs) may be obtained from motoneurons innervated by the vestibulo-spinal tract, particularly from the cervical region of the spinal cord. The MVEP has been studied principally as a noninvasive clinical tool for evaluation of normal otolith vestibular function, since traditional nystagmographic methods only really assess canal function. However, we have been interested in what “auditory” function acoustic sensitivity sacculus may have in the perception of sound. The results of a series of experiments are reported which may be summarized as follows. (1) MVEP has a frequency tuning [Todd, Cody, and Banks, Hear. Res. (in press)], with a maximum response frequency between 300 and 350 Hz and a band-width of about 3 octaves. (2) MVEP can be obtained to natural acoustic stimuli [Todd and Cody, J. Acoust. Soc. Am. **107**, 496–500 (2000)], such as dance music, above about 90 dB SPL. (3) MVEP can be obtained to continuous sounds. Given the above then, there are a number of possibilities where acoustic sensitivity of the sacculus may play a role in human perception.

4:00

4pPP7. Acute effects of nicotine on physiological responses of the auditory systems of nonsmokers. Ashley W. Harkrider (Dept. of Audiol. and Speech Pathol., Univ. of Tennessee, 457 S. Stadium Hall, Knoxville, TN 37996, aharkrider@yahoo.com) and Craig A. Champlin (Univ. of Texas, Austin, TX 78712)

The role of cholinergic mechanisms in the auditory system was investigated by assessing the acute effects of nicotine, a cholinergic drug, on aggregate responses within the auditory pathway. In a single-blind procedure, evoked potentials (auditory brainstem, middle-latency, long-latency, steady-state responses) and electroencephalograms were measured from 20 normal-hearing, nonsmokers (10 male) under two conditions (nicotine, placebo). After the drug session, plasma tests indicated a subject's nicotine concentration. Results indicated that transdermal administration of nicotine to nonsmokers does appear to acutely affect the transmission of acoustic information. The measures most affected by nicotine administration likely originate from the midbrain and cortex and are believed to involve processing of auditory information related to vigilance and the screening of sensory input. The effects of nicotine on evoked potentials or the electroencephalogram do not appear to be dependent on the sex of the subject or the hemisphere from which the responses are measured. [Work supported by AAA and NIDCD.]

4:15

4pPP8. Standing waves in the brain and bone conduction hearing. Martin Lenhardt, Alan Madsen, James Genova, and Howard Neal (Prog. in Biomed. Eng., Virginia Commonwealth Univ., Richmond, VA 23298-0168, lenhardt@hsc.vcu.edu)

Bone conduction hearing studies have a long history; however, in the last 25 years the topic of bone conduction hearing has been virtually abandoned. Delivering sound, including speech, to the head has certain advantages in high-noise communication. Difficulty arises in localization since bone conduction is a binaural perception, at least for tones under 10 kHz. At higher frequencies a robust lateralization of tones was observed for stimuli of 20 digital tones in the range of 4–18 kHz and 20–34 kHz. Inverting the listener's body altered perception and transmission by bone conduction. Some tones in areas of resonance were altered in amplitude when the body was inverted, presumably due to increased cerebral spinal fluid pressure. Modeling of the skull as a sphere with rigid boundary conditions generally predicted direct observations. These findings also account for the observation of perception jumping from ear to ear with changes in frequency during ultrasonic bone conduction listening.

Session 4pSA**Structural Acoustics and Vibration: Tire Vibration**

J. Stuart Bolton, Chair

*School of Mechanical Engineering, Purdue University, 1077 Ray W. Herrick Laboratories,
West Lafayette, Indiana 47907-1077****Invited Papers*****2:00****4pSA1. Wave-number domain representation of tire vibration.** J. Stuart Bolton and Yong-joe Kim (1077 Ray W. Herrick Labs., School of Mech. Eng., Purdue Univ., West Lafayette, IN 47907-1077)

In the work to be described here, wave-number decomposition techniques have been used to study tire vibration. In the experimental component of this work, a tire was driven radially at a point on its treadband. Measurements of the resulting radial treadband vibration were made at approximately 200 points around the treadband circumference by using a laser Doppler velocimeter. From an inspection of the resulting space-frequency data, it was possible to identify the frequency ranges in which the tire responded either modally or nonmodally. Further, by performing a circumferential wave-number decomposition of the space-frequency data, the propagation characteristics of the wave types that contributed to the response of a tire could be determined. It was observed that a small number of circumferentially propagating waves combine to control the response of a tire in both low- and high-frequency ranges. The cut-on and propagation characteristics of these waves are closely related to propagating waveguide modes of flexural wave-bearing systems. Together, these results are consistent with a curved waveguide model of a tire. It is therefore suggested that a membrane-like waveguide model combined with a propagating wave representation of the tire vibration may offer an efficient way of representing tire vibration.

2:30**4pSA2. Two-dimensional experimental analysis of tire vibration.** Jacob Klos, Robert J. Bernhard, and Paul Cvar (1077 Ray W. Herrick Labs., School of Mech. Eng., Purdue Univ., West Lafayette, IN 47907-1077)

This paper presents an experimental analysis of the velocity response of a treadless tire in the two-dimensional wave-number domain. The out-of-plane velocity response was measured on a stationary tire using a Doppler laser vibrometer for a two-dimensional 2-cm by 2-cm grid of points over the frequency band from 70 to 1500 Hz. From the data, three distinct regions of behavior can be identified: simple standing wave patterns, waveguide behavior, and nearly cylindrical wave propagation. Using 2-D wave-number reduction techniques, the measured spatial velocity distribution at each frequency was transformed into 2-D wave-number space. Different wave types in the tire carcass can be identified including the dominant flexural and in-plane waves (due to Poisson's strain). At frequencies below the ring frequency of the tire cross section, a qualitative comparison is made to published results for 2-D dispersion curves of cylinders. At higher frequencies, a comparison is made in the wave-number domain, to a coupled plate model of the tire carcass and models of wave behavior in inhomogeneous plates. A qualitative explanation of the behavior of the tire is identified from the comparisons.

3:00**4pSA3. A one-dimensional wave model for a pneumatic tire.** R. J. Pinnington (Inst. of Sound and Vib. Res., Univ. of Southampton, Highfield, Southampton SO17 1BJ, UK, rjp@isvr.soton.ac.uk)

A wave equation for one-dimensional waves out of plane waves of an infinite tire belt are derived. The dynamics of the belt are regarded as those of a Timoshenko beam under tension, that is, uniformly supported by the sidewall stiffness. The sidewall is modeled here as a static stiffness controlled by the internal air pressure and the geometry of the tire cross section. The fourth-order differential wave equation is solved yielding two wave numbers as a function of frequency. These describe all the wave types of the belt. Below about 100 Hz the sidewall stiffness controls the deformation and there is only motion local to the excitation. At higher frequencies

waves propagate around the belt with four different controlling mechanisms, tension, bending, shear, and longitudinal stiffness. The tire is made circular by the application of the boundary conditions, which is mainly a uniform force applied on a line across the width. The input and transfer mobilities of the belt are output, showing standing wave behavior below about 500 Hz, above which there are only propagating waves and near-field waves as on an infinite belt.

3:30–3:45 Break

3:45

4pSA4. Modeling of tire/road interaction. Wolfgang Kropp (Dept. of Appl. Acoust., Chalmers Univ. of Technology, S-41296 Gothenburg, Sweden)

In order to reduce tire/road noise a deep understanding of the noise generation mechanisms involved is required. The key to reach this understanding is a correct model of the interaction between a tire and a rough road. This includes the modeling of normal and tangential contact between tire and road and the modeling of the wave propagation in the tire. A tire model is presented, which includes the radial, lateral, and tangential motion of the tire. The tire is modeled as a double-layered plate where the two layers are formulated by the general field equations including external tension. The tire model also takes into account the local deformation of the rubber surface due to external forces. This allows for solving directly the contact problem applying Green's functions (i.e., impulse response functions) of the vibrating body. The model also takes into account the coupling between adjacent points at the surface of the tire. Numerical results will be presented for both the tire model and contact model.

Contributed Papers

4:15

4pSA5. Tire tread block noise experiments including effects of pavement porosity. Sarah Jones, Dean Karnopp, and Nesrin Sarigul-Klijn (Transportation Noise Control Ctr. (TNCC), Dept. of Mech. and Aero. Eng., Univ. of California at Davis, Davis, CA 95616-5294)

A tread block impact device has been developed to record the pressure pulse signature generated by a single tread block impacting a surface. High-speed video images are used to relate the impact progression with the instantaneous acoustic signal. The acoustic pressure pulses generated by this device can be associated with the changing volume velocity of the air displaced as the tread block approaches the surface. A simple monopole source can be used to predict the acoustical pulse that is radiated from impact. Experimentation and theoretical calculations will be correlated, as well as discussion of previous research. Noise due to a sequence of blocks on a tire can be inferred from single pulse results and the sound power level can be examined as a function of the speed of the tire. By using impact surfaces with known area removal and, hence porosity, experimental relationships between the pressure pulse signature and impact surface characteristics can be determined. By evaluating the frequencies generated by tread block impacts and the effect of porosity on noise reduction, tuned pavements which could attenuate the most irritating frequencies could be developed. The theoretical and experimental models may prove useful in suggesting means of reducing this component of traffic noise.

4:30

4pSA6. Predicted ground effects for the sound fields near a reflecting wall. Raymond S. H. Tang and Kai Ming Li (The Dept. of Mech. Eng., The Hong Kong Polytechnic Univ., Kowloon, Hong Kong, mmkml@inet.polyu.hk)

In many environmental noise prediction schemes, an empirical correction factor of 2.5 or 3.0 dB(A) is to be added for the calculation of noise 1 m in front of a reflecting facade. In this paper, theoretical and experi-

mental studies have been conducted to examine the validity and accuracy of such approximations. The theoretical analysis is based on the representation of the sound waves in their integral forms. By imposing the impedance boundary conditions on the absorbing ground and the vertical wall, it is possible to extend the classical Weyl–Van der Pol formula for inclusion of the effect of a reflecting facade. As expected, the analytical solution is composed of four terms (a source and three image sources) as a result of the absorbing ground and the reflecting facade. Experimental measurements have been conducted indoors to validate the theoretical predictions. It has been demonstrated that the simple empirical correction factor is not adequate in most cases for the assessment of the facade noise levels above an absorbing ground. [Work supported by the Hong Kong Polytechnic University.]

4:45

4pSA7. Efficiency of barriers for transport noise insulation in the near-field region. S. A. Rybak (N. N. Andreyev Acoust. Inst., Moscow 117036, Russia, rybak@tarsks.msk.su) and S. A. Kostarev (Lab. of Acoust. and Vib. Tunnel Assoc., Moscow 107217, Russia)

The efficiency of noise-insulating barriers needs to be carefully calculated in the near-field region where the Fresnel's parameter is not valid. The correct calculative method for the diffraction across the barrier is developed, taking into account both the resonance absorption at the boundary of the barrier and its thickness for the low-frequency range. The comparison of the calculus and experiments for the barriers arranged along the transport lines was made and the agreement is sufficiently good.

Session 4pSC

Speech Communication: Modeling Speech Processing and Production (Poster Session)

Elliot Saltzman, Chair

Department of Physical Therapy, Boston University, 635 Commonwealth Avenue, Boston, Massachusetts 02215

Contributed Papers

All posters will be on display from 2:00 p.m. to 5:00 p.m. To allow contributors an opportunity to see other posters, contributors of odd-numbered papers will be at their posters from 2:00 p.m. to 3:30 p.m. and contributors of even-numbered papers will be at their posters from 3:30 p.m. to 5:00 p.m.

4pSC1. Waveguide modeling as a tool for fitting a hearing aid. Grzegorz Szwoch, Andrzej Czyzewski, and Bozena Kostek (Sound Eng. Dept., Tech. Univ. of Gdansk, Poland, kid@sound.eti.pg.gda.pl)

A physical modeling method is used to develop a model of the acoustical elements of a hearing aid and to examine its properties. This model enables one to design an acoustical element having desired frequency characteristics, thus providing the acoustical system well fitted to the individual patient's needs. The acoustical system of the hearing aid is treated as a waveguide—a long and narrow tube. Modeling methods similar to the ones applied to human vocal tract synthesis are used. The system is divided into several cylindrical sections. Physical phenomena present in the acoustical system (wave reflections, frequency-dependent damping, etc.) can be implemented directly in the model. By adjusting the parameters of the model, one is able to examine acoustical properties of various systems, having different shapes and dimensions. This method may be helpful for the audiologists during the process of hearing aid fitting.

4pSC2. Using Markov models to assess articulation errors in young children. H. Timothy Bunnell, Debra M. Yarrington, and James B. Polikoff (Speech Res. Lab, duPont Hospital for Children, 1600 Rockland Rd., Wilmington, DE 19803)

Digital recordings of children producing the names "Rhonda" and "Wanda," and/or "Toto" and "Coco" were made using the microphone input to a Toshiba laptop computer (16-bit samples, 22 050-kHz sampling rate) with an AKG C410/B head-mounted condenser microphone. These names were associated with animated characters in a mock video game running on the laptop under the control of a Speech Language Pathologist. The children, ranging in age from four to six years, were undergoing speech therapy at the Alfred I. duPont Hospital for Children for one or both of two common articulation errors: /w/ substituted for /r/; and/or /t/ substituted for /k/. The initial segment in each recorded utterance was classified by laboratory staff as either r/w or t/k, and assigned a goodness rating. Discrete Hidden Markov phoneme Models (DHMMs) trained using data recorded from normally articulating children were then used to classify the same utterances and results of the automatic classification were compared to the human classification. Results indicate that appropriately trained DHMMs can provide accurate classification of utterances from children in speech therapy. This technology could support articulation drill on home computer systems as an adjunct to speech therapy. [Work supported by Nemours Research Programs.]

4pSC3. Speech intelligibility and acoustic characteristics of Mandarin speakers with cerebral palsy. Jing-Yi Jeng, Gary Weismer, and Ray Kent (Dept. of Communicative Disord. and Waisman Ctr., 1975 Willow Dr., Univ. of Wisconsin-Madison, Madison, WI 53706)

Reduced speech intelligibility is often associated with individuals having cerebral palsy. The purpose of the present study was to investigate the speech intelligibility and acoustic characteristics of three subtypes of cerebral palsy (athetoid, spastic, mixed) among adult speakers of Mandarin. A Mandarin word intelligibility test was constructed using phonetic contrasts known from previous studies to be especially sensitive to the dysarthria associated with cerebral palsy. These included tonal contrasts, whose relationship to speech intelligibility deficits in persons with cerebral palsy is not well understood. Acoustic measures consisted of formant frequencies, segment durations, and measures of tone. Preliminary results show that speech intelligibility of athetoid speakers is often worse than that of spastic speakers, and that there are also reliable acoustic differences between the groups. A multidimensional scaling analysis of the perceptual results and regression models of the relationship of the acoustic to the perceptual variables will also be reported. [Work supported by NIH DC00319.]

4pSC4. Compensatory articulatory behavior to immediate structural perturbation of the oral cavity. Masaaki Honda, Tokihiko Kaburagi, and Akinori Fujino (NTT Commun. Sci. Labs./CREST JST 3-1 Morinosato-Wakamiya, Atsugi-shi, Kanagawa 2430198, Japan, hon@idea.brl.ntt.co.jp)

Speech compensation to the structural perturbation of the oral cavity was investigated by using an artificial palate whose thickness in the alveolar region is dynamically changeable. The artificial palate was constructed with a base and thin rubber attached to the base, and the rubber balloon was dynamically controlled by means of an external air flow. The articulatory behavior with the structural perturbation was recorded using an electromagnetic articulographic system. A set of VCV syllables uttered in the carrier sentence, containing the consonants /s, sh, t, n/ and the vowels /a, i, u/, was used for the test. Unexpected perturbation of the palate shape, a change in thickness of 1 to 4 mm, caused significant speech articulation errors for the fricative consonants, fictive consonants /s, sh/ changed to the stop consonants /t, ch/. On the contrary, the stop consonant /t/ was very robust to unexpected perturbation. After five to seven speaking trials with the modified palate shape with or without audio feedback, these articulation errors disappeared and the subject could produce correctly perceived

consonants. Then, the articulatory compensation in the alveolar region significantly appeared in the tongue tip position but was small in the jaw and the lip positions.

4pSC5. On the correlation between orofacial movements, tongue movements, and speech acoustics. Jintao Jiang, Abeer Alwan, Patricia Keating (Depts. of Elec. Eng. and Linguist., UCLA, Los Angeles, CA 90095), and Lynne Bernstein (House Ear Inst., Los Angeles, CA 90057)

This study is a first step towards quantifying the relationship between external orofacial movements, tongue movements, and the acoustics (AC) of speech sounds. The database analyzed consisted of six sentences and 23 CV syllables spoken by a male and a female; each utterance was repeated eight times. A Qualysis (optical motion capture system) and an EMA (electromagnetic articulography) system were used to characterize orofacial and tongue movements, respectively; principal component analysis was applied to both types of data to reduce the dimensionality. Acoustic features were represented by linear spectral pairs. The datastreams were aligned using a synchronization technique, and a multilinear regression technique was applied to quantify the correlation between them. The results are in general agreement with the study of Yehia *et al.* [AVSP-1997, pp. 41–44], which examined two sentences by a male talker, but also shows that correlations are highly context and speaker dependent. On average, the highest correlations were found between the optical (orofacial) and EMA (tongue) features. The results and their implications will be discussed. [Work supported in part by NSF KDI award 9996088.]

4pSC6. Planning and forming articulatory movements from phoneme-specific representatives. Takesi Okadome and Masaaki Honda (NTT CS Labs./CREST, JST, Morinosato-wakamiya 3-1, Atsugi-shi, Kanagawa pref 243-0198, Japan, houmi@idea.br1.ntt.co.jp)

A computational model of sensorimotor planning of speech production is described. The model uses phoneme-specific representatives to predict the trajectories of articulatory movements for continuous speech. As phoneme-specific representatives, the model uses acoustics for the vowels and articulatory features for the consonants. The acoustics are represented by LPC cepstral segments. The articulatory features of the consonants are those of the kinematics of the articulators; that is, the position and velocity of the articulator. The model first recovers the kinematic features from the acoustic segment for the vowels using the acoustic-to-articulatory inverse mapping method [S. Suzuki *et al.*, Proc. ICSLP 6, 2251–2254 (1998)]. Using the kinematic features of the vowels together with those for the consonants, the model determines the minimum-acceleration trajectory that coincides with the extremum of the time integral of the square of the magnitude of acceleration of the articulator [T. Okadome *et al.*, Proc. IEEE SMC 2, 469–474 (1999)]. The model predicts both the qualitative features and quantitative details experimentally observed.

4pSC7. Demonstrating effects of parameter dynamics on gestural timing. Elliot Saltzman (Haskins Labs., New Haven, CT and Dept. of Physical Therapy, Boston Univ., 635 Commonwealth Ave., Boston, MA 02215, esaltz@bu.edu) and Dani Byrd (Haskins Labs., New Haven, CT and USC, Los Angeles, CA 90089-1693)

Speech movements have been modeled as goal-directed gestures whose activation-gated dynamics are those of critically damped point attractors. Recent work by ourselves and others has suggested that gestural parameter values are not fixed but rather vary as a function of a gesture's

position in the word and phrase. For example, lowering of gestural stiffness values has been hypothesized to underlie gestural lengthening adjacent to phrasal boundaries. Additionally, evidence exists suggesting that prosodic structure affects the relative timing between gestures. In earlier work, we have hypothesized that slowing at phrase boundaries can be implemented using nonconstriction gestures called pi-gestures (prosodic gestures) that act on the parameter values of all concurrently active constriction gestures in proportion to the activation level of the pi-gesture. In order to understand the potentially complex consequences of these parameter dynamics, we present a series of simulations involving different temporal arrangements of two constriction gestures coordinated with one another and with a prosodic gesture. The within- and between-gesture temporal patterning of the constriction gestures will be evaluated under the influence of pi-gestures of greater and lesser strength (activation) and duration. Simulation results will inform interpretation of natural kinematic articulatory data collected in movement tracking experiments. [Work supported by NIH.]

4pSC8. Kinematic factors underlying timing in connected speech. Stephen M. Tasko (Army Audiol. and Speech Ctr., Walter Reed Army Medical Ctr., Washington, DC 20307-5001) and John R. Westbury (Univ. of Wisconsin, Madison, WI)

Current understanding of the kinematic factors underlying speech timing has been largely derived from short, simple speech tasks such as citation words. However, timing analyses based upon acoustic segment durations have shown that timing rules differ for citation words and connected speech. Thus it is not clear how well temporal correlates of citation-style speech movement might generalize to connected speech. A better understanding of the articulatory bases of timing patterns associated with connected speech may aid in the refinement of natural-sounding articulatory-based speech synthesis. This study sought to determine temporal correlates of articulator fleshtime movements in orally read speech recorded from 20 normal adult speakers of American English. Position histories associated with the tongue blade, tongue dorsum, jaw, and lower lip were segmented into series of movement strokes, and measures of distance, maximum speed, and duration were determined for each stroke. Canonical correlation analysis showed that for all articulators, these three measures were highly interrelated, and that 80% of the variance in movement duration could be accounted for by movement distance and maximum movement speed. These results suggest the possibility of predicting articulator timing based on a small number of kinematic measures. [Work supported by NIH Grant Nos. DC00820, DC03723, and DC03659.]

4pSC9. Articulatory kinematic and formant frequency changes across vocalic nuclei: Some preliminary analyses. Gary Weismer and Jing-Yi Jeng (Dept. of Communicative Disord. and Waisman Ctr., 1975 Willow Dr., Univ. of Wisconsin-Madison, Madison, WI 53706)

There are few data relating changes in articulatory position to changes in vocalic formant frequencies. This kind of information is useful for speech-production theory in general and specifically in understanding the relationship in neurogenic speech disorders between movement effects and resulting speech-intelligibility deficits. The purpose of this study was therefore to assemble a fairly large kinematic/acoustic database on normal speakers, for a few words that have proven useful in our studies of the intelligibility deficit in dysarthria. Data were part of the x-ray microbeam database, and for the present study included the words "hail," "shoot," and "coat." The position histories of four tongue pellets and two lip pellets as well as the time histories of the first and second formant frequencies were studied for 45 young adult speakers. The F2 transition extent (change in F2 across a defined time interval) was significantly correlated with some combination of the pellet time histories for each word,

and the magnitude of the correlation showed clear gender effects. Results will be discussed in terms of the linkage between planar kinematic data and acoustic data in predicting speech-intelligibility deficits. [Work supported by NIH DC03723 and DC000820.]

4pSC10. Effects of central dopaminergic stimulation by apomorphine on producing sentence focus in speech production in Parkinson's disease—a preliminary report. Q. Emily Wang (Dept. of Commun. Disord. and Sci., Rush Univ., Rush-Presbyterian-St. Luke's Medical Ctr., Chicago, IL 60612), Yi Xu (Northwestern Univ., Evanston, IL 60208), Katie Kompoliti, and Christopher G. Goetz (Rush Univ., Rush-Presbyterian-St. Luke's Medical Ctr., Chicago, IL 60612)

It is well documented that when a sentence focus is produced, the f_0 -range of the word under focus is expanded while the f_0 -range of the post-focus words is suppressed. It is also a well-known fact that individuals with Parkinson's disease (PD) exhibit reduced ability in pitch modulation, resulting in a perceptual quality of "monopitch" in connected speech. Although the use of dopaminergic stimulation improves other physical symptoms such as balance and walking in PD, it is unclear whether a similar effect will be found on vocal folds, producing an improvement in patients' ability in pitch modulation. In this study, five non-demented English-speaking individuals with a clinical diagnosis of idiopathic PD (Hoehn and Yahr stage 3 and 4, off) were examined for their ability in pitch modulation in association with Apomorphine, a nonergot dopamine agonist. Two six-word sentences with four sentence focus variations each were used. Subjects randomly produced three repetitions of each focused version at baseline and two clinical conditions: injected with placebo or apomorphine. Perception tests and acoustic analyses were conducted to determine changes in these patients' ability to produce sentence focus. Implications of the findings are discussed in terms of speech motor control.

4pSC11. A quasi-one-dimensional model for speech production. Wei Zhao, Zhaoyan Zhang, Steven H. Frankel, and Luc Mongeau (School of Mech. Eng., Purdue Univ., West Lafayette, IN 47907)

A quasi-one-dimensional model for speech production was developed. The unsteady quasi-one-dimensional Euler equations with temporal and spatial area variation were solved using a fourth-order accurate Runge-Kutta scheme for time integration and a sixth-order compact finite-difference scheme for spatial discretization. A model for wall friction was used, and a model for the pressure loss associated with flow separation at the glottis was derived. The code was validated by calculations of steady isentropic flow through a nozzle and an acoustic standing wave in a tube with uniform cross-sectional area. A simple geometry was then used to simulate the acoustic wave generated in the vocal tract. The results are in qualitative agreement with data from experiments obtained using a dynamic model of the larynx. [Work supported by NIH DCO 3577-02, RO1 grant from NICDC.]

4pSC12. Articulatory speech synthesis and the analysis of boundary conditions. Jun Huang, Stephen Levinson (Univ. of Illinois, 405 N. Mathews Ave., Urbana, IL 61801, jhuang@ifp.uiuc.edu), Scott Slimon, and Don Davis (Electric Boat Corp., Groton, CT 06340)

In this work, a computational fluid dynamics (CFD) approach was used to model the unsteady fluid flow in idealized human vocal tracts. The speech waveform synthesis was based on the slightly compressible Reynolds-averaged Navier-Stokes (RANS) equations. A K-epsilon turbulence model has been used to represent the effects of turbulence. The

vocal tract geometry was determined from the mid-sagittal cut through the three-dimensional vocal tract shape. The moving vocal tract shapes during the speech articulation were estimated using interpolation technique. The excitation signal for the simulation was accomplished by specifying a time-varying area at the inlet to the bottom of the pharyngeal cavity and a sawtooth waveform was used to represent the area variation in time. Different boundary conditions were applied at the inlet of the vocal tract. One is based on specifying the volume velocity and the other is based upon specifying the particle velocity, respectively. We will further analyze the effects of different boundary conditions using an analogy of transmission line T-network. Finally, we will present the waveform and spectrum of some synthesized voiced sounds based on the above CFD approach and boundary conditions. [This work was supported by Motorola and NSF.]

4pSC13. Glottal pressure profiles for a diameter of 0.04 cm. Ronald C. Scherer and Daoud Shinwari (Dept. of Commun. Disord., Bowling Green State Univ., Bowling Green, OH 43403, ronalds@bgnet.bgsu.edu)

Computer models of phonation often rely on aerodynamic equations for flow through the glottis. Usually the aerodynamic equations have come from empirical work with steady-flow models made from hard material. The equations simplify the pressure-flow-geometry relations through the larynx. The model used here (model M5) has 14 pressure taps on the vocal folds to give rather complete pressure profiles. Pressure profiles will be reported for symmetric glottal shapes, a minimal diameter of 0.04 cm, and nine glottal angles (uniform; convergent, and divergent 5, 10, 20, and 40 deg) for transglottal pressures of 3, 5, 10, and 15 cm H₂O. The glottis is rectangular in the anterior-posterior direction, with a glottal length of 1.2 cm. Results indicated that the pressure coefficient for the glottal entrance ranged from 1.09 to 1.60 with an average of 1.33 (compared to van den Berg's 1.375). The value decreased as flow increased for any specific glottal angle. The transglottal pressure coefficient ranged from 1.00 to 1.82 with an average of 1.28. The average value for the uniform glottis was 1.63. Values decreased as flow increased. Pressure profiles will be compared to predictions from current glottal aerodynamic equations. [Work supported by NIH Grant 1R01DC03577.]

4pSC14. Syllable-boundary magnitudes from jaw movement patterns. Bryan Pardo and Osamu Fujimura (Speech and Hearing Sci., The Ohio State Univ., 1070 Carmack Rd., Columbus, OH 43210-1002)

The C/D model [O. Fujimura, J. Acoust. Soc. Jpn. (E) **13**, 39–48 (1992)] describes rhythmic patterns of utterances by amplitude-controlled pulse trains, each pulse representing a syllable or a boundary. The syllable magnitude (pulse height) controls mandibular movement. Assuming a direct proportional relation between the two variables, the syllable magnitude for each occurrence of five, nine, and Pine under various prominence conditions was inferred using Erickson's Pine Street dialogue data [Erickson, Fujimura, and Pardo, Lang. Speech **41**, 395–413 (1998)], which contained repeated corrections of the street address. A computational procedure has been devised to process semi-automatically microbeam data produced by two male and two female speakers. The timing of each syllable pulse was determined using the iceberg method [O. Fujimura, J. Acoust. Soc. Am. **99**]. Syllable durations were determined in proportion to the syllable magnitudes. Boundary magnitudes were represented by the gaps among syllable durations. The syllable and boundary durations thus determined reasonably agreed with acoustic durations. [This work was supported in part by NSF Grant Nos. SBR-951199B, SBR-9809046, and BCS-9977018 (O. Fujimura) and by ATR/MIC, Japan.]

4pSC15. Effect of vocal fold tissue shear properties on phonation threshold pressure in a physical model of the larynx. Roger Chan and Ingo Titze (Natl. Ctr. for Voice and Speech, Dept. of Speech Pathol. and Audiol., The Univ. of Iowa, Iowa City, IA 52242, chan@shc.uiowa.edu)

Previous analytical studies have shown that phonation threshold pressure (P_{th}), the minimum subglottal pressure required to initiate and sustain vocal fold oscillation, is directly related to the viscous shear properties (dynamic viscosity or viscous shear modulus) of the vocal fold mucosa in vibration [Titze, J. Acoust. Soc. Am. **83**, 1536–1552 (1988); Chan, Ph.D. dissertation (1998)]. This relationship was verified empirically with a physical model of the larynx [Titze *et al.*, J. Acoust. Soc. Am. **97**, 3080–3084 (1995)], but no biological materials were tested in that study. This study attempted to establish the effects of biologically based tissue shear properties on P_{th} . Biomaterials with varying viscoelastic shear properties (abdominal subcutaneous fat, 0.01%–1.0% hyaluronic acid, and hyaluronic acid mixed with fibronectin) were incorporated into the artificial vocal fold mucosa of the physical model and self-sustained oscillation was induced. Results showed that no stable phonation could be established for fat, while higher P_{th} were consistently observed for higher concentrations of hyaluronic acid (with or without fibronectin), in correlation with their differences in viscous shear properties. These findings support the notion that hyaluronic acid could potentially be an optimal surgical bioimplant for the vocal fold mucosa, because it likely facilitates the ease of phonation. [Work supported by NIDCD.]

4pSC16. Measuring vocal quality with speech synthesis. Bruce R. Gerratt and Jody Kreiman (Div. of Head/Neck Surgery, UCLA School of Medicine, 31-24 Rehab, Los Angeles, CA 90095)

Previous research has demonstrated that listeners do not agree well when using traditional rating scales to measure pathological voice quality. Although these findings may indicate that listeners are inherently unable to agree in their perception of such complex auditory stimuli, an alternate explanation implicates the particular measurement method—rating scale judgments—as the culprit. A novel method of assessing quality—listener-mediated analysis by synthesis—was devised to assess this possibility. In this new approach, listeners explicitly compare synthetic and natural voice samples, and change speech synthesizer parameters to create acceptable auditory matches to voice stimuli. Thus this method replaces unstable internal standards for qualities like breathiness and roughness with externally presented, infinitely variable stimuli, thus overcoming one major hypothetical source of disagreement in rating scale judgments. To assess the reliability of this method, listeners were asked to adjust the signal-to-noise ratio for 13 pathological voices so that the resulting synthetic stimuli matched the original voices as well as possible. In a separate experiment, listeners judged the noisiness of the same stimuli using a traditional visual-analog rating scale. The relative reliability and validity of the two methods of quality assessment will be discussed, as will implications for models of voice quality perception.

4pSC17. A comparison of linguistic and pathological breathiness using the LF model. Melissa A. Epstein (Dept. of Linguist., UCLA, 3125 Campbell Hall, Los Angeles, CA 90095-1543, melissae@ucla.edu)

This study investigated the hypothesis that breathiness used contrastively or to enhance a contrast in a language is as “breathy” as breathiness resulting from vocal pathology. Vowel samples were taken from four male and four female native speakers of English and four male and four female speakers with pathologically breathy voices. In the English data, a single pitch period was extracted and analyzed from an allophonically modal sample (following /ʔ/), a vowel-initial allophonically breathy sample (following /h/), and a vowel-medial allophonically breathy sample.

A representative vowel-medial cycle was selected for speakers with vocal pathology. To determine relative breathiness, vowels were inverse filtered and fit with the LF model of the glottal flow derivative; measurements of open quotient, glottal skew, excitation strength (E_e) and the shape of the return phase were made from the LF smoothed pulses. Breathiness occurring allophonically after intervocalic /h/ did not differ significantly on any parameter from the vowels produced by speakers with pathologically breathy voices. E_e proved to be the best measure of relative breathiness. It distinguished the pathological samples from the normal samples (allophonically modal and allophonically breathy together); it also distinguished the allophonically breathy samples from the allophonically modal samples.

4pSC18. Recognition of digit strings in noisy speech with limited resources. Douglas O’Shaughnessy and Marcel Gabrea (INRS-Telecommunications, 900 dela Gauchetiere West, P.O. Box 644, Montreal, QC H5A 1C6, Canada)

Automatic recognition of sequences of spoken digits (e.g., telephone or credit card numbers) can be accomplished with excellent accuracy, even in speaker-independent applications over telephone links. However, even such relatively simple recognition tasks suffer decreased performance in adverse conditions, such as significant background noise or fading on portable telephone channels. If one further imposes significant limitations on the computing resources to be dedicated to a recognition task, then robust, limited-resource speech recognition remains a suitable challenge, even for a vocabulary as simple as the digits. Since connected-digit recognition over telephone lines is a very practical application, the amount of computer resources needed for a given level of recognition accuracy was investigated for different levels and types of acoustic noise. Rather than use a traditional hidden Markov model approach with cepstral analysis, which is computationally intensive and does not always work well under adverse acoustic conditions, simpler spectral analysis was used, combined with a segmental approach. The limited nature of the vocabulary (i.e., ten digits) allows this simpler approach. High recognition accuracy is maintained despite a massive decrease (versus traditional methods) in both memory and computation.

4pSC19. The application of segment-based orthogonal transform in signal bias removal. Wern-Jun Wang (Dept. of Commun., Natl. Chiao Tung Univ., Taiwan, ROC), Eng-Fong Huang (Chunghwa Telecommunication Labs., Taiwan, ROC), and Sin-Horng Chen (Natl. Chiao Tung Univ., Taiwan, ROC)

The iterative technique for signal bias removal (SBR) to minimize the effects of corrupted speech with unknown noisy sources was proposed for robust speech recognition. The SBR method was applied by considering a multiplicative spectral bias or equivalently, an additive cepstral bias. Two problems of this approach can be found. One is that it needs sufficient iterations to attain better results and the other difficulty is the lower accuracy of bias estimation. To overcome these drawbacks, a novel approach of SBR incorporated with orthogonal transform is adopted in this study. Instead of the conventional frame-based process, this method is a segment-by-segment approach. The basic idea of this approach is that the segment constituted by any feature element will be modeled using orthogonal transform coefficients. Owing to the characteristics of orthogonal transform, the high-order coefficients are bias-free and only the zeroth order coefficient is bias-corrupted. The high-order coefficients can therefore be used for bias estimation and the real bias will be deduced from the zeroth-order coefficient. Effectiveness of the method was confirmed by an

experiment of multispeaker noisy continuous Mandarin speech recognition. The significant improvement on computation and recognition rate can be achieved by this method.

4pSC20. Modeling stop-consonant releases for synthesis. Helen M. Hanson and Kenneth N. Stevens (Sensimetrics Corp., 48 Grove St., Somerville, MA 02144-2500 and MIT Res. Lab. Electron., Cambridge, MA 02139, hanson@sens.com)

This study is part of a project leading to rule-based speech synthesis using the Hlsyn synthesizer. In Hlsyn, stop-consonant releases are generated by controlling the time variation of a constriction that is formed by the lips, the tongue blade, or the tongue body. In order to generate a consonant release that is perceptually acceptable and that has acoustic characteristics that match those of normal speech, it was found that the trajectory of the consonant release (cross-sectional area versus time) had to exhibit an initial rapid rise, followed by a delay in which the rise was interrupted, followed finally by a final rise. The burst at the consonant release is generated by the airflow through the constriction during the time that the rise is delayed. A model of the time course of this three-stage release of the articulator has been developed, taking into account the role of the intraoral pressure, the tapering formed by the articulator, and the glottal area. The model shows that the delay between the initial and final area increases progressively longer for labials, alveolars, and velars, in agreement with acoustic data on the burst duration. [Work supported in part by NIH Grant Nos. MH52358 and DC00075.]

4pSC21. Quantization errors in formant estimation. Gautam K. Vallabha and Betty Tuller (Ctr. for Complex Systems & Brain Sci., Florida Atlantic Univ., 777 Glades Rd., Boca Raton, FL 33431, vallabha@walt.ccs.fau.edu)

The locations of formants in a speech signal are usually estimated by computing the linear predictive coefficients (LPC) over a sliding window and estimating the peaks in the spectrum of the resulting all-pole LPC filter. The peaks are estimated either by solving for the roots of the LPC polynomial or by computing its DFT and finding the peaks in the magnitude spectrum. Three different sources of errors in this analysis were investigated using synthesized vowels: (1) F_0 quantization: In addition to the expected result that error increases with F_0 , it was found that the absolute error decreases with higher formant frequencies. (2) LPC+root solving: The location and bandwidth of a peak are usually estimated from the location of a pole on the z plane. It was found that this approximation is invalid if the formants have high bandwidths or if they are either too close or too far from each other. (3) LPC+DFT: In order to compensate for the quantization error introduced by the short-term DFT, a three-point parabolic interpolation scheme is usually employed for better estimation of peak locations. It was found that this compensation scheme is effective only for high-bandwidth formants. [Work supported by NIMH.]

4pSC22. Pathologies not included: Survey of premillennium speech recognition software. Renetta Tull (Dept. of Communicative Disord., Trace R & D Ctr., Univ. of Wisconsin—Madison, 1500 Highland Ave., Madison, WI 53705, tull@waisman.wisc.edu)

An October 1999 survey of several leading speech recognition software manufacturers revealed that the speech technology industry is unprepared to serve users with vocal pathologies in 2000. Companies surveyed had neither prepared for consumers with vocal pathologies nor had im-

mediate plans to alter systems to accommodate the pathological voice by 2002. Examples of vocal pathologies include leukoplakia, carcinoma, hyperfunction, polyps, and nodules. While the aforementioned pathologies involve direct impairment to the vocal mechanism (e.g., swollen pharyngeal mucous membranes, vocal fold polyps), neurological disorders such as Parkinsons disease also cause vocal pathologies (e.g., dysarthrias, spasmodic dysphonia). Some companies are beginning to incorporate natural language processing for consumers with aphasia; however, many of the top producers of speech recognition engines have not considered the pathological voice within their disability market. Acknowledging that the speech technology market is still new, speech recognition software developers note that lack of time and manpower prohibits concentration on groups with special needs. This research examines the feasibility of altering linguistic grammars of templates from healthy, prototypical voices to accommodate the phonetic changes that are observed in the pathological voice. Backus-Naur form and Chomskys syntactic structures are used to modify existing grammar templates.

4pSC23. The perceived pitch of synthesized vowels with alternate pulse cycles. Xuejing Sun (Dept. of Commun. Sci. and Disord., Northwestern Univ., Evanston, IL 60208, sunxj@nwu.edu)

Speech with alternate pulse cycles (APC), also known as diplophonia, double pulsing, etc., often brings problems to pitch determination algorithms. Although this phenomenon has been studied in light of nonlinear theory recently, to date, no study has been done to measure the perceived pitch of speech with APC. In the present study, a perception experiment was conducted to investigate the perceived pitch of speech with APC. Two synthetic vowels, /a/ and /i/, with fundamental frequencies at 140 and 220 Hz, both amplitude and frequency modulated, respectively, were used as stimuli. The modulation levels ranged from 0 to 0.9. The subjects were asked to judge the pitch of modulated vowels in reference to the synthetic vowels without modulation. The results show that the modulation level affects the perceived pitch in a nonlinear manner, and frequency modulation seems to have greater effects on the perceived pitch than amplitude modulation. In addition, different fundamental frequencies have different effects on the perceived pitch across modulation levels, whereas different vowel types appear to make no significant differences. Based on these results, a pitch determination algorithm was developed, and the performance of the algorithm appeared promising.

4pSC24. Improved voice activity detection of noisy speech. Min-sung Koh and Margaret Mortz (School of Elec. Eng. and Computer Sci., Washington State Univ., Spokane, WA 99202-1648, mkoh@eecs.wsu.edu)

This paper reports on the performance of a new algorithm for detecting voice activity in the presence of additive Gaussian noise. The proposed algorithm enhances the noisy speech before determining voice activity. In the speech enhancement, two noise reduction stages are used. The first stage uses Donoho's wavelet-based noise-reduction method. The second stage uses nonlinear spectral magnitude enhancement in the Fourier domain. After signal conversion back to the time domain and renormalization, simple energy thresholds are used to detect the presence of voice activity. Experimental results show that the new method can achieve good performance with preprocessing a speech signal even though a simple energy-based voice-activity detector is used. A new algorithm shows sig-

nificantly better performance when it is compared with the voice-activity detector specified in the ITU standard G.729 Annex B and other methods proposed recently.

4pSC25. Speech processing using the average localized synchrony detection. Ahmed M. A. Ali, Jan Van der Spiegel (Dept. of Elec. Eng., Univ. of Pennsylvania, 200 South 33rd St., Philadelphia, PA 19104, ahm@ee.upenn.edu), and Paul Mueller (Corticon, Inc., King of Prussia, PA 19406)

A new auditory-based speech-processing system based on the biologically rooted property of average localized synchrony detection (ALSD) is proposed. The system is a modification to the generalized synchrony detector (GSD) [S. Seneff, *J. Phonetics* **16**, 55–76 (1988)]. It generates a pseudospectrogram of the speech signal by detecting periodicity, while reducing the response to the individual harmonics of the fundamental frequency and the sensitivity to implementation mismatches. This is achieved without sacrificing the frequency resolution. Hence, it presents a more consistent and robust representation of the formants. The system is evaluated for its formant extraction ability while reducing spurious peaks. It is compared with other auditory-based front-end processing systems in the tasks of vowel, stop and fricative classification on clean speech from the TIMIT database, and in the presence of noise. The results illustrate the advantage of the ALS system in extracting the formants and reducing the spurious peaks, while preserving the frequency resolution. They also indicate the superiority of the synchrony measures over the mean rate in the presence of noise. [Work supported by Catalyst Foundation.]

4pSC26. Relative contributions of passband and filter skirts to intelligibility of bandpass speech: Some effects of amplitude and context. James A. Bashford, Jr., Richard M. Warren, and Peter W. Lenz (Dept. of Psych., Univ. of Wisconsin—Milwaukee, Milwaukee, WI 53201, bashford@uwm.edu)

Recently, Warren and Bashford [*J. Acoust. Soc. Am.* **106**, L47–L52 (1999)] reported over 90% intelligibility for everyday sentences reduced to a 1/3-octave band (center frequency 1500 Hz, slopes 100 dB/octave, peak levels 75 dB). Surprisingly, when the bandpass sentences were partitioned using 2000-order FIR filtering, the rectangular 1/3-octave passbands had only 24% intelligibility, while the filter skirts separated by a 1/3-octave notch had an intelligibility of 83%, despite their severe spectral tilts. The present study found that when monosyllabic word lists were substituted for the sentences, the whole-band intelligibility was 26%, the passband 4%, and the filter skirts 16%. Intelligibility also was measured for 1/3-octave sentences with peak levels ranging from 85 down to 35 dB: The whole band intelligibilities ranged from 90% to 68%, and the filter skirt pairs each had three to four times the passbands' intelligibility. These findings demonstrate the ability to compensate for extreme distortions of the spectral profile of speech. Hence, it is necessary to consider the contribution of frequencies outside the nominal passband as defined by the 3 dB downpoints along the filter slopes when dealing with filtered speech. [Work supported by NIH.]

4pSC27. Implementation of high-quality text-to-speech synthesis for limited domain applications. Sunil R. Shukla and Thomas P. Barnwell III (Ctr. for Signal and Image Processing, Georgia Inst. of Technol., GCATT Bldg., 250 W. 14th St. NW, Atlanta, GA 30318, shukla@ece.gatech.edu)

Though current TTS systems can produce unrestricted, intelligible speech using a limited database of prerecorded sounds, the quality of speech is highly artificial. By using larger speech segments, such as a combination of words and diphones, the number of segment boundaries

are reduced drastically, resulting in high-quality TTS that has a humanlike natural voice. Such a system can be useful for limited domain applications such as driving directions, weather advisories, etc. Experiments were conducted to generate natural sounding speech from words and diphones by modifying the pitch and duration of the units to match the prosody characteristics. The Festival TTS system, developed at the CSTR at the University of Edinburgh, was used to aid in the implementation of the experiments. The method for synthesis is pitch-synchronous residual LP. This method allows for producing smooth transitions between segments. A database of residuals and parameters for segments (words and diphones) is prepared. During synthesis, the segments are concatenated and the desired prosody is matched using time-domain modifications of pitch and duration. Using natural prosody, waveforms were synthesized, resulting in close to natural sounding speech. This demonstrates synthesized speech can sound identical to the original speaker, given the correct prosody.

4pSC28. Robust speech localization using formant phase subtraction. Parham Aarabi and Vaughan Pratt (Gates 4B, Dept. of Comput. Sci., Stanford Univ., Stanford, CA 94305, parham@stanford.edu)

This paper examines the benefits of a context aware time difference of an arrival-based speech localization system. This approach uses information about the formant structure of the speaker in order to compute a more robust time difference of arrival between two microphones. This ability allows microphone arrays to localize speech in environments with signal-to-noise ratios as low as -10 dB, when the noise source is another speaker, and -4 dB, when the noise source is white Gaussian.

4pSC29. Comparing single- and two-channel telephone speech enhancement for elderly hearing-impaired listeners. Amy E. Sheffield, Ashok Krishnamurthy (Dept. of Elec. Eng., The Ohio State Univ., Columbus, OH 43210, sheffield.20@osu.edu), Stephanie Davidson, and Lawrence L. Feth (The Ohio State Univ., Columbus, OH 43210)

Elderly listeners with sensorineural hearing loss have difficulty understanding speech over the telephone. Hearing aids and telephone amplifiers help, but many elderly individuals do not have the access to these devices. One possible solution to the problem is to "preprocess" the speech at the talker before transmitting it to the hearing-impaired listener. The purpose of this research is to compare processing techniques applied at the talker end that can significantly improve the speech understanding of these persons. Intelligibility enhancement of single- and two-channel amplitude compression schemes are compared by objective and subjective means. The articulation index (AI) and the diagnostic rhyme test (DRT) are used as the objective and subjective measures, respectively. In a preliminary study, we used normal listeners with speech that is passed through Moore's [(1999)] hearing loss model. The hearing loss model simulates the effects of loudness recruitment and reduced frequency selectivity as well as elevated thresholds of hearing. In addition, elderly hearing-impaired subjects were tested for a subset of the processing schemes. We present results comparing relative intelligibility of the various processing schemes. [Work supported by a grant from the Ohio State University College of Social and Behavioral Sciences.]

Session 4pUW

Underwater Acoustics, Signal Processing in Acoustics and Acoustical Oceanography: Model Based Processing for Sources in Motion II

Peter G. Cable, Cochair

BBN Technologies, Union Station, New London, Connecticut 06320

William M. Carey, Cochair

*Aerospace and Mechanical Engineering Department, Boston University, 110 Cummington Street,
Boston, Massachusetts 02215*

Chair's Introduction—2:00

Invited Papers

2:05

4pUW1. A review of matched field processing and moving sources. A. Tolstoy (ATolstoy Sci., 8610 Battailles Court, Annandale, VA 22003, atolstoy@ieee.org)

Over the past two decades or so there has been considerable work devoted to matched field processing (MFP), with particular emphasis on source localization. Most of that work has assumed a stationary source, a simpler problem in many ways. However, there has also been work done to investigate and even take advantage of source motion. Examples include: early work by Song (1990) and Song and Baggeroer (1990) based partially on research by Shan *et al.* (1985); work by Zala and Ozard (1991) generalized the Linear and Capon processors for source motion; Tran and Hodgkiss (1991) analyzed 200-Hz data for low speed sources. Alternately, one might consider the source to be fixed while the array moves, introducing synthetic aperture techniques as seen in Stergiopoulos and Sullivan (1989). MFP for a moving source can be a more challenging problem than for the fixed source but also offers new and different techniques with much promise. [Work supported by ONR.]

2:20

4pUW2. Matched field processing with Doppler and the degrees of freedom in an acoustic channel. Arthur B. Baggeroer (MIT, Cambridge, MA 02139, abb@arctic.mit.edu)

Matched field processing which incorporates source motion, or Doppler, has been examined by several investigators. There are two motivations: (i) long integration times improve signal-to-noise ratios, but this leads to signal gain degradation if the source transits a single resolution cell, and (ii) mainlobes interfere constructively and sidelobes destructively. The MFP Doppler processing can be separated into incoherent and coherent methods. In the former a series of short-term MFP outputs are incoherently combined while in the latter a time-varying replica is used. In both methods the number of degrees of freedom (DOF), i.e., the dimension of the representation basis set, in the acoustic signal is a critical parameter since it determines both MFP resolution and sidelobe properties. The DOF are limited by the minimum between number of sensors and the number of modes propagating in the channel. We examine the concept of degrees of freedom in an acoustic channel and the performance of Doppler-based MFP methods. [Work supported by ONR, Code 321US, Undersea Signal Processing.]

Contributed Papers

2:35

4pUW3. Spatial modulation using vertical arrays in an underwater acoustic telemetry channel. Daniel Kilfoyle, James Preisig (Appl. Ocean Phys. and Eng. Dept., W.H.O.I, Woods Hole, MA 02543, dkilfoyle@whoi.edu), and Arthur Baggeroer (MIT, Cambridge, MA 02139)

Multiple propagation paths between sound sources and receivers in an underwater acoustic channel are a common component of underwater acoustic telemetry. The temporal reverberation that results may adversely affect performance measures such as bit error rate and output signal-to-noise (SNR) ratio. With the use of transducer and hydrophone arrays, however, resolvable paths may also be exploited as additional bandwidth

in a manner completely analogous to frequency bandwidth using a parallel channel model. In this discussion, we describe an experiment conducted in Buzzards Bay, MA, that utilized moving, vertical arrays to spatially modulate and demodulate two acoustic communication signals that fully occupied identical time-frequency spaces. Decomposition of the composite shallow water channel into two, parallel communication channels was accomplished solely with spatial degrees of freedom. Performance is compared with a conventional, single signal transmission. In addition, *in situ* channel impulse response measurements are analyzed to explain the success of the spatial modulation experiment. Performance improvements comparable to those yielded by a 3–6-dB SNR increase are achieved without using additional signal power. We will conclude with a brief discussion regarding the challenges as well as potential applications of the spatial modulation technique.

4p FRI. PM

4pUW4. Tracking performance of least-squares algorithms in time-varying ocean channels. James C. Preisig (Dept. of Appl. Ocean Phys. and Eng., Woods Hole Oceanogr. Inst., Woods Hole, MA 02543)

Adaptive least-squares algorithms are widely used to estimate the state of or compensate for the effects of time-varying acoustic channels in the ocean environment. A special case of such time-varying channels is those through which the signals emitted by moving sources propagate. Current methods of characterizing the tracking behavior of these adaptive algorithms are limited to cases where the algorithm accurately models the time-varying behavior of the channel or for which this behavior has a particularly simple form. A new technique is presented which allows the performance of adaptive least-squares algorithms to be characterized for more general classes of channel fluctuations. The technique shows that the performance can be interpreted in terms of model mismatch. New applications enabled by the technique are presented. [Work supported by ONR.]

3:20

4pUW5. Motion compensation for a synthetic aperture sonar. John E. Piper (Coastal Systems Station, Code R21, 6703 W. Hwy. 98, Panama City, FL 32407-7001, piperje@ncsc.navy.mil)

Synthetic aperture sonar is capable of producing high-resolution images. The method typically requires many small corrections to the timing of the data to compensate for motion perturbations. A redundant phase center approach is used to calculate these corrections. This approach uses channels that spatially overlap between pings. This allows construction of a correlation function that can be used to estimate motion-induced time lag differences. To improve the accuracy of this method the correlation function is fit to a quadratic function to find its maximum. Experimental data from a variety of at-sea tests is used in this analysis. Results indicate that significant improvements to the image resolution can be gained with this approach.

Invited Paper

3:35

4pUW6. Adaptive and synthetic aperture processing structures for high-resolution mine-hunting sonars deploying planar or cylindrical arrays. Stergios Stergiopoulos and Amar Dhanantwari (DCIEM, 1133 Sheppard Ave. West, Toronto, ON M3M 3B9, Canada)

The developments reported in this paper have been devoted to minimizing the influence of partially directional noise fields as well as maximizing the array gain of mine-hunting sonar systems by means of signal processing. The effort of this study included the design of a generic beamforming structure that allows the implementation of nonconventional signal-processing techniques in integrated active/passive sonar systems. The nonconventional processing schemes of this study are adaptive and synthetic aperture beamformers that have been shown experimentally to provide improvements in array gain for signals embedded in partially correlated noise fields. The reported results show that synthetic aperture and adaptive processing schemes with near-instantaneous convergence can be implemented in mine-hunting sonars deploying 3D arrays. Moreover, practically realizable angular resolution improvements provided by the nonconventional beamformers are equivalent with those provided by the conventional beamformer of a three-times-longer physical aperture and for broadband FM type of signals. The same set of results demonstrates also that the combined implementation of a synthetic aperture and the subperture STMV adaptive scheme suppresses significantly the sidelobe structure of cw pulses for mine-hunting operations. In summary, the present paper attempts to address characterization of source signals detected by planar or cylindrical arrays of mine-hunting sonar systems with 3D image resolution capabilities.

Contributed Papers

3:50

4pUW7. Exploiting target motion for active matched-field depth estimation. Granger Hickman and Jeffrey Krolík (Duke Univ., Box 90291, Durham, NC 27708)

This paper concerns the use of complex multipath propagation for discriminating the depth of a moving target with a low-frequency active sonar. Although matched-field processing techniques, which exploit full-wave modeling of multipath propagation, have been previously demonstrated in passive sonar applications, their use in active sonar has been precluded by uncertainty in the complex backscattering characteristics of the target. In this paper, a matched-field technique is proposed which exploits changes in the complex moving target return seen between consecutive sonar pings at a horizontal receive array. Because the method exploits only changes in the coherent combination of multipath arrivals, it does not require *a priori* knowledge of the target backscattering characteristics assuming its aspect is only slowly changing from ping to ping. In this paper, simulation results are presented which are modeled after a realistic Mediterranean scenario taken from the Shallow-Water Active Classification Experiment performed in November 1995. Matched-field depth estimation simulation results indicate that target depth accuracies of better than 20% of the channel depth can be achieved at ranges of 10 to 15 km. [Work supported by ONR.]

4:05

4pUW8. Design and development of software tool for high-power sensor array design. P. M. Joseph (Sharjah College, P.O. Box 5398, Sharjah, UAE, pmjosep@hotmail.com)

The transmitting and receiving characteristics of sonar transducer arrays depend on various factors like its geometry, number of elements, source strength, relative phases of the elements, etc. An array with narrow main beam and low sidelobes is desirable for a variety of applications. But, these are conflicting requirements. One of the requirements can be achieved only at the cost of the other. It is the designer's choice to select the parameters depending on the nature of application. Due to mechanical and electrical limitations, the measured array parameters are found to differ from the theoretically formulated ones, especially at higher drive levels. An array design package has been developed and presented in this paper for predicting the optimum array, based on the requirements of the user, taking into account various high-power electrical and mechanical limitations. In order to make the package more user friendly, the array design software has been split into different modules, each carrying out a specific task like array shape selection, element shape selection, power level selection, mode selection, etc.

Session 5aPA

Physical Acoustics: Microelectromechanical Systems (MEMS) and Nonlinear Acoustics

Levent Degertekin, Chair

School of Mechanical Engineering, Georgia Institute of Technology, Atlanta, Georgia 30332

Contributed Papers

9:00

5aPA1. Plate equations and equivalent circuit for micromachined piezoelectrically actuated flextensional transducers. Gokhan Percin and Butrus T. Khuri-Yakub (Edward L. Ginzton Lab., Stanford Univ., Stanford, CA 94305-4085, percin@StanfordAlumni.Org)

In the absence of analytical expressions for the equivalent circuit parameters it is difficult to calculate the optimal design parameters and dimensions and to choose suitable materials. The influence of coupling between flexural and extensional deformation and coupling between structure and acoustic volume on the dynamic response of piezoelectrically actuated flextensional transducer is analyzed using three analytical methods: classical plate theory, Mindlin plate theory, and simple variational procedure. The transducer design is based on a flextensional transducer that excites the axisymmetric resonant modes of a clamped circular plate. It is constructed by depositing a thin piezoelectric annular plate onto a thin, edge clamped, circular plate. An ac voltage is applied across the piezoelectric material to set the compound membrane into flexural vibration. The devices have a range of operating resonance frequencies starting from 450 kHz up to 4.5 MHz. The device is manufactured by silicon surface micromachining and implemented in the form of two-dimensional arrays. Individual elements are made of thin silicon nitride membranes covered by a coating of piezoelectric zinc oxide. Classical thin plate theory, Mindlin plate theory, and variational methods are applied to derive two-dimensional plate equations for the transducer, and to calculate the coupled electromechanical field variables such as mechanical displacement. These methods use classical kinematic relations for a plate and a variational equation for the coupled electromechanical field to reduce three-dimensional field equations to two-dimensional plate equations. As a result, three different exact solutions to corresponding systems are obtained. An equivalent circuit of the transducer is also obtained from these solutions.

9:15

5aPA2. Microfluidic device for amplification of ultrasonic pressure waves with application to MEMS hydrophones. Daniel N. Pascual, Debora A. Compton, Tom G. Bifano, Robin O. Cleveland, and Allan D. Pierce (Mech. Eng. Dept., Boston Univ., 110 Cummings St., Boston, MA 02215, dpascual@bu.edu)

A micro-electro-mechanical systems (MEMS) device has been developed to amplify the fluid motion associated with ultrasound waves in water. Sound waves were sensed by the displacement of a small silicon membrane. Two structures were used in association with the membrane to increase the membrane motion: an acoustic horn and an acoustic resonator. An acoustic horn typically provides moderate signal gain across a broad bandwidth. An acoustic resonator provides significant gain in a small frequency band. Horns and resonators were constructed in bulk (100) silicon using KOH etching. Measurements of the membrane motion, in response to ultrasound in the 1- to 5-MHz range, were made with a laser Doppler velocimeter (LDV). Horn gains were measured to be on the order of 5 over the bandwidth used. Resonator gains of up to 3 were measured with a Q of 4. The LDV measurements also demonstrated that the membrane has a modal response which may be undesirable for the detection of acoustic waves. Both the horn and resonator structures have the potential to increase the sensitivity of most types of MEMS based acoustic sensors.

9:30

5aPA3. Characterization of single-crystal silicon paddle oscillators: Dynamic behavior and loss mechanisms. Scot F. Morse, Brian H. Houston, Douglas M. Photiadis (Naval Res. Lab., Code 7136, 4555 Overlook Ave. SW, Washington, DC 20375, morse@lpsa2.nrl.navy.mil), Joseph F. Vignola, and Xiao Liu (SFA, Inc., Largo, MD 20774)

Micro-scale mechanical devices continue to be of interest in a broad range of sensor and device applications. Many fundamental issues pertaining to the operation and performance of these devices remain to be thoroughly investigated. This is especially true in the recently developed field of nanoelectromechanical systems (NEMS). Resonant devices are now being produced in single-crystal silicon with released features as small as 30 nm. Besides novel device applications these structures may be used to explore fundamental properties and mechanisms that include internal friction as a function of scale. Measurements and analysis are reported here that, for the first time, characterize the detailed vibratory response of micro-scale oscillators ($\sim 100 \mu\text{m}$). The devices are placed in a vacuum chamber, excited inertially, and measured optically with a laser Doppler vibrometer (LDV) instrument. Single-point swept frequency measurements are used to investigate the vibrational spectrum of the oscillators and determine the Q. Spatial scans allow for mode identification and help to identify energy loss mechanisms into the supporting structure. In addition to observing the fundamental torsional and translational modes, we also observe numerous flexural modes of the devices. Both linear and nonlinear behavior are observed. [Work supported by ONR.]

9:45

5aPA4. Behavior characteristics of silicon micromachined electrostatic ultrasonic transducers. Li-Feng Ge and Jingping Shao (Dept. of Automation, Anhui Univ., Hefei 230039, PROC)

The plate-on-air-spring model for electrostatic ultrasonic transducers with grooved backplates [L.-F. Ge, IEEE Trans. UFFC **46**, 1120-1127 (1999); Sci. China, Ser. A **42**, 1308-1315 (1999)] is applied to simulate the behavior of silicon micromachined transducers with pit-array-textured backplates. Thus a subdiaphragm on an individual cavity is treated as a circular or square tensile plate supported by an elastic medium. A program is developed for simulating the physical model. The calculated results reveal that such transducers possess the same basic behavior characteristics as grooved transducers, i.e., multi-peaked or multimode responses. The fundamental mode is determined by the air-gap spring stiffness and the electrostatic negative stiffness produced by the electrostatic attraction, and higher modes (i.e., bending modes) mainly by the bending stiffness of the diaphragm and the tension in it. Therefore, the useful operation modes for such transducers have two single modes (i.e., the fundamental mode and the first bending mode), and a compound mode including the two single modes simultaneously. The analytical results well agree with the experimental observations. [Work supported by NNSF (#69974001), NCP, ANSF, and AU211F.]

10:00

5aPA5. Focusing of acoustical shock waves at a caustic cusp.

François Coulouvrat (Lab. Mod. Mécanique, UPMC & CNRS (UMR 7607), 8 rue du Cap. Scott, 75015 Paris, France, coulouvrat@ccr.jussieu.fr)

This paper investigates theoretically the focusing of weak acoustic shock waves at a caustic cusp. Near the cusp, a diffraction boundary layer is introduced; its characteristic length scales are determined based on the Pearcey function which governs the linear acoustic field. With these adequate variables, the nonlinear wave equation is then shown to reduce to the KZ equation. The proper matching to the geometrical acoustics approximation provides the complete boundary conditions for any transient signal, so that the problem is numerically tractable. A similitude is found for an incoming step wave, and is compared to focusing at a smooth caustic. The linear approximate solution is obtained in the general case. Application for an incoming N-wave to simulate sonic boom superfocusing is also discussed.

10:15

5aPA6. Focusing of shock waves at smooth caustics. Thierry Auger and François Coulouvrat (LMM, UPMC/CNRS, 8 rue du Cap. Scott, 75015 Paris, France, thierry.auger@airbus.aeromatra.com)

Focusing of weak shock waves at smooth caustics is investigated both theoretically and numerically. The pioneering works of Buchal and Keller (1960) and Ludwig (1966) about geometrical theory of diffraction are extended to take into account 3D heterogeneity and fluid motion, along with nonlinear effects. We thus recover the nonlinear Tricomi equation derived by Guiraud (1965) and Hayes (1968). In most situations, nonlinear effects are shown to be negligible for incoming smooth signals. However, for incoming weak shock waves, the linear theory predicts an unphysical, infinite amplification of the signal discontinuity. To recover finite amplitudes, it is necessary to take into account nonlinearities. The resulting equation to be solved is the so-called nonlinear Tricomi equation, similar to a transonic equation. To solve it, an iterative algorithm is based on an unsteady version of the equation. The algorithm is a modification of the pseudospectral code used for solving the KZ equation of nonlinear acoustics. Convergence of the method is analyzed. The results are illustrated by simulations performed in various situations, and the application to sonic boom is discussed.

10:30–10:45 Break

10:45

5aPA7. Some numerical result properties of nonlinear and dispersive equations. Hichem Hadouaj (Académie Navale, 7050 Menzel Bourguiba, Tunisia, North Africa)

It is known that the envelope-solitary shear horizontally surface waves at the top of an elastic film superimposed on a nonlinear elastic substrate are governed by a nonlinear Schrödinger equation at the interface. And also, the cubic Schrödinger equation using the complex amplitude is the model equation which describes the envelope optical solitons propagation in fibers. The generalized Zakharov system has found a number of applications in various physical problems, such as interaction of intramolecular vibrations giving rise to Davydov solitons with acoustic disturbances, interaction of high-frequency and low-frequency gravity disturbances in an atmosphere, and so on. Some numerical simulations of soliton propagation and soliton-soliton collisions in a generalized Zakharov system are presented. It envisages soliton-soliton interactions in the damped ac-driven nonlinear Schrödinger equation, which is the simplest model of charge-density wave conductor or of a cold plasma driven by an external ac

electric field. Numerical simulations on collision of two solitons with zero initial velocity at infinity are given, and it is confirmed that after the collision the soliton separates at a finite velocity, i.e., a two-soliton breather does not exist in this model.

11:00

5aPA8. A simple method for comparing waveform distortion of nonlinear surface waves in different cubic crystals. R. E. Kumon and M. F. Hamilton (Dept. of Mech. Eng., Univ. of Texas, Austin, TX 78712-1063)

Simulations of monofrequency, nonlinear surface acoustic waves propagating in cubic crystals were performed using the theory of Hamilton, Il'inskii, and Zabolotskaya [J. Acoust. Soc. Am. **105**, 639–651 (1999)]. Nonlinearity matrix elements, which describe the coupling strength of harmonic interactions, provide a useful tool for characterizing the nature of the waveform distortion and can be computed without integrating large systems of equations. In cases where the surface is a plane of mirror symmetry, the nonlinearity matrix elements are real valued, and the longitudinal components of the waveforms typically form either compression or rarefaction shocks. However, in cases where the surface has lower symmetry, the nonlinearity matrix elements are complex valued, and the resulting waveform distortion is typically asymmetric, sometimes with low-frequency oscillations forming near peaks and shocks. A simple transformation based on the phase of the dominant nonlinearity matrix element was shown to provide a reasonable approximation of the asymmetric distortion in many cases. Comparisons will be presented between approximate and exact solutions in the (111) plane for several real crystals. These results are consistent with measured surface wave pulses reported previously [R. E. Kumon *et al.*, J. Acoust. Soc. Am. **103**, 2926(A) (1998)]. [Work supported by ONR.]

11:15

5aPA9. Shock formation in the presence of entropy gradients. Hao Lin and Andrew J. Szeri (Dept. of Mech. Eng., Univ. of California, Berkeley, CA 94720-1740)

The steepening of a normal compression wave into a shock in a homentropic flow field is well understood through the method of characteristics. In a non-homentropic flow field, however, shock formation from a compression wave is more complex. The effects of entropy (or sound speed) gradients on shock formation from a compression wave are determined using a wave front expansion in Cartesian and in spherical polar coordinates. The latter problem has application to the intense energy focusing of sonoluminescence, particularly when applied to a spherically collapsing gas. The principal result is an analytical criterion for the time and place of shock formation, for a wave propagating into a field of smoothly varying entropy.

11:30

5aPA10. Nonlinear surface waves in isotropic elastic substrates with thin films. Won-Suk Ohm and Mark F. Hamilton (Dept. of Mech. Eng., Univ. of Texas, Austin, TX 78712-1063)

Surface waves in isotropic elastic solids are nondispersive. Considered here is the effect of dispersion, due to a thin solid film on the surface of an isotropic substrate, on nonlinear propagation. The analysis is performed with the theoretical model developed by Zabolotskaya [J. Acoust. Soc.

Am. **91**, 2569 (1992)] augmented to include dispersion. The dispersion relation is obtained from linear theory but is otherwise exact. Approximations leading to a relation of the Korteweg–de Vries or Benjamin–Ono type, sometimes proposed for the study of dispersive nonlinear surface waves, are not employed. Both loading and stiffening films are considered. Loading films decrease the surface wave speed. Stiffening films increase the surface wave speed until it equals the shear wave speed in the substrate, where cutoff of the surface wave mode occurs. Computations were performed with elastic moduli for a fused quartz substrate, which exhibits negative nonlinearity in the sense that rarefaction rather than compression shocks form in the horizontal particle velocity waveform. Gold (loading) and alumina (stiffening) films are considered. Weak, moderate, and strong dispersion (corresponding to increasing film thickness) relative to nonlinearity are examined. Stationary wave solutions are obtained for the alumina film. [Work supported by ONR.]

11:45

5aPA11. Propagation of short pulse through arbitrary relaxing media with resonant properties. New approach. German A. Maksimov and Vladimir A. Larichev (Dept. of High Density Energy Phys., Moscow Eng. Phys. Inst., Kashirskoe sh.31, Moscow 115409, Russia, maximov@dpt39.mephi.ru)

The new analytical representation of the fundamental solution (Green's function) describing the short pulse propagation in a medium with a single process of resonant relaxation is presented in the report. This analytical solution is based on the generalized local response function of linear media [V. A. Larichev and G. A. Maksimov, *Acoust. Phys.* **44**, No. 46, 709–716 (1998)]. It contains well-known Lorentz and Zener models of relaxing media, like particular cases. The changing of the pulse shape at propagation, described by the obtained solution, shows a variety of forms of pulse propagation and general laws of pulse dynamics beginning from pure relaxation behavior and up to the resonant one. [Work supported by ISTC in Moscow.]

SATURDAY MORNING, 3 JUNE 2000

HENRY/GRADY ROOM, 8:30 A.M. TO 12:00 NOON

Session 5aPP

Psychological and Physiological Acoustics: Potpourri (Poster Session)

Marjorie R. Leek, Chair

Army Audiology and Speech Center, Walter Reed Army Medical Center, Washington, D.C. 20307-5001

Contributed Papers

All posters will be on display from 8:30 a.m. to 12:00 noon. To allow contributors an opportunity to see other posters, contributors of odd-numbered papers will be at their posters from 8:30 a.m. to 10:15 a.m. and contributors of even-numbered papers will be at their posters from 10:15 a.m. to 12:00 noon.

5aPP1. Consonant recognition in noise with temporal cues: I. Effects of temporal envelope enhancement on identification performance.

Frederic Apoux (LPE, UMR CNRS 8581, Inst. de Psych., Univ. Paris V, 71 Av. Edouard Vaillant, 92774 Boulogne-Billancourt Cedex, France and Intrason France), Christian Lorenzi (Univ. Paris V, 92774 Boulogne-Billancourt Cedex, France), and Frederic Berthommier (ICP/INPG, 38031 Grenoble Cedex, France)

This study sought to determine whether consonant identification performance in noise can be improved by applying an expansive nonlinearity to the temporal envelope of the speech stimuli. Consonants in /a-consonant-a/ context were processed through four broad bandpass filters, half-wave rectified and lowpass filtered at 500 Hz to extract the temporal envelope from each band. The envelope from each band was used to modulate a band-limited noise. In each band, the temporal envelope was either left intact or expanded by raising it to the power 1.414 or 2. Three signal-to-noise ratios (S/N) were used: -6, 0, and +6 dB. Twelve young normal-hearing listeners were tested twice in each condition, one week apart, using a forced-choice paradigm. Overall, identification performance increased with S/N and training. For each S/N and testing session, small but consistent improvements in identification performance of 2%–6% were observed when raising the envelopes to the power 1.414 or 2. Information transmission analysis performed on the aggregate confusion matrix showed that improved identification with expanded envelopes was mainly due to an increase in the reception of voicing.

5aPP2. Consonant recognition in noise with temporal cues: II. Effects of temporal envelope enhancement on response times.

Frederic Apoux (LPE, UMR CNRS 8581, Inst. de Psych., Univ. Paris V, 71 Av. Edouard Vaillant, 92774 Boulogne-Billancourt Cedex, France and Intrason France), Olivier Crouzet, and Christian Lorenzi (Univ. Paris V, 92774 Boulogne-Billancourt Cedex, France)

In an effort to evaluate the importance of envelope cues in speech recognition, response times and percentages of correct responses were measured simultaneously in normal-hearing and hearing-impaired listeners, using a forced-choice identification paradigm. Naturally produced /aCa/ disyllables presented in a white noise background (+6 dB signal-to-noise ratio) were used to create temporal-only stimuli having instantaneous amplitudes identical to the natural speech stimuli, but flat spectra. The temporal envelope of the /aCa/ disyllables was either left intact or enhanced by raising it to the power 2. In the first experiment, 13 young normal-hearing listeners (24–27 years) were tested. The results show that envelope enhancement yields a general improvement in both percent correct identification (4.75%) and response times (63 ms). In the second experiment, nine hearing-impaired listeners (53–71 years) were tested at a comfortable listening level. Overall, temporal expansion fails to show better consonant identification in terms of percentage of correct responses, but yields a significant improvement (66 ms) in terms of response time. These results suggest that: (1) the perceptual effects of speech-envelope enhancement cannot be completely accounted for by percent correct-responses measurements, and (2) response times seem to be more sensitive to the effects of envelope enhancement.

5aPP3. New techniques to reduce observer inconsistency in psychoacoustic experiments. Judi A. Lapsley Miller (Psychophysics Lab., Victoria Univ. of Wellington, New Zealand, jmiller@nsnml.navy.mil)

One of the biggest problems still facing experimental psychophysics is observer inconsistency. Observer inconsistency degrades performance, hampering comparisons with other observers and theoretical predictions. By repeating an experiment multiple times, using the same stimuli, observer inconsistency can be reduced and asymptotic errorless performance estimated by using group operating characteristic (GOC) analysis and function of replications combined estimation (FORCE) analysis, respectively. These new techniques are described and illustrated with results from various psychoacoustic experiments, showing that humans are sometimes capable of performing as well as an ideal observer once observer inconsistency is reduced.

5aPP4. Detection advantage with coherent amplitude modulation for speech filtered into narrow bands. Joseph W. Hall III, Emily Buss, and John H. Grose (Div. Otolaryngol./Head and Neck Surgery, Univ. of North Carolina School of Medicine, Chapel Hill, NC 27599, jwh@med.unc.edu)

Introduction of masker amplitude modulation can improve detection thresholds for a masked signal in a number of masking paradigms. In some cases this detection advantage depends on the coherence of modulation across a relatively wide range of frequencies. In the current experiment 43 spondees were recorded from a single male talker and employed as the signal in a forced choice, closed set word identification task. Both the spondee signal and the speech-shaped Gaussian noise masker were filtered into six narrow bands. Either the signal or the masker was amplitude modulated by an 8-Hz raised cosine, and that modulation was either coherent or random across bands. Thresholds were consistently lower when the masker was modulated with coherent rather than random patterns of AM. No difference between coherent and random AM was observed when the speech signal was modulated. This finding suggests that the advantage observed with coherent masker AM may not have been driven simply by the coincidence of speech energy across frequency appearing in masker "dips." Other explanations will be discussed, including a possible interpretation in terms of comodulation masking release.

5aPP5. Critical speech bands: Spectral resolution of contrasting temporal speech patterns. Eric W. Healy and Sid P. Bacon (Psychoacoust. Lab., Dept. of Speech and Hearing Sci., Arizona State Univ., P.O. Box 871908, Tempe, AZ 85287-1908, ewh@asu.edu)

The high intelligibility of narrow-band sentences is attributable to contrasting temporal patterns in different frequency regions within the narrow band [E. W. Healy, Dissertation, University of Wisconsin-Milwaukee (1998)]. In the current study, the spectral resolution for contrasting temporal speech patterns was measured by partitioning the narrow-band spectrum. Sentences from the HINT were filtered to a minimum overall bandwidth, centered at 1500 Hz, which allowed relatively high intelligibility of component words. The overall bandwidth was 2/3, 1, or 3/2 octaves. Each overall band was digitally filtered (several thousand dB/octave) into 1, 2, 4, 6, or 10 partitions. The envelope from each partition was used to amplitude-modulate a corresponding band of low-noise noise, and the modulated carriers were combined and presented to normal-hearing listeners. Intelligibility increased and reached an asymptote as the number of partitions increased. The number of partitions necessary for asymptotic performance was 4, 6, and 10 for the 2/3-, 1-, and 3/2-octave overall bandwidths. For each, this corresponds to a partition bandwidth of approximately 1/6 octave. These results indicate that, in the region surrounding 1500 Hz, resolution of individual temporal speech patterns approximates the critical band measured using traditional psychoacoustic methods and nonspeech stimuli. [Work supported by NIDCD.]

5aPP6. Effects of signal frequency and masker level on the Schroeder phase effect. Rene H. Gifford and Sid P. Bacon (Psychoacoustics Lab., Dept. of Speech and Hearing Sci., Arizona State Univ., Tempe, AZ 85287-1908, Rene.Gifford@asu.edu)

Positive-Schroeder ($m+$) and negative-Schroeder ($m-$) harmonic complexes have identical amplitude spectra but different phase spectra. The corresponding waveforms are mirror images of one another. The $m+$ complexes usually produce less masking, revealing a so-called Schroeder phase effect (SPE). Many have hypothesized that the SPE results from a combination of the $m+$ complexes producing a more peaked response along the basilar membrane along with the influence of fast-acting compression. The present study examined the effects of signal frequency and masker level on the SPE in a simultaneous-masking paradigm. The signal frequency was 600, 1200, 2400, or 4800 Hz. The masker covered a spectral range from about one octave below to one octave above the signal frequency. Fundamental frequencies were chosen individually based on comparing data for $m+$ and sine-phase maskers. The overall level of the masker ranged from 50 to 95 dB SPL. The SPE generally increased with increasing signal frequency. For all signal frequencies, the SPE increased as the masker level increased from its lowest level. At the higher frequencies, however, the SPE decreased significantly with level. The results are consistent with a greater degree of compression at high frequencies and moderate levels. [Work supported by NIDCD.]

5aPP7. The effects of aspirin on a psychophysical estimate of basilar membrane compression. Michelle L. Hicks and Sid P. Bacon (Psychoacoustics Lab., Dept. of Speech and Hearing Sci., Arizona State Univ., Tempe, AZ 85287-1908, mhicks@asu.edu)

It has been suggested that the amount of basilar membrane (BM) compression to tones at characteristic frequency can be estimated psychophysically by examining the slope of a growth-of-masking function for a masker that is lower in frequency than the signal [A. J. Oxenham and C. J. Plack, *J. Acoust. Soc. Am.* **101**, 3666–3675]. In this study, growth-of-forward masking functions were measured before and during administration of a moderate dose of aspirin that caused 5–15 dB of temporary hearing loss. The level of a 200-ms sinusoidal masker required to mask a 5-ms, 4-kHz sinusoidal signal presented 2 ms after the offset of the masker was determined for signal levels from 30 to 90 dB SPL. The masker was either the same frequency, or lower in frequency than the signal (signal-masker ratios of 1.2 to 1.8). High-pass noise was presented continuously to prevent off-frequency listening. In general, the aspirin, by presumably altering outer hair cell function, resulted in growth-of-masking functions (masker level versus signal level) for the off-frequency maskers that were steeper by as much as a factor of 2. These results support the argument that the slopes of the functions reflect BM compression, and that aspirin reduces that compression. [Work supported by NIDCD.]

5aPP8. The influence of aspirin on temporal effects in simultaneous masking with noise and tonal maskers. Sid P. Bacon and Michelle L. Hicks (Psychoacoustics Lab., Dept. of Speech and Hearing Sci., Arizona State Univ., Tempe, AZ 85287-1908, spb@asu.edu)

Temporal effects in simultaneous masking were measured before and during the administration of a moderate dose of aspirin that caused about 5–15 dB of temporary hearing loss. The 10-ms, 4.0-kHz signal was presented at the beginning or in the temporal center of a 400-ms masker. The masker was either a broadband noise presented at a spectrum level of 10 or 20 dB SPL or a tone with a frequency of 5.2, 5.4, or 5.6 kHz presented at a level of 80 dB SPL. Consistent with previous results [D. McFadden and C. A. Champlin, *J. Acoust. Soc. Am.* **87**, 2634–2642 (1990)], aspirin tended to reduce the temporal effect with the broadband masker (i.e., the overshoot) by decreasing the threshold for the signal at masker onset. Aspirin also tended to reduce the temporal effect in the presence of the tonal masker, but in this case by increasing the threshold for the signal in the temporal center and, to a lesser extent, at masker onset. This differen-

tial effect of aspirin suggests that the mechanisms underlying the two temporal effects may differ. [Work supported by NIDCD.]

5aPP9. Forward masking in normal-hearing, hearing-impaired, and noise-masked normal listeners. Paula P. Henry and D. Wesley Grantham (Vanderbilt Bill Wilkerson Ctr., 1114 19th Ave. S., Nashville, TN 37212, paula.henry@vanderbilt.edu)

This study examined forward-masking functions in normal-hearing, hearing-impaired, and noise-masked normal listeners using a 300-ms, one-octave 80 dB SPL masker centered at 4000 Hz and a 20-ms, 4000-Hz pure-tone signal. Thresholds were measured as a function of interstimulus interval (0 to 150 ms), and a time constant (τ) was estimated from the best exponential fit to this function. No significant differences were found between τ values obtained from the normal-hearing (mean: 44 ms) and hearing-impaired (mean: 45 ms) subjects. Thus, temporal resolution in this task was not worse for the impaired listeners than for the normal listeners. Under the noise-masked condition, normal-hearing subjects completed the forward-masking task in the presence of a continuous 4000-Hz, one-octave noise, whose level was chosen to shift each subject's threshold to match that of a hearing-impaired subject of equivalent age and gender. The mean τ value obtained in the noise-masked condition was 8 ms, which was significantly smaller than the τ values obtained from either group in the absence of a background masker. This curious finding suggests that use of a continuous noise masker to simulate hearing loss is not effective in this temporal resolution task. [Work supported by NIDCD.]

5aPP10. Grouping and adaptation in modulation detection interference. Andrew J. Oxenham (M.I.T., 77 Massachusetts Ave., Cambridge, MA 02139) and Torsten Dau (Boston Univ., Boston, MA 02215)

Modulation detection interference (MDI) can be reduced by gating the masker and signal carriers asynchronously, possibly due to grouping effects. However, as both carriers are generally modulated throughout their presentation, two other factors may also play a role: First, the response to modulation may adapt, thereby reducing the effective masker modulation depth prior to signal presentation. Second, analogous to intensity perception, a dynamic change in modulation depth (increment detection) may be easier to detect than an overall difference across two presentations (level discrimination). A MDI paradigm with 1- and 4.3-kHz carriers was used. The 16-Hz sinusoidal signal modulation was masked by noise modulation with a passband from 8 to 24 Hz. In one experiment, the gating synchrony of carrier and modulator were independently varied. In another experiment, both masker and signal modulated a single 4.3-kHz carrier, with gating synchrony between masker and signal varied. Results suggest that the asynchronous gating of either the carriers or modulators alone is not sufficient to produce a release from MDI. Thresholds were higher for the single carrier than for the two carriers, and gating the masker modulation on prior to the signal did not improve performance. [Supported by NIDCD and Max Kade Foundation.]

5aPP11. Modulation detection interference with narrow-band noise interferers. Jesko Verhey (Carl von Ossietzky Universitaet Oldenburg, AG Medizinische Physik, D-26111 Oldenburg, Germany, jesko@medi.physik.uni-oldenburg.de) and Torsten Dau (Boston Univ., Boston, MA 02215)

The detection of amplitude modulation of a sinusoidal (target) carrier can be impaired by the presence of a modulated sound (interferer) with a different carrier frequency, referred to as modulation detection interference (MDI). Typically, sinusoidally amplitude-modulated tones remote in frequency are used as the interferer, and the target modulation threshold is increased if the interferer- and target-modulation rates are similar. In the present study, narrow-band noise was used as the interferer. Detection thresholds for modulations of 5, 10, and 50 Hz imposed on tone carriers

were measured as a function of noise bandwidth (0–200 Hz), the spectral distance between the carrier and interferer (0.5–4 octaves), the carrier level (40–80 dB SPL), and the carrier-interferer level difference (0–30 dB). In all conditions, the intrinsic fluctuations of the noise interfered with modulation detection, with maximum interference at noise bandwidths about twice the modulation frequency. This bandwidth dependence corresponds directly to that found in “on-frequency” conditions with narrow-band noise carriers; however, the interference is much smaller. An across-channel modulation-filterbank model is used to describe the similar bandwidth dependence and the different overall effectiveness of the on- and off-frequency interferer. [Work supported by the Deutsche Forschungsgemeinschaft.]

5aPP12. Temporal resolution of frequency-modulated and virtual frequency signals. Nandini Iyer, Lawrence L. Feth (Dept. of Speech and Hearing Sci., The Ohio State Univ., 110 Pressey Hall, 1070 Carmack Rd., Columbus, OH 43210, iyer.8@osu.edu), and Ramkumar Sridharan (The Ohio State Univ., Columbus, OH 43210)

Discrimination of “glide” versus “step” transition was investigated for two classes of signals: frequency modulated (FM) tones and virtual frequency signals [Lublinskaya (1996); Anantharaman *et al.* (1997)]. A “virtual frequency signal” refers to the frequency transition perceived in an amplitude-modulated two-tone complex when the amplitude of one of the tones increases with duration, while that of the other decreases. In the FM tone experiment, the glide signal traverses a certain frequency range in a linear trajectory, whereas the step signal traverses the same frequency range in a series of discrete steps. In the virtual frequency signal experiment, the glide signal has a linear amplitude transition whereas the step signal has a staircase-like amplitude transition. Listeners heard these signals in an adaptive 2Q, 2AFC paradigm. The experimental parameters for both classes of signals were center frequency (500, 1000, and 4000 Hz) and number of steps. The signal duration was retained constant at 250 ms. The smallest ERB separation for which the step signal could be distinguished from the glide signal was determined. Results from these experiments will be discussed along with implications for temporal resolution and spectral integration of virtual frequency signals. [Work supported by a grant from The Ohio State University College of Social and Behavioral Sciences.]

5aPP13. Filter shapes for long-duration signals and for short-duration signals as a function of signal delay from masker onset. Elizabeth A. Strickland (AUS Dept., Purdue Univ., West Lafayette, IN 47907)

Filter shapes were measured as a function of level for long-duration (350-ms) signals, and for short-duration (10-ms) signals delayed from masker onset by 2 or 202 ms. Masker duration was 400 ms for the long-duration and long-delay signals, and 200 ms for the short-delay signal. Signal frequencies were 1 and 4 kHz. The level of a notched-noise masker was varied to find threshold. This study extends previous studies by fixing the signal level rather than the masker level, and by measuring filter shapes for several signal levels. Filter widths for the long-duration signal and the long-delay signal could be fit for the multiple levels, keeping efficiency constant. Filter widths were similar for these two conditions. This agrees with previous results measured with the masker level fixed [B. A. Wright and H. Dai, *J. Acoust. Soc. Am.* **95**, 931–938 (1994)]. Filter widths for the short-delay signal were broader, and efficiency varied with level. The broader filter at a short delay as compared to a long delay is also consistent with results measured with masker level fixed [B. A. Wright, *J. Acoust. Soc. Am. Suppl. 1* **86**, S1, S121]. Results will be discussed in terms of overshoot. [Work supported by NIH (NIDCD).]

5aPP14. Coincidence network for estimating tone frequency from firing-rate wavelength. Rachod Thongprasirt and Timothy A. Wilson (Dept. of Elec. Eng., Rm. EN204C, The Univ. of Memphis, Memphis, TN 38152, tawilson@memphis.edu)

A signal-processing auditory model was constructed. It was composed of four stages: the cochlear-filtering stage, the transduction stage, the neural-adaptation stage, and the spike-generation stage. A linear fourth-order gammatone filter with traveling-wave delay was used in the cochlear-filtering stage. The output from the neural adaptation was the rate of spike generation. The cochlear wavelength was defined as the distance between successive peaks of the cochlear traveling wave. It was shown that for a given tone input there is an approximately linear relationship between the cochlear wavelength and the peak location of the cochlear traveling wave. Furthermore, this linear relationship still holds, within the low-frequency range, between the firing-rate wavelength and its peak location. A simple coincidence-counting network was constructed to estimate the input frequency in the low-frequency range. Each coincidence element had inputs from two different places along the cochlear length. Its output was then processed by another neural level. The results of the frequency estimation will be presented.

5aPP15. Characteristics of efferent mediated enhancement of distortion-product otoacoustic emissions. Xiao-Ming Sun and David F. Dolan (Kresge Hearing Res. Inst., Univ. of Michigan Med. School, 1301 East Ann St., Ann Arbor, MI 48109)

Efferent stimulation can enhance distortion-product otoacoustic emissions (DPOAEs) (Siegel and Kim, 1982), but the stimulus conditions in which enhancement occurs have not been clearly described. To clarify the efferent mediated enhancement of cochlear active process, $2f_1-f_2$ DPOAE I/O function was measured in pigmented guinea pigs. For $f_2/f_1 = 1.20$, the crossed olivocochlear bundle (COCB) stimulation suppressed the DPOAE evoked by low primary levels, the rapid growth portion of an I/O function. It enhanced it in the saturating portion evoked by moderate levels, 45 to 65 dB SPL when $L_1-L_2=0$ to 20 dB, and suppressed it with further increase in stimulus level. For $f_2/f_1 = 1.08$, enhancement appeared only when $L_1=75$ dB SPL and $L_1-L_2=30$ to 40 dB. For $f_2/f_1 = 1.45$, enhancement appeared when $L_1=45$ to 55 dB SPL and $L_1-L_2=-10$ to -20 dB. COCB stimulation tends to alter the DPOAE I/O function toward linear growth. Both suppression and enhancement were significantly reduced by strychnine. The results suggest that effect of efferent activation on evoked OHC activity is not a simple suppression, but a modulation of the nonlinearity of its growth function. [Work supported by grants: NIH/NIDCD P01-DC00078 and T32-DC00011.]

5aPP16. Acoustic effects of middle-ear cavities in the cat family. Gregory T. Huang, William T. Peake (Res. Lab of Electron., MIT, 77 Massachusetts Ave., Cambridge, MA 02139), Annette M. Taberner (Harvard—MIT Div. of Health Sci. & Technol., Cambridge, MA 02139), John J. Rosowski, and Craig A. Atencio (Mass. Eye & Ear Infirmary, Boston, MA 02114)

The acoustic effect of the middle-ear air-filled space is to add an impedance in series with that of the tympanic membrane and ossicular chain. In the cat family, a bony septum divides the space into two volumes, coupled through a small opening (foramen); this configuration is well modeled as a lumped circuit of two parallel compliances (the cavities) coupled by an acoustic mass and resistance (the foramen). The model impedance magnitude has a minimum and a maximum, whose frequencies are predictable from structure. A narrow-band notch in hearing sensitivity near 2–4 kHz is caused by the pole. Measurements of cavity impedance in four felid species are consistent with this structure-based model. Acoustic measurements in ears of three species with the septum removed show that the minimum and maximum in impedance magnitude are eliminated, as predicted by the lumped model. However, new impedance maxima appear at 8–12 kHz; these maxima are assumed to result from distributed resonances in the enlarged cavity. Thus the inferred effect of the septum on

hearing is to introduce a notch in the midfrequencies, while avoiding such notches at higher frequencies where they could interfere with sound-source localization. [Work supported by NIH and NSF.]

5aPP17. Specificity of learning for the discrimination of sinusoidal-amplitude-modulation rate. Matthew B. Fitzgerald and Beverly A. Wright (Audiol. and Hearing Sci. Prog., 2299 N. Campus Dr., Northwestern Univ., Evanston, IL 60208-3550)

There is evidence that the learning of tonal-frequency discrimination is specific to the trained frequency. This study examined whether learning of the discrimination of sinusoidal-amplitude-modulation (SAM) rate was also specific to the trained rate, and whether that learning was specific to the trained task. SAM-rate-discrimination thresholds of six listeners trained 1 h per day for 6 days with a 150-Hz rate improved from 24 Hz before to 12 Hz after training. Comparisons between proportional improvements of trained and control listeners revealed that trained listeners learned substantially more than controls on the trained task, but learned only slightly or no more than controls on: (1) the SAM-rate task at two untrained rates (30 and 300 Hz) and (2) three untrained tasks (tonal-frequency discrimination, ripple-noise discrimination, modulation detection) at the trained rate/frequency. Thus, as with tonal-frequency discrimination, learning of SAM-rate discrimination is specific to the trained rate. However, the mechanism(s) underlying this discrimination are probably separate from those governing tonal-frequency discrimination and other related discrimination or detection tasks. These results enhance our understanding of perceptual learning, the processing of amplitude modulation, and the perception of pitch. [Work supported by NIDCD.]

5aPP18. The emergence of cross-modal complex temporal-processing abilities in children. Kelley M. Wright, Janet M. Weisenberger, and Jennifer D. McDonnell (Speech and Hearing Sci., Ohio State Univ., Columbus, OH 43210)

Studies of temporal processing of multi-item visual, auditory, and tactile stimulus sequences in children with central auditory-processing disorders have shown considerable variability even among the children in the normal control group, suggesting that some aspects of complex temporal processing may continue to develop throughout childhood. In the present study, two experiments were performed. In the first, groups of normally developing children aged 8, 10, 12, and 14 years were tested with three-item sequences of auditory tones (analogous to the pitch pattern sequence test widely used in assessing central auditory-processing abilities), sequences of colored boxes, and sequences of tactile vibrations presented to different locations on the forearm. Children were asked to identify the sequence presented. Results showed significant improvements in performance with age, with most of the difference between the 8- and 10-year-old groups. A second experiment examined younger children, aged 6, 7, and 8 years, with slightly modified versions of the same tasks. Again, a significant improvement in performance with age was observed. These improvements were evident for all sensory modalities, although overall performance levels were highest for the visual task and lowest for the tactile task. Results are discussed in terms of more general development of perceptual and cognitive abilities in childhood. [Work supported by NIH.]

5aPP19. Temporal envelope processing in developmental dyslexics. Christian Lorenzi, Christian Fullgrabe (LPE, Univ. Paris V, 71 Av. Edouard Vaillant, 92774 Boulogne-Billancourt Cedex, France), and Annie Dumont (Hopital Robert Debre, 75019 Paris, France)

This study evaluates temporal-envelope processing in six dyslexic children (mean age: 10.10 years), six normal control children (11.6 years), and six normal control adults (24.8 years) by measuring: (1) temporal modulation transfer functions (TMTFs), i.e., the detection thresholds of sinusoidal amplitude-modulation applied to a white noise carrier, as a function of modulation frequency (4, 16, 64, 256, or 1024 Hz), (2) recog-

dition performance for vowel-consonant-vowel stimuli over five sessions; speech stimuli were either unprocessed or processed to degrade the spectral information. Modulation thresholds were similar in normal children and adults; for both groups, TMTFs were low pass in shape. TMTFs measured in dyslexics were bandpass in four children, low pass and flat in the two others. Overall, modulation thresholds were higher in dyslexics than in normal children at 4 and 1024 Hz. Unprocessed-speech recognition performance was perfect in normal children and adults, and impaired in dyslexics. Processed-speech recognition performance was poorer in normal and dyslexic children compared to normal adults. Performance improved across sessions in normal children and adults, but remained constant in dyslexics. These results support the hypothesis that some dyslexics show abnormal auditory temporal-envelope processing. This may, in turn, explain their difficulties with speech perception. [Work supported by the Cognitique program.]

5aPP20. Amplitude-modulation detection in two patients with left- or right-hemisphere damage. Christian Lorenzi (LPE, UMR CNRS 8581, Inst. de Psych., Univ. Paris V, 71 Av. Edouard Vaillant, 92774 Boulogne-Billancourt Cedex, France), Jocelyne Wable, Christine Moroni, Catherine Belin, and Bruno Frachet (Hopital Avicenne, 93009 Bobigny Cedex, France)

The modulation depth required for the detection of sinusoidal amplitude-modulation applied to a white noise carrier was measured as a function of modulation frequency, giving temporal modulation transfer functions (TMTFs). The modulation frequency was 4, 16, 64, 256, or 1024 Hz. Measurements of intensity discrimination thresholds for white noise were performed as a control task. In both tasks, white noise stimuli had a 2-s duration and were presented monaurally at 75 dB SPL. Two adult listeners with left- or right-hemisphere damage and three normal-control

adult listeners were tested in each ear. Lesions encompassed the primary auditory cortex in the two patients. All listeners had normal audiometric thresholds. The modulation detection data showed no sign of ear asymmetry. For all listeners, the TMTFs displayed a typical low-pass characteristic. Modulation detection was 10–14 dB poorer-than-normal at modulation frequencies between 4 and 256 Hz in the patient with left-hemisphere damage. In contrast, modulation detection was normal in the patient with right-hemisphere damage. Intensity discrimination was within the normal range (0.5–1 dB) for both patients. These results support the hypothesis that cortical structures within the left hemisphere play a crucial role in temporal envelope processing. [Work supported by the Fondation de l'Avenir and the Cognitique Program.]

5aPP21. A kinship with unusual auditory perception. D. B. Daly and D. M. Daly (Box 210855, Dallas, TX 75211, openmike@alumni.stanford.org)

Craig, while investigating vibrotactile sensation for blind readers, encountered several individuals with extraordinary perceptual ability [Science **196**, 450 (1977)]. We report a two-generation kinship (three polyglots) with extraordinary ability to classify certain sets of synthetic acoustic stimuli. With sets of three or four formant /ge/-/ye/ which controls typically identify as containing two or three sounds (/e/, /ge/, /ye/), the mother and son could distinguish as many as six classes of sounds; the father could distinguish five. The additional classes involved stimuli with brief transitions; one class encompassed control transition durations. Each individual could also classify these same stimulus sets using two or three classes; these results were homogeneous ($p < 0.0001$) with the results from a population of 45 age/gender-matched controls. These results require fine-grained differences in rate or duration of basilar membrane movements be represented in auditory cortex.

SATURDAY MORNING, 3 JUNE 2000

CANADIAN ROOM, 8:15 TO 11:00 A.M.

Session 5aSC

Speech Communication: Phonetic Perception and Variability in Speech

James R. Sawusch, Chair

Department of Psychology, State University of New York at Buffalo, Park Hall, Buffalo, New York 14260

Contributed Papers

8:15

5aSC1. On the perception of voicing for plosives in noise. Marcia Chen and Abeer Alwan (Dept. of Elec. Eng., UCLA, 405 Hilgard Ave., Los Angeles, CA 90095)

Previous research has shown that the VOT and first formant transition are primary perceptual cues for the voicing distinction for syllable-initial plosives (SIP) in quiet environments. This study seeks to determine which cues are important for the perception of voicing for SIP in the presence of noise. Stimuli for the perceptual experiments consisted of naturally spoken CV syllables (six plosives in three vowel contexts) in varying levels of additive white Gaussian noise. In each experiment, plosives which share the same place of articulation (e.g., /p, b/) were presented to subjects in identification tasks. For each voiced/voiceless pair, a threshold SNR value was calculated. It was found that the perception of voicing has a strong dependence on vowel context with /Ca/ syllables being significantly better discriminated than /Ci/ and /Cu/ syllables. In addition, labials consistently had a higher threshold SNR (or more easily confusable) than alveolars and velars. Threshold SNR values were then correlated with measurements of several acoustic parameters of the speech tokens. It was found that VOT

did not appear to influence the perception of voicing in noise as much as the first formant transition. [Work supported in part by NIH and an NSF Fellowship.]

8:30

5aSC2. The role of duration in the perception of vowel quality. James M. Hillenbrand, Michael J. Clark (Speech Pathol. and Audiol., Western Michigan Univ., Kalamazoo, MI 49008), and Robert A. Houde (RIT Res. Corp., Rochester, NY 14623)

Test signals consisted of formant synthesized versions of 300 utterances selected from a database of 1668 /hVd/ syllables spoken by 45 men, 48 women, and 46 children [Hillenbrand *et al.*, J. Acoust. Soc. Am. **97**, 3099–3111 (1995)]. Four versions of each utterance were synthesized: (1) an *original duration* set in which the duration of each vowel was matched as closely as possible to that of the original utterance, (2) a *neutralized duration* set in which the synthesis control parameters were linearly stretched or contracted to produce a fixed vowel duration of 272 ms (the mean of all 300 utterances), (3) a *short duration* set in which the contours were resampled to produce a fixed duration of 144 ms (two standard deviations below the mean), and (4) a *long duration* set in which the

contours were resampled to produce a fixed duration of 400 ms (two standard deviations above the mean). The great majority of the signals were accurately identified under all four conditions, suggesting that the overall effect of duration on vowel identity is modest. Significant duration effects were observed, however, especially in the short-duration condition in which the majority of /æ/ tokens were heard as /ε/ and many tokens of /a/ and /ɔ/ were heard as /ʌ/. [Work supported by NIH.]

8:45

5aSC3. Thresholds for second formant transitions in front vowels. Shawn Goodman and Diane Kewley-Port (Dept. of Speech and Hearing Sci., Indiana Univ., Bloomington, IN 47405, shgoodma@indiana.edu)

Formant dynamics in vowel nuclei contribute to accurate vowel classification in English. This study examined listeners' ability to discriminate dynamic second formant transitions in female vowels /i, i, e, ε, æ/. Acoustic measurements were made from the nuclei (20% to 80% of vowel duration) of vowels in /bVd/ context in short sentences. Based on measurements along with informal listening, natural sounding vowel nuclei were synthesized. Synthesis parameters were selected to yield 12 discrimination conditions: initial frequency value for F_2 (2525, 2272, or 2068 Hz), slope direction (positive or negative), and duration (110 or 165 ms). F_0 and F_1 were both steady-state. Stimuli with steady-state F_2 s were compared to those in which final frequency varied. Five listeners were tested under adaptive tracking using a modified 2AFC task with feedback to determine ΔF , the threshold for frequency extent (the difference in frequency between the initial and final F_2 values). Analysis showed that initial frequency, slope direction, and duration all contributed to significant differences in ΔF . Results indicated that listeners attended to differences in frequency extent (Hz), not formant slope (Hz/s). Formant extent thresholds were 3 to 18 times smaller than extents measured in the natural speech tokens. [Work supported by NIHDCD-02229.]

9:00

5aSC4. Production and perception of English vowels. Byunggon Yang (English Dept., Dongeui Univ., 24 Kayadong, Pusan 614-714, South Korea, bgyang@hyomin.donguei.ac.kr)

This study examined the relation between production of nine English vowels and perception of the synthesized vowels by 14 American male and female speakers. Fant's bandwidth equations were employed to dynamically tune to the varied formant frequency values. A remarkable r^2 value was obtained from the regression analysis between the center formant frequency values at which the subjects perceived the same vowel quality from the discriminatory test and the formant values of the synthesis models. Males and females perceived the synthesized vowels in the same way with converging center formant values and similar ranges of the same vowel quality. There was a strong link between production and perception of male and female speakers. The average r^2 value was very high, which suggests a very lawful relation between production and perception. From the individual analyses we found that listeners adjusted the criteria for vowel discrimination in relation to their own vocal tracts. This result suggests that speaker normalization can be pursued by finding an individual regression equation between the reference and observed formant patterns. Also, the present data proved that human listeners possess a very accurate perceptual mechanism to extract invariant features from complex acoustic stimuli.

9:15

5aSC5. The obligatory use of vowel duration in speaking rate normalization. Kathleen M. Measer and James R. Sawusch (Dept. of Psych., SUNY at Buffalo, Buffalo, NY 14260, kmmeaser@acsu.buffalo.edu)

Speech segment duration varies with speaking rate and phonetic identity. This study explored how listeners normalize for speaking rate when variation in segmental duration was a result of phonetic identity. Listeners identified the initial consonant in series that ranged from "beat" to "wheat" and "bead" to "weed." Even though the endpoints of the series

were spoken at the same speaking rate, the vowel duration in "bead"–"weed" was longer than "beat"–"wheat." Luce and Charles-Luce [J. Acoust. Soc. Am. 78, 1949–1957 (1985)] showed that vowel duration is an acoustic correlate of final stop voicing. Even though the vowel duration difference was an acoustic correlate of the final consonant voicing, listeners responded to the series as if vowel duration variation were also a cue to speaking rate. Listeners gave more /b/ responses to the "bead"–"weed" series. These data are inconsistent with models of perception in which acoustic correlates are uniquely assigned to sources in speech production. Instead, speaking rate normalization appears to be an early, autonomous process and the perceptual utilization of duration information for multiple aspects of speech perception may be obligatory. [Work supported by NIDCD Grant No. R01DC00219 to SUNY at Buffalo.]

9:30–9:45 Break

9:45

5aSC6. On idiolectal differences in speaking rate: A comparison of spontaneous and read speech. Robert E. Remez, Janelle Barnes (Dept. of Psych., Barnard College, 3009 Broadway, New York, NY 10027, remez@columbia.edu), Dalia Shoretz (Harvard Univ. Med. School, Boston, MA 02115), Maya Chatav, and Rebecca Piorowski (Barnard College, New York, NY 10027)

Individuals who share a dialect nonetheless differ consistently in phonetic expression. The origin of such idiolectal differences has been sought in sociolinguistically and in physiologically motivated accounts. Our study focused on characteristic differences in the habitual rate of speech production within a dialect, aiming to calibrate the contribution of neuromuscular constraint in idiolect differences. We sampled and analyzed the temporal characteristics of spontaneous and read speech produced by a set of adolescent female talkers from Brooklyn, New York. Spontaneous speech was elicited by open-ended questions, while the Farm Script was used as the text for production of read speech at comfortable and at rapid rates. Overall, the rate of rapid read speech was correlated with the rate of comfortable spontaneous speech and comfortable reading, indicating a clear contribution of neuromuscular constraints to characteristic differences in speech rate. However, some talkers exhibiting the ability to produce rapid speech at the fastest rates adopted comfortable speaking rates at the slower end of the distribution. In light of these findings, we doubt that idiolectal differences in speaking rate can be accountable solely by appeal to differences between individuals in agility of oral articulators. [Work supported by NIH DC00308.]

10:00

5aSC7. Effects of talker voice on lexical neighborhood and speech perception. Liza K. Zimack and James R. Sawusch (Dept. of Psych., SUNY at Buffalo, Buffalo, NY 14260, lkz2@acsu.buffalo.edu)

Previous research has shown that the number of words that a syllable is similar to (lexical neighborhood) influences the perception of the syllable and its phonemes. For example, in "bowth" "powth" and "bows" "pows" series, "bowth" is similar to more words than "powth," while "pows" is similar to more words than "bows." Listeners report more syllables beginning with /b/ in "bowth" "powth" and more that begin with /p/ in "bows" "pows." However, with one particular talker, no effects of lexical neighborhood were found. This research focused on two explanations for the lack of neighborhood effects for this talker. One possibility is that the previous method of determining lexical neighborhood is inaccurate for this talker. Alternatively, the time course for phonetic and lexical information processing may be different for this talker. In this research, listeners identified the final consonant in two nonword CVC series: "faesh" "faech" and "taesh" "taech." With a syllable final target phoneme, lexical neighborhood influences should build up prior to the processing of the target phoneme. Implications of the data for models of word recognition and perception with different talkers will be discussed. [Work supported by NIDCD Grant No. R01DC00219 to SUNY at Buffalo.]

10:15

5aSC8. Representational specificity of lexical form in the perception of spoken words. Conor T. McLennan (Lang. Percept. Lab., State Univ. of New York at Buffalo, Park Hall, Buffalo, NY 14260-4110, conor@deuro.fss.buffalo.edu), Paul A. Luce (State Univ. of New York at Buffalo, Buffalo, NY 14260-4110), and Jan Charles-Luce (State Univ. of New York at Buffalo, Buffalo, NY 14260-4110)

Exemplar and traditional abstractionist models of lexical representation were evaluated by examining flapping in American-English. According to abstractionist models, flaps are considered neutralized surface realizations of the underlying phonemic forms /t/ and /d/, with fully specified /t/ and /d/ forms of words represented in memory. Exemplar models, on the other hand, assume flapped tokens of spoken words have full representational status. An investigation of the representation of flaps in American-English was conducted using the repetition priming paradigm. Participants shadowed spoken words in two blocks of trials. Stimuli in the first block were primes and those in the second block targets. Primes and targets consisted of flapped and carefully articulated bisyllabic words. Time to shadow target words in the second block was measured as a function of prime type. The shadowing time results provide evidence for underlying abstract representations. The implications of these results for the theories of form-based lexical representations will be discussed.

10:30

5aSC9. Voice onset time of infant-directed speech in Korean. Soyoung Lee and Barbara L. Davis (Dept. of Commun. Sci. and Disord., Univ. of Texas, Austin, TX 78712)

Voice onset time (VOT) of infant-directed speech (IDS) was compared to adult-directed speech (ADS) at the sentence level in Korean. Previous research has not investigated VOT at the sentence level. Korean mothers were asked to describe pictures to their infants and to a Korean adult using subject-object-verb structure. Each picture showed an object noun and a verb containing stop consonants in prevocalic position. VOT of lenis,

fortis, and aspirated stop consonants for nouns and verbs in IDS and ADS was measured. In general, Korean mothers used two different strategies in IDS compared to ADS: longer VOT for aspirated stop consonants and shorter VOT for lenis/fortis stop consonants. These strategies may make a clear distinction between aspirated and lenis/fortis conditions. However, VOT of aspirated verbs was shorter than for nouns in IDS. When mothers show exaggerated pitch accent at the end of sentences in IDS, noun and verb phrases are likely to be combined together. Verbs are no longer at the phrase initial position, so VOT of aspirated verbs becomes shorter. These results indicate that Korean mothers do not consistently provide an acoustically prominent phonetic environment for important syntactic structures in IDS to their infants.

10:45

5aSC10. Perceptual assessment of intoxication level. Harry Hollien, Rebecca Hill (Inst. for Adv. Study of the Commun. Processes, Univ. of Florida, Gainesville, FL 32611), and Camilo A. Martin (Veteran's Administration Medical Ctr., Gainesville, FL 32608)

Data are reviewed and summarized from ten perceptual experiments wherein groups of listeners assessed the levels of intoxication for various types of talkers speaking at specific windows of sobriety and inebriation. The speakers included both sexes. Subjects were selected on the basis of a number of rigorous criteria; they produced several classes of speech when sober and at three highly controlled levels of intoxication. Listener groups included prescreened university students and professionals trained in appropriate areas. The task was to judge intoxication level from speech samples. It was found that speakers were judged as drunker than they actually were at mild levels of intoxication but that auditors sharply underestimated their involvement for severe intoxication. Data from three subprocedures led to the identification of a number of trends. However, no significant differences were found when the judgments by men and women were contrasted, nor were they significant between professionals and lay auditors or between the classes of speech produced. [Research supported by NIAAA.]

SATURDAY MORNING, 3 JUNE 2000

ENGLISH ROOM, 8:10 A.M. TO 12:30 P.M.

Session 5aUW

Underwater Acoustics: Scattering, Reverberation and Noise

Kevin D. LePage, Cochair

SACLANT Undersea Research Center, Vl. San Bartolomeo, La Spezia 19138, Italy

Nicholas C. Makris, Cochair

Department of Ocean Engineering, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, Massachusetts 02139

Chair's Introduction—8:10

Contributed Papers

8:15

5aUW1. Acoustic wave reflection from a rough surface over a random fluid half-space. Jin-Yuan Liu (Inst. of Undersea Technol., Natl. Sun Yat-sen Univ., Kaohsiung, Taiwan 804)

Plane-wave reflection from a rough surface overlying a fluid half-space with a sound-speed distribution subject to random perturbation is considered. A theory based upon a boundary perturbation method in conjunction with a formulation derived from Green's function for the coherent field in the random medium has been applied to a typical oceanic envi-

ronment to study its effects on the plane-wave reflection. By considering the coherent field itself, it may be obtained straightforwardly through a procedure consistent with the formalisms currently employed in rough surface scattering. The results show that both the rough surface and medium inhomogeneities may reduce the plane-wave reflection; however, the characteristics of the curves representing their effects are different, enabling us to identify the dominant scattering mechanism. The results for the coherent reflection due to individual scattering mechanism are compatible to those found in the existing literature. [Work supported by NSC of Taiwan, R.O.C.]

5a SAT. AM

5aUW2. A unified model for reverberation and submerged target scattering in shallow water. Purnima Ratilal and Nicholas C. Makris (MIT, 77 Massachusetts Ave., Cambridge, MA 02139)

A unified model for reverberation and submerged object scattering in a stratified medium is developed from a full-field theory for scattering in a waveguide. The advantage of the unified approach is that it enables consistent predictions to be made of the target-echo-to-reverberation ratio as a function of range and depth. A full-field waveguide scattering theory is especially necessary for shallow-water applications since simpler but approximate methods based on free-space scattering, such as the sonar equation, can be in error by many orders of magnitude depending on target shape, size relative to the wavelength, and environment. The model is applied to determine whether submerged targets can be detected above seafloor reverberation for a variety of water column and sediment types, sonar and target locations, and target shapes and sizes. The model is also applied to determine conditions necessary for discrete geomorphological features of the seafloor and sub-seafloor to be detected above diffuse seafloor reverberation.

5aUW3. Reverberation characteristics in shallow water with a strong thermocline. Ji-Xun Zhou, Xue-Zhen Zhang, and Guoliang Jin (School of Mech. Eng., Georgia Inst. of Technol., Atlanta, GA 30332-0405, jixun.zhou@me.gatech.edu)

The first joint China-U.S. ocean acoustics experiment was conducted in the summer of 1996 in a far offshore area in the Yellow Sea with strong thermoclines. The water depth was about 76 m with a flat seabed. The sediments in the area consisted of silty sand with mean grain diameter of 0.0643 mm. Shallow-water reverberation data were collected by using a 32-element vertical array that spanned the water column. Explosive sources of 1 kg were detonated both above and below the thermocline. The Yellow Sea '96 reverberation data exhibit strong receiver and source depth dependence. Reverberation intensity level (RL), received in the thermocline, shows a regular interference structure with increasing time (distance). This can be explained by the R-mode theory. In general, seabottom scattering dominated the Yellow Sea '96 reverberation process. Therefore, the seabottom scattering strength was derived, by inversion, from the reverberation and propagation data. A resonant volume-scattering source in the water column was also observed in the measurements. This demonstrates that water column effects (scattering from internal wave packets, fish groups, etc.) may also be significant for the shallow-water reverberation at some frequencies. [Work supported by ONR.]

5aUW4. Bistatic reverberation in range-independent waveguides. Kevin D. LePage (SACLANT Undersea Res. Ctr., VI. San Bartolomeo, 19138 La Spezia, Italy)

The basis for predictions of the temporal evolution of bistatic reverberation in range-independent waveguides is obtained under a modal framework using the narrow-band approximation. Approximate expressions for the field incident on the bottom at any given time are obtained as a function of modal group speed, dispersion, bandwidth, and azimuthal Fourier order. These functions, combined with the range-dependent spatial correlation functions of Gaussian distributed scatterers, which are also approximated as a function of Fourier order, provide the basis for the prediction of reverberation intensity and three-dimensional spatial coherence as a function of time after shot in multipath environments. The predicted characteristics of the reverberation are shown for several scenarios, emphasizing the differences caused by frequency, bandwidth, waveguide parameters, and receiver aperture characteristics.

5aUW5. Critical comparisons of the measured forward scattering from a clay bottom and a sand bottom. Tokuo Yamamoto and Hua Sun (Appl. Marine Phys. Div., RSMAS, Univ. of Miami, Miami, FL 33149)

Low grazing angle, medium frequency (3 to 15 kHz) forward scattering from a clay bottom off the coast of Gulf Port, Mississippi was measured using a 32-channel bilinear array and a broadband acoustic source. The clay bottom has a sharp velocity gradient with the sound speed at the interface slightly less than that of sound speed in water. The forward scattering was also measured at a sand-bottom site on the Florida Atlantic coastal shelf. The sound speed of the sand bed was roughly 1700 m/s. The measured bottom scattering-strength patterns are very similar both in magnitudes (-60 to -20 dB) and the distributions (roughly elliptical with the source and receiver being the foci). The apparent similarity, however, results from two different physical mechanisms. The waveform and time analyses show that acoustic energy penetrates deep into the clay bottom and scattered within the sediment volume, whereas acoustic energy is scattered at the surface of the sand bottom. In the clay bed, the sound waves are upward refracted due to the velocity gradient, showing a caustic-like phase shift by 0 to 90 deg. [Work supported by ONR Ocean Acoustics Program.]

5aUW6. Multiple Doppler shifts in the field scattered from an object moving in a horizontally stratified waveguide. Yisan Lai and Nicholas C. Makris (MIT, 77 Massachusetts Ave., Cambridge, MA 02139)

A method is developed to solve the problem of bistatic scattering from a horizontally moving object in a horizontally stratified waveguide. The target motion induces both frequency and wave-number shifts in the scattered field. For long-range propagation in a waveguide, the incident field can be decomposed into a discrete set of plane waves that propagates at the equivalent angles of the trapped modes. Doppler effects in the scattered field depend on the direction of the incident and scattered field with respect to the direction of motion. Each incident mode contributes a number of Doppler-shifted scattered components. For a monochromatic source that excites N incident modes in the waveguide, there will be roughly N^2 distinct harmonic components in the scattered field. The effect of these multiple Doppler shifts on the scattered field will be illustrated through a number of canonical examples.

5aUW7. Multiple scattering effects produced by a cylindrical shell in shallow water. Angie Sarkissian (Naval Res. Lab., Code 7132, 4555 Overlook Ave., Washington, DC 20375-5350, angie@aquanrl.navy.mil)

Multiple scattering effects produced by a finite ribbed cylindrical shell placed in a shallow-water environment are evaluated for the monostatic geometry. A combination of finite elements/infinite elements methods and the method of superposition is applied to compute first-order multiple scattering effects produced by a cylinder placed close to the shallow-water surface. Results are shown for a cylinder having a length to a diameter ratio of 6.8 at a $ka=10$, where k is the wave number in the fluid at the location of the cylinder and a is the radius of the cylinder. [Work supported by ONR.]

10:15

5aUW8. Visualization of the energy flow in and around a fluid-loaded elastic cylindrical shell. Cleon E. Dean and James P. Braselton (Phys. Dept., P.O. Box 8031, Georgia Southern Univ., Statesboro, GA 30460-8031, cdean@gasou.edu)

The energy flow in and around a fluid-loaded elastic cylindrical shell can be visualized by the instantaneous elastodynamic Poynting vector field and the time-averaged complex elastodynamic Poynting vector field. The instantaneous field can be used to animate and follow the energy flow around the cylindrical shell. The real and imaginary parts of the complex elastodynamic Poynting vector field can be used to show the traveling and standing wave components of the energy field, respectively. Special attention is given to the fluid/elastic boundary and energy coupling between the solid and fluid media. Numerous vector field plots and several animations will be shown.

10:30

5aUW9. Laboratory measurements of low-frequency acoustic scattering from a buried canonical target. Harry J. Simpson, Brian H. Houston (Naval Res. Lab., Washington, DC 20375-5350), and Carl K. Frederickson (Univ. of Central Arkansas, Conway, AR)

Finding and identifying proud and buried objects in littoral waters is of growing interest. Current acoustic-based systems can readily detect objects sitting proud; however, identification is considerably more difficult. For a buried target, even detection may prove difficult especially using higher-frequency systems. These higher-frequency systems are limited in range and sound penetration, which makes imaging buried targets extremely difficult. The research presented here is concerned with the use of relatively low-frequency acoustics (1 to 20 kHz) to overcome transmission and penetration limitations. This research includes quantifying the acoustical interaction with a buried target for ensonification angles above and below the critical angle. The target used for these measurements was an aluminum cylinder that was 20 cm long and 30 cm in diameter. A right-angle cone was machined in one end of the cylinder to enhance backscattering. Free-field measurements and completely buried measurements have been conducted. Of particular interest is the scattering from the buried canonical target for ensonification angles below the critical angle. These measurements are used to quantify the evanescent field interactions and the roughened interface enhanced scattering from the buried canonical target. [Work supported by ONR.]

10:45

5aUW10. Universal features of all-orders scattering in Biot theory. Eric Smith (Appl. Res. Labs., Univ. of Texas, Austin, TX 78713)

Collective effects of multiple scattering in a fluid-saturated porous medium are treated self-consistently within the context of Biot theory. Like other effective-medium theories, Biot's can be sensibly perturbed with random heterogeneity in any of the modulus or density parameters, with arbitrary covariances. When the perturbations are Gaussian and white, they can largely be absorbed back into the mean effective-medium parameters at long wavelength, by a successive iteration based on pertur-

bation theory. This iteration begins with wave numbers at the grain scale, and ends with wave numbers near those of the acoustic waves of interest. Remarkably, a large range of arbitrary perturbations can be shown in this way to focus their collective scattering on the slow compressional wave, only mildly refracting the fast wave. Inertial and modular effects, and their implications for acoustic localization, will be discussed.

11:00

5aUW11. High-frequency scattering from heterogeneous marine sediments. Anatoliy N. Ivakin (Andreev Acoust. Inst., Shvernika 4, Moscow 117036, Russia, aniva@glasnet.ru)

A model is presented for high-frequency acoustic scattering from marine sediments with continuous and discrete heterogeneities. The process of penetration into the sediment is described in terms of a combined wave-ray approach which gives a simple expression for the sound intensity within sediment taking into account main contributions of both ray and lateral components of the full-field solution. The process of volume scattering is described in terms of the scattering cross section per unit volume defined for continuous and discrete types of sediment heterogeneity. In the case of a continuous medium, it is defined through the spatial spectra of fluctuations of acoustical parameters. In the case of discrete heterogeneity, the scattering cross section per unit volume is expressed through the size distribution and individual scattering functions of discrete scatterers. Existing data on sediment heterogeneity are analyzed to provide inputs for the seabed scattering model. Angular-frequency dependencies of the backscattering strength are calculated and comparison with available experimental results is presented and discussed.

11:15

5aUW12. Influences of temperature and pressure on the material and acoustic scattering properties of zooplankton. Dezhong Chu (Dept. of Appl. Ocean Phys. and Eng., Woods Hole Oceanogr. Inst., Woods Hole, MA 02543, dchu@whoi.edu), Peter H. Wiebe, Timothy K. Stanton, Terence R. Hammar, Kenneth W. Doherty, Jack Zhang, Benjamin D. Redeer (Woods Hole Oceanogr. Inst., Woods Hole, MA 02543), Mark C. Benfield (Louisiana State Univ., Baton Rouge, LA 70803), and Nancy J. Copley (Woods Hole Oceanogr. Inst., Woods Hole, MA 02543)

It is well known that small variations (a few percent) in sound speed and density contrasts of the fluidlike weakly scattering zooplankton result in a significant variation in the target strength of zooplankton and ultimately in acoustically derived estimates of zooplankton biomass. Since pressure and temperature often vary considerably as a function of depth, the material properties of zooplankton are expected to change accordingly. To accurately determine zooplankton biomass from the acoustic scattering data, knowledge of the relationship between their material properties and these oceanographic parameters is required. *In situ* measurements of sound-speed contrast of the zooplankton species, *Calanus finmarchicus*, were made with the Acoustic Properties Of zooplankton (APOP) system in Wilkinson Basin (Gulf of Maine) in August 1999. The sound speed of this animal in the mixed layer above the thermocline (0–18 m) was significantly smaller than that right below the thermocline. Below the thermocline, temperature became nearly constant with depth so that only pressure was changing. The measured sound-speed contrast below the thermocline (60–180 m) increased slowly with increasing pressure. Mod-

els of the material properties of weakly scattering zooplankton that take into account the temperature and pressure dependence are proposed and numerical examples are presented. [Work supported by NSF, NOAA/NURP, and ONR.]

11:30

5aUW13. FDTD modeling of low-frequency sea surface reverberation in the presence of a near-surface bubble layer. Richard S. Keiffer, Robert A. Zingarelli (Naval Res. Lab., Stennis Space Center, MS 39529, keiffer@nrlssc.navy.mil), and Jorge C. Novarini (Planning Systems, Inc., Long Beach, MS 39560)

There exists a well-documented discrepancy between the low to moderate frequency ($150 \text{ Hz} < f < 1.5 \text{ kHz}$), low-grazing angle (< 30 degrees) reverberation-derived backscattering strengths collected at sea and the predictions provided by rough surface scattering models. One of the central features of the data/model comparisons in this regime is the strong wind speed dependence exhibited in the data and the very weak wind speed dependence predicted by surface scattering models. In a previous study [Keiffer *et al.*, *J. Acoust. Soc. Am.* **97**, 227–234 (1995)], a heuristic model was used to explore the hypothesis that bubble-induced refraction may modify the insonification of the air/sea interface and significantly enhance the surface reverberation in this frequency range. In the present work, a finite difference time domain (FDTD) solution to the linear acoustic wave equation is exercised in numerical experiments designed to unambiguously demonstrate the significant impact that near-surface, bubble-induced refraction can have on the reverberation time series and, therefore, on the surface scattering strengths.

11:45

5aUW14. A time-variant wedge assemblage method for acoustic Doppler backscattering from 2-D sea surfaces. Richard S. Keiffer (Naval Res. Lab., Stennis Space Center, MS 39529, keiffer@nrlssc.navy.mil) and Jorge C. Novarini (Planning Systems, Inc., Long Beach, MS 39560)

Recently, the details of a time domain model for the acoustic impulse response of static two-dimensional (2-D) sea surfaces based on wedge diffraction theory were reported [R. S. Keiffer and J. C. Novarini, *J. Acoust. Soc. Am.* **107**, 27–39 (2000)]. That approach, which is called the wedge assemblage (WA) method, was shown to accurately predict the acoustic backscattering from static, 2-D, fully developed ($U = 20 \text{ m/s}$) sea surfaces in the low to moderate frequency range. In this presentation, concepts from linear time-variant filter theory are used to extend that modeling capability to the problem of scattering from moving surfaces. Numerical simulations of the time-variant impulse response function have been computed for backscattering from (pressure-release) moving sinuso-

ids and standing waves. The results are shown as Doppler spectra over a broad band range of source frequencies (50 Hz to 2 kHz). Also presented are simulations of monostatic backscattering from time-evolving, fully developed seas as a function of azimuthal orientation. In the crosswind orientation, it is found that the Doppler shifts are substantially smaller than the first-order perturbation theory prediction and similar to the predictions from higher-order perturbation theory by M. Pourkaviani and J. F. Willemssen [*J. Acoust. Soc. Am.* **90**, 426–432 (1991)].

12:00

5aUW15. Sea-surface generated noise dependence on wind speed and white cap coverage. Douglas H. Cato (Defence Sci. and Technol. Organisation, P.O. Box 44, Pyrmont, NSW, Australia, doug.cato@dsto.defence.gov.au)

Breaking waves are an important source of noise in the ocean, and evidence from laboratory measurements indicates that the noise is generated by bubbles formed in wave breaking. Measurements at sea, however, have found that the noise correlates better with wind speed than any property of the surface waves. The relationship between noise intensity I and wind speed U is usually of the form $I = aU^n$ (where a is a constant), but values of n vary between studies over the range of 1 to 3. Measurements of noise near Australia show that, where other components of noise are low, the result is consistently $n = 3$ for a wide range of conditions. It is suggested that lower values of n result where non-wind-dependent noise contributes significantly to the levels at lower wind speeds, so decreasing the slope of the dependence on wind speed. Studies of white cap coverage have found a similar dependence on wind speed, with observed values of n being close to 3, thus providing a link between noise, wind speed, and wave breaking. Some preliminary results of noise measurements concurrent with measurements of white cap coverage to test this link are discussed.

12:15

5aUW16. Six-year records of low-frequency ambient sound in the North Pacific. Rex K. Andrew, Bruce M. Howe, and James A. Mercer (Appl. Phys. Lab., Univ. of Washington, 1013 NE 40th St., Seattle, WA 98105)

Spectra of ambient sound have been collected over the past 6 years at 13 locations around the North Pacific. Data were collected on single hydrophones for 3 min every 6 min and power spectra in 1-Hz bands from 0–500 Hz were calculated. Results from this 6-year data set are presented including probability statistics (e.g., the fraction of time a sound level is exceeded), whale, ship, and wind statistics, and long-term trends. These results build on those presented previously for 2 years [Curtis *et al.*, *J. Acoust. Soc. Am.* **106**, 3189–3200 (1999)]. [Work supported by Ocean Acoustics, ONR.]